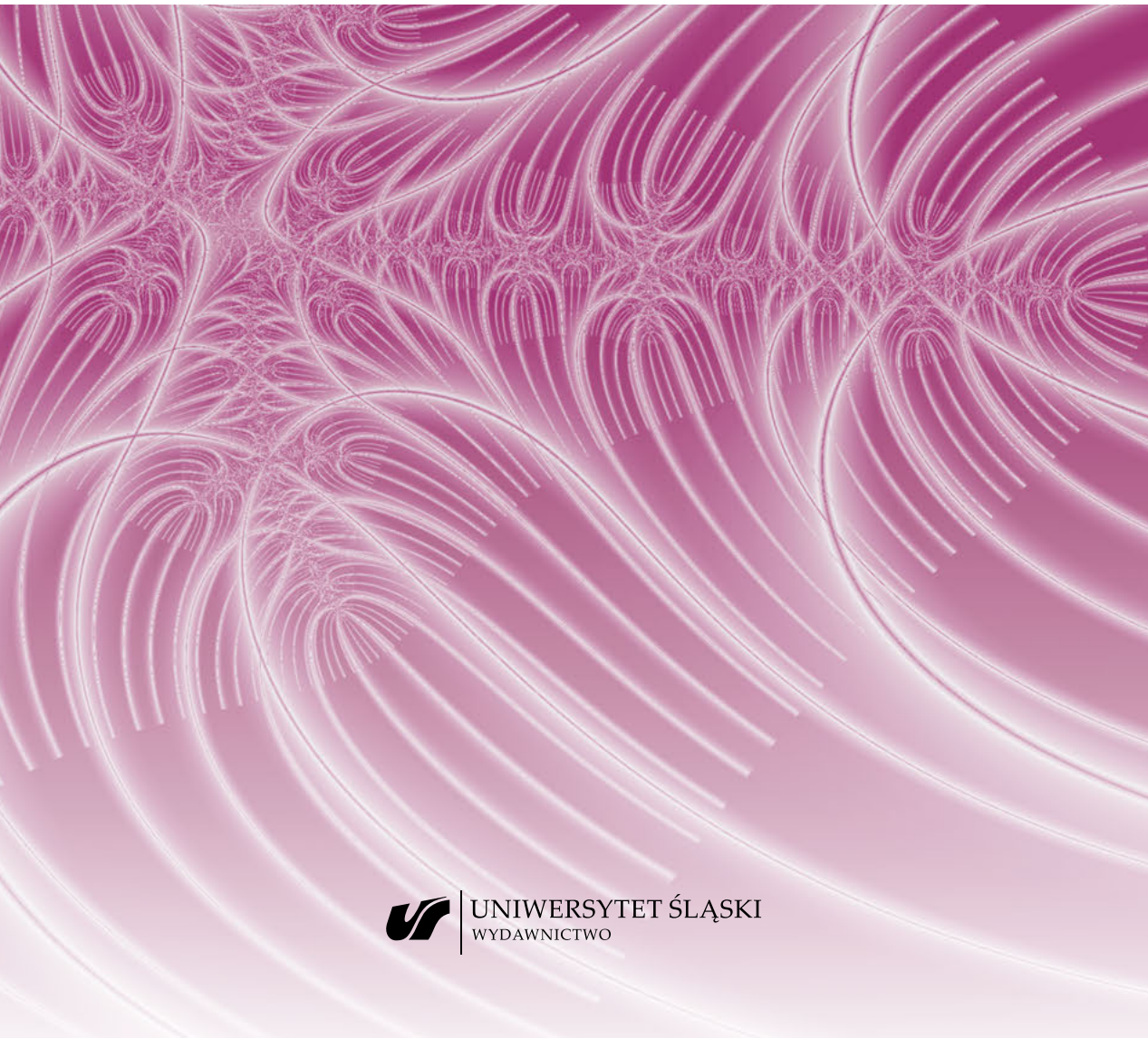
Theory and Practice of Second Language Acquisition

Vol. 8 (1), 2022



UNIWERSYTET ŚLĄSKI
WYDAWNICTWO

Theory and Practice of Second Language Acquisition

Vol. 8 (1), 2022

Wydawnictwo Uniwersytetu Śląskiego • Katowice 2022

This publication is indexed in the following databases:
Arianta, BASE (Bielefeld Academic Search Engine), BAZHUM, CEEOL, CEON, ERIH PLUS, GIGA, Index Copernicus, INFO NA, Nukat, MLA Directory of Periodicals, Most wiedzy, OAI-PMB Data Provider Registry, POLINDEX (PBN), Public Knowledge Project Index, Scopus, WorldCat.

Contents

Articles

Reviews

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 5–8
https://doi.org/10.31261/TAPSLA.12376

Preface

Theory and Practice of Second Language Acquisition was founded as a forum of discussion for Polish as well as foreign scholars and seems to have fulfilled its mission as a journal on the rise. The present volume marks the beginning of the eighth year of its presence in the scholarly world. The journal has become more and more popular as we get more and more interesting submissions from both Polish and foreign researchers. Indeed, since its foundation, every consecutive issue of the journal has welcomed contributions from many renowned researchers, such as Peter MacIntyre, David Singleton, Larissa Aronin, Sarah Mercer, Tammy Gregersen, and Jean-Marc Dewaele, to name just a few. Also, a fast growing number of OA uploads has been observed as an indication of the journal's popularity, as is the queue of the articles already accepted and awaiting their turn to be included in the next volumes to be published. This is why we have decided to increase the number of research papers published in a single volume for the second time: in the first years of the journal's existence there were six, last year seven, and starting with issue 8(1) *TAPSLA* is going to include eight research contributions, followed by two book reviews. It is the journal's ambition to demonstrate new trends and unknown venues for research in SLA, focusing both on theoretical discussion and the practical solutions to problems based upon them. We aim not only to publish and share with our readers contributions from well-known and respected scholars but would like to promote young researchers from all over the world, who often present fresh and innovative ideas or open up new perspectives on issues already researched. In other words, the journal hopes to become a venue for the exchange of ideas between well-established academics and those inspired by them. In terms of its content, the journal presents contributions on issues ranging from purely linguistic and cognitively-oriented research on language acquisition processes to psycho- and sociolinguistic studies, always trying to feature the most recent developments in terms of topic choice, as well as in

the methodology of research they employ. We publish our journal through an open access system, where the entire production process is executed online and the final product is available to everyone, thus offering an opportunity to share ideas through a broad, effective, and economical mode of dissemination. We aim at keeping high standards and quality, which are guaranteed by the international Editorial Board of *TAPSLA*, whose members are well-known Polish and foreign experts on a wide range of second language acquisition issues. The journal is indexed in numerous databases, including Scopus. As the journal is published by the prestigious Polish academic publisher, Wydawnictwo Uniwersytetu Śląskiego (University of Silesia Press), the copyediting and technical side of the production are done by an experienced team of editors at the University of Silesia Press. Updated information and all the issues published so far are available on the journal webpage at www.tapsla.us.edu.pl.

The topics of the present volume extend between focus on selected language skills influenced by a range of psycholinguistic variables, learning environments and instruction types, issues related to teacher wellbeing, and textbook analysis. The opening text "Language Teacher Wellbeing in the Workplace: Balancing Needs" by Sonja Babic, Sarah Mercer, Astrid Mairitsch, Johanna Gruber, and Kirsten Hempkin is the only one in the volume concentrating on teachers. Since it has been proven that teachers' wellbeing correlates positively with the quality of teaching and overall learners' success, the authors decided to investigate the factors which contribute to the construct. The study is especially valuable for its diversified and multinational perspective and inclusion of not only psychological and individual variables, but also influences of wider educational and sociological environments. The second paper, by Xavier Martin-Rubió and Irati Diert-Boté, titled "Catalan Law and Business Students in Italy: The Impact of a Stay Abroad on Fluency and Accuracy," shifts the perspective towards the learners. The author applies a qualitative perspective and investigates the factors responsible for diverse values of fluency and accuracy measures among three subjects participating in the study, in an attempt to provide some useful tips for both teachers and learners on how to make the most of an organized mobility abroad. The third text in the present volume, "The Impact of Individual Interest and Proficiency on Self-efficacy Beliefs in Foreign Language Listening" by Hyang-Il Kim, presents a dynamic picture of how individual interest is linked to basic and advanced skill self-efficacy against the background of listening activities aimed to expand learners' language proficiency. Statistical tests applied by the author reveal the links which were largely uninvestigated so far and offer very useful recommendations for both language teachers and learners. The topic of self-efficacy is further expanded by Bogusława Maria Gosiewska-Turek in her paper "Dyslexia, Self-efficacy, and Language Instruction in Foreign Language Learning—A Mixed Quantitative-qualitative Study," this time bringing into the picture the variables of the language-related

impairment of dyslexia and the style of language instruction. The results show that properly organized teaching can offer at least partial compensation for the inefficiencies stemming from neuropsychological limitations of the learners. The next paper in the present volume, by Agnieszka Ślęzak-Świat, titled "Development of Digital Literacy—Translanguaging and Transmedia Note Taking Formats for Academic Reading," can be best described as a diagnostic study attempting to identify the major difficulties which the students may have with efficient note taking. The author focuses on two selected note taking techniques—translanguaging and transmedia formats—revealing the potential areas for improvement in this respect. The study shows that there is much room for raising learners' awareness and competence in dealing with summarizing notes and that there is a visible need for explicit instruction related to such skills. The sixth paper, "Investigating Effects of Integrated Reading and Writing Skills Instruction in Enhancing Students' Critical Thinking Skills in EFL Classroom," written by Ayalew Tilahun, Berhanu Simegn, and Mulugeta Teka, concentrates on the potential of a selected language instruction procedure on the development of learners' critical thinking abilities. The author attempted to challenge the traditional separation of language skills, which is still a common practice in Ethiopia, in order to find support for a modification of teaching practices. The results of the carefully constructed quasi-experimental study reported in the paper demonstrate that there is a statistically significant positive influence of skill integration on learners' critical thinking, which in turn is very likely to translate into better achievement in foreign language development. The seventh contribution to the present volume, "Turkish EFL Learners' Acquisition of Psych Verbs and Unaccusative Verbs: A Replication Study on Underpassivization and Overpassivization" by Seray Tanyer and Samet Deniz, reports on an investigation of a peculiar aspect of grammar among Turkish learners of English. The authors wanted to find out whether processability constraints influence the acquisition of two types of verbs, focusing on the passive rather than active learners' skills. A statistically significant difference between unaccusative and psych verbs was noted, but no clearly interpretable correlations were found. The last research paper in the present volume, also placed in the Turkish context, is the study by Tan Arda Gedik and Yağmur Su Kolsal titled "A Corpus-based Analysis of High School English Textbooks and English University Entrance Exams in Turkey." This time, however, the focus is shifted towards the potential incompatibility between the content of English national level university entrance exams and the content of high school textbooks which should prepare the student for them. The authors employ corpus linguistics tools to demonstrate the inefficiency of the textbooks which are commonly used in Turkey.

Traditionally, the volume concludes with two book reviews. The first one, contributed by Pilar Safont, evaluates *Twelve Lectures on Multilingualism*

edited by David Singleton and Larissa Aronin in 2019. The reviewer speaks very highly of the up-to-date and much needed content, pointing to the fact that we now live in a multilingual world where multilingual contacts have become an everyday practice and a norm. Given the state-of-the-art relevance of the topics covered in the volume and the excellent choice of contributors—the top names in the research on multilingualism—the reviewer describes the book as "a must-read" in any graduate or postgraduate course on languages in contact, multilingual education or third language acquisition." The second review, by Caterina Hauser, presents a book focusing on another very important face of the present times in language education: *New Technological Applications for Foreign and Second Language Learning and Teaching* (2020), edited by Mariusz Kruk and Mark Peterson. The reviewer applauds the relevance of the topics covered in the book, pointing to the necessity and inevitability of efficient exploitation of technological development in the area of language instruction and learning. The younger brothers of CALL—Virtual Reality, social media and even chatbots—are more and more often successfully introduced into the classroom by language teachers and eagerly utilized by language learners, so a volume focusing on their applicability is undoubtedly another must-read. The reviewer stresses yet another advantage of the book: its coverage of not only practical applications of new devices and software, but also of ethical considerations related to personal data protection, privacy and potential threats stemming from their use.

The papers published in the present issue, although offering mainly empirical accounts, do not neglect solid theoretical foundations and overviews. They encompass a very broad range of individual researcher's empirical work in varied teaching contexts, presented from a variety of perspectives. We hope that this innovative and creative research, especially in the context of its potential for practical applications, will be of interest to other scholars. The practical solutions to problems proposed by some of the authors can be adapted for many other teaching and learning contexts. We hope that all types of readers—researchers, teachers and students—will find the articles not only useful but also inspirational. More than anything else, we would like to thank all the authors in this volume and, as is our usual practice, to extend our invitation to all Polish and foreign researchers and academics to share their work with us by submitting it to our journal.

https://orcid.org/0000-0003-0626-0703 *Danuta Gabryś-Barker*
https://orcid.org/0000-0003-0308-4337 *Adam Wojtaszek*

Articles

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 11–34
https://doi.org/10.31261/TAPSLA.11514

Sonja Babic
University of Graz, Austria
https://orcid.org/0000-0003-1004-3090

Sarah Mercer
University of Graz, Austria
https://orcid.org/0000-0003-2558-8149

Astrid Mairitsch
University of Graz, Austria
https://orcid.org/0000-0002-9885-3399

Johanna Gruber
University of Graz, Austria
https://orcid.org/0000-0003-3355-7748

Kirsten Hempkin
Univerza v Mariboru, Slovenia
https://orcid.org/0000-0003-4035-0665

Language Teacher Wellbeing in the Workplace: Balancing Needs

Abstract

Teachers who experience high wellbeing in their workplace teach more effectively, have better relationships with learners, and high attainment among their learners (Mason, 2017). To understand what contributes to language teacher wellbeing, we examined the three pillars of positive psychology (Seligman, 2011) and drew in particular on work in Positive Organizational Scholarship (Cameron & Spreitzer, 2012) to explore institutional and personal factors which teachers perceived as influential for their wellbeing. The paper reports on insights from 15 language teachers in 13 different countries. This sampling technique ensured a diverse set of perspectives on this topic. Data were gathered through in-depth, semi-structured interviews which were analyzed using Thematic Analysis (Braun & Clarke, 2006). The analysis revealed five main themes the teachers perceived as relevant for their wellbeing including workplace culture, social relationships, sense of meaning and purpose, language teacher status, and physical wellbeing. The findings highlight that wellbeing is not just a personal and subjective phenomenon, but it is also collectively and socially determined. The study concludes with a reflection on implications for practice, policy makers, and school leaders as well as a consideration of issues of individuality to address in future research.

Keywords: language teacher wellbeing, workplace, positive organizational scholarship, job satisfaction

Introduction

Wellbeing has become a topic of increased global interest (Calvo et al., 2012; Cho, 2014), particularly during the global pandemic in 2020 (MacIntyre et al., 2020; Zacher & Rudolph, 2020). In several countries, it has even formed the basis of economic policy. For example, New Zealand has adopted the Happiness Index, India has introduced the Ease of Living measurement in 2019, and Bhutan has utilized the Gross National Happiness (GNH) since 2008. In a survey conducted in the UK during the 2020 lockdown, a YouGov poll found that eight out of ten people in the UK would want the government to highlight health and wellbeing more (Harvey, 2020). This growing awareness of the importance of wellbeing generally has also been witnessed within education. For example, the OECD (2019) and PISA (2015) have included measurements of learner wellbeing since 2015, and countries such as Finland (Government of Finland, 2018), and the Republic of Ireland (Government of Ireland, 2019) have incorporated wellbeing initiatives into key education policies.

Language teaching is no exception to this trend of growing interest in wellbeing. Most notably, there has been a recent increase in studies focusing on the wellbeing of language teachers within language education (e.g., Hiver & Dörnyei, 2017; Mercer & Gregersen, 2020; Moskowitz & Dewaele, 2019; Ončevska Ager & Mercer, 2019) including papers published during the COVID-19 pandemic (e.g., MacIntyre et al., 2020).

The language teaching profession is often plagued by high levels of stress, growing burnout rates, and unfavorable work conditions, which can consequently lead to language teachers across the globe leaving the profession (Swanson, 2008). It seems that sustaining energy, commitment, motivation, and maintaining optimal levels of wellbeing over time has become increasingly challenging for teachers across the globe (Day & Gu, 2010; Sulis et al., in preparation). Therefore, it is now more important than ever to gain a more comprehensive understanding of the factors that influence teachers' sense of wellbeing in order to better understand how institutions and educational systems can best support teachers in their professional roles and help them thrive.

In this study, we explore the language teachers' perspective on factors which positively influenced and supported their wellbeing as language professionals. We placed a particular focus on the perceived contribution of organizational and school contextual factors given that most research to date has tended to emphasize the individual and personal traits affecting wellbeing with less consideration of context-specific characteristics. Indeed, MacIntyre highlighted that, apart from positive experiences (and emotions) and positive character traits, "positive institutions have been the least well studied" (2016, p. 5). Therefore, given that "wellbeing emerges from a blend of personal and

professional factors as well as contextual factors, in particular, our perception of our environment" (Mercer & Gregersen, 2020, p. 3), we accounted for personal but focused particularly on institutional factors that seem to contribute to the wellbeing of language teachers and promote the quality retention of these language professionals.

Literature Review

What Is Wellbeing and Why Is It Important?

To understand wellbeing, Seligman's (2011) PERMA framework has been widely used. It encompasses five wellbeing components: positive emotions, engagement, relationships, meaning, and accomplishments. The PERMA model stems from Positive Psychology (PP) which suggests there are three pillars of wellbeing: positive experiences, positive individual traits, and positive institutions. The positive-experiences pillar is concerned with one's satisfaction with the past, happiness in the present, and hope for the future. Positive individual traits refer to characteristics that typically lead to a fulfilling life, such as courage, perseverance, and the capacity to love. Finally, positive institutions deal with the systemic designs and strengths that are integral in developing and maintaining flourishing communities (Seligman & Csikszentmihalyi, 2000). These suggest that when investigating wellbeing, it is important to understand all three strands and the interplay between these, as wellbeing emerges from the dynamic interaction of personal and contextual components in people's lives.

Wellbeing is not only "valuable because it feels good, but also […] because it has beneficial consequences" (Diener & Seligman, 2004, p. 1). Studies have shown that higher levels of wellbeing are associated with increased levels of happiness and productivity (Krekel et al., 2019), higher income, social awareness, and connectedness (Diener & Seligman, 2004), better social equity (White, 2010), and physical health and longevity (Diener & Seligman, 2004; Lyubomirsky et al., 2005). Given the number of scientifically proven benefits, we believe that improving people's wellbeing should become key aim of researchers, policy makers, and other relevant stakeholders for all social groups, but especially teachers who are known to work in a highly stressful profession but who can also influence learner wellbeing.

Indeed, experiencing higher levels of wellbeing is not only beneficial for teachers but also for their students (Roffey, 2012). Teachers who enjoy high wellbeing teach more creatively (Bajorek et al., 2014), cultivate better relation-

ships in the classroom (DeVries & Zan, 1995), attain higher levels of achievement among learners (Briner & Dewberry, 2007), and have fewer discipline problems (Kern et al., 2014). Finally, physically and mentally healthy teachers can better "cope with the daily challenges of teaching languages" (Mercer & Gregersen, 2020, p. 1). In sum, wellbeing is a key factor contributing to good teaching practice.

Factors Affecting Teacher Wellbeing

Teacher wellbeing can be affected by numerous factors, including how the individual teachers feel and think about their work (Sulis et al., in preparation). Some of these factors encompass, for example, self-efficacy (Wyatt, 2018), resilience (Gu & Day, 2007), and optimism (e.g., Luthans, Youssef-Morgan, & Avolio, 2015). However, wellbeing is not only determined by personal characteristics and psychological states, it is also influenced by social and contextual factors including a number of institutional factors. For example, positive organizational factors enhancing teacher wellbeing typically include positive social relations with administrators, parents, and colleagues (Butt & Retallick, 2002), a supportive school climate (Day et al., 2007), and teachers identifying with the school's values (Skaalvik & Skaalvik, 2011). Organizational factors known to adversely affect teacher wellbeing include discipline issues (Skaalvik & Skaalvik 2011), excessive workload (Smithers & Robinson, 2008), potential for interpersonal conflicts (Skaalvik & Skaalvik 2007), role conflict and ambiguity (Travers & Cooper, 1996), lack of adequate resources and facilities (Aldrup et al.,2017), responsibility for evaluation (Kyriacou, 2001), and accountability demands (Rogers, 2012). Other threats to teacher wellbeing, in many countries, include negative media depictions of the teaching profession, which further lower already fragile teacher morale (McCallum & Price, 2010).

Understanding Language Teacher Wellbeing

Teaching has been identified as one of the most stressful occupations globally (Ingersoll, 2011). Record numbers of teachers in general (Borman & Dowling, 2008), and language teachers in particular, are choosing to leave the profession (Chang, 2009; Worth et al., 2017). Key factors foregrounded by Mason (2017) that contribute to language teacher attrition include heavy workload within foreign language teaching programs, detrimental legislation, a lack of institutional support, a lack of appreciation and respect, and low linguistic self-efficacy.

Language teachers face a number of challenges unique to language teaching (Piechurska-Kuciel, 2011). In terms of the psychological factors, for example, they typically experience a heightened emotional burden in their work due to the personally meaningful content and interpersonal relations in language classroom (Golombek & Doran, 2014). Language teaching and learning are "inherently emotional endeavours" (King & Ng, 2018, p. 141) and teaching languages thus requires increased emotional understanding from the teacher in comparison to other subjects (King & Ng, 2018). Another psychological challenge is foreign language anxiety (Horwitz, 1996), as some teachers may have to manage their own insecurities in the target language, while also dealing with the anxieties of their learners. In terms of social and contextual factors, languages have relatively low status compared to other school subjects and are undervalued within some schools and societies (Mason, 2017). Moreover, language teachers, especially those working in the private sector, may experience precarious working conditions such as income insecurity, zero-hour contracts, and an untenable workload (MacIntyre et al., 2019; Walsh, 2019; Wieczorek, 2016).

In this study, we seek to explore and better understand both the personal and contextual factors that contribute to language teacher wellbeing with an aim of drawing attention to individual and systemic strategies which could be employed to support teacher wellbeing. To do this, we draw on PP, which is a branch of psychology that looks at how and why people flourish in life (e.g., Seligman, 2011) and what makes life most worth living (Seligman & Csikszentmihalyi, 2000). PP is concerned with "positive experiences like happiness and engagement, positive traits like character strengths and talents, positive relationships like friendship and love, and the larger institutions like family and school that enable these" (Peterson et al., 2008, p. 20). Typically, work in PP has tended to focus largely on the first two pillars of wellbeing (positive experiences and individual traits) with notably less work examining the role and nature of positive institutions.

To ensure a thorough consideration of the final pillar as well as the other two dimensions of wellbeing, we take a deliberate Positive Organizational Scholarship (POS) approach (Cameron & Spreitzer, 2012). POS refers to the study of factors that positively influence workplace culture and thriving work conditions within organizations and institutions (Cameron & Caza, 2004). POS is a scientific, theory-based and rigorous investigation of the factors that positively influence workplace culture and thriving work conditions within organizations and institutions (Cameron & Caza, 2004). Furthermore, it focuses on "positive processes, on value transparency, and on extending the range of what constitutes a positive organizational outcome" (Caza & Caza, 2008, p. 21). To gain further knowledge into the school ecologies and aspects that facilitate flourishing of language professionals, as perceived by professionals themselves,

we employed POS principals in our research. In this way, we hope the study will ensure balance in the consideration of both individual as well as institutional factors affecting wellbeing.

Methodological Design

Research Questions

This paper aims to investigate language teacher perspectives on personal and institutional factors that they feel contribute to their wellbeing in relation to the workplace and seeks to answer the following research question:

What factors on a personal and institutional level do language teachers feel affect their wellbeing?

The study was designed to explore diversity and engage with a range of teachers in various settings in order to generate a broad understanding of the possible factors involved in teacher wellbeing across institutional and cultural contexts. As such, the sampling procedure was designed to be as widely encompassing as possible.

Procedure

Participants in this study were recruited via the researchers' social media networks, such as Twitter and Facebook, as well as through email contacts. Social media platforms were used to share the call for participation with the international audience. The main advantage of this sampling method was to expand the geographical scope of this project and not limit it to one specific country. We encouraged anyone who received or saw the call to forward it to their networks, which has created a snowball effect (Goodman, 1961). Given the positive psychology-informed approach, this study was designed to only examine those language professionals' perspectives who felt that their current workplaces support their wellbeing. This was also ethically a desirable approach to ensure people were able to focus on the aspects that benefited their wellbeing, without an undue focus on the detrimental factors. As such, the call for participation specifically looked for EFL teachers who felt that, on the whole, their institution is a positive place to work in terms of their wellbeing. We asked for participation from teachers who identified as follows:

- I am currently teaching EFL.
- On the whole, my institution is good for my wellbeing.

As a next step, all participants were sent a link to a short biodata questionnaire to contextualize their responses. This questionnaire comprised items about participants' demographics as well as questions about their teaching situation, for example, information about the institution they were working at, the type of teaching contract they had, and any additional work responsibilities they had to fulfill. As a next step, participants were asked for an interview at a time and date of their convenience. The consent sheet was also sent to the teachers before the interview to read and sign. It outlined the purpose of the study, what is expected from them, how data will be used and stored, and it assured anonymity and confidentiality. Due to the variety of geographical locations of the participants, all interviews were conducted online (through Skype, WeChat, WhatsApp or Facebook Messenger). The interviews were transcribed by the team members. During transcription, any real names and places were removed or changed to protect participants' identities. All the interviews were transcribed for content, including anything which contributed to meaning such as pauses, laughter, and sighs. The audio files were deleted immediately following transcription.

Context and Participants

In this study, 15 participants from 13 different countries volunteered to participate. These include Argentina, Indonesia, Ukraine, Turkey, Slovakia, Belarus, Hungary, Poland, Serbia, Nicaragua, China, two participants from Japan, and two from Slovenia. Fourteen participants were female and one was male. The participants were working in various educational institutions and four participants identified as language school owners as well as teachers. They were of different ages (between 30 and 50 years old) and in different career phases. Their teaching experience ranged between nine months and 26 years. The majority (n = 13) of the participants were working full time, whereas two teachers were employed part time. Nine teachers had permanent contracts, five temporary, and one participant was working on zero-hours contract.

Further details about participants' demographic and contextual information can be found in Table 1. All participants have been given a pseudonym to protect their identity.

Table 1

Participants' demographic and contextual information ordered alphabetically by their country of residency

Participant	Gender	Age	Country they are currently working in	Educational institution	Duration of current employment	Part time/ full time employment	Type of contract
Elena	Female	46	Argentina	Private language school (school owner)	17 years	Full time	Permanent
Darja	Female	n/a	Belarus	University or other tertiary institution	9 years	Part time	Temporary
Bao	Female	30	China	University or other tertiary institution	11 months	Full time	Permanent
Lili	Female	50	Hungary	Primary school, grammar school, and university or other tertiary institution	26 years	Full time	Permanent
Sari	Female	44	Indonesia	University or other tertiary institution	5 years	Full time	Permanent
Thomas	Male	34	Japan	University or other tertiary institution	9 months	Full time	Temporary
Niko	Female	47	Japan	University or other tertiary institution	9 years	Full time	Permanent
Abigail	Female	38	Nicaragua	Non-government organization	2 years	Part time	Temporary
Oliwia	Female	43	Poland	University or other tertiary institution	2 years	Full time	Temporary
Tatjana	Female	47	Serbia	Private language school (school owner), and university or other tertiary institution	21 years	Full time	Temporary
Mina	Female	39	Slovakia	Private language school (school owner)	10 years	Full time	Zero-hour contract
Ema	Female	38	Slovenia	University or other tertiary institution	6 years	Full time	Permanent
Hana	Female	45	Slovenia	University or other tertiary institution	18 years	Full time	Permanent
Ayla	Female	40	Turkey	University or other tertiary institution	15 years	Full time	Permanent
Oksana	Female	36	Ukraine	Private language school (school owner)	9 years	Full time	Permanent

Research Instruments

Fifteen in-depth, semi-structured interviews were conducted to encourage a conversational manner between the researcher and the participant, while, at the same time being guided by predetermined questions to create comparable content (Dörnyei, 2007). Semi-structured interviews also hold the advantage that participants are offered the chance to discuss topics they feel are important and which can deviate from the interviewer's questions (Dörnyei, 2007). These semi-structured interviews were based on a protocol which encompassed eleven sections covering the teachers' career trajectory, general information about their workplaces, physical space and resources, organizational structure, workplace culture, social relationships, autonomy, opportunities for professional development, work-life spillover, institution in its larger socio-cultural system, and overall statements about their general wellbeing. Each interview lasted between 45 minutes and one hour. A corpus of 108 836 words was generated.

Data Analysis

To analyze the data, we followed a Thematic Analysis approach (Braun & Clarke, 2006). To familiarize ourselves with the data, each researcher individually began the analysis by reading, commenting, and memoing on the transcripts. During this process, we met twice for discussions as a team to share ideas and generate initial code list. The interviews were coded in Atlas.ti. Three researchers coded four interviews each, and one coded three interviews. In the third group meeting, we compared and discussed the code lists, which we then combined into one set of codes (n = 38). This list included categories, such as social relationships, teacher autonomy, workplace culture, physical wellbeing, societal appreciation, pay, and continuing professional development. The interviews were revisited in light of this shared code list and any anomalies or exceptions were added and discussed with the whole team. Throughout the coding process, a joint notebook for memos was created, which was shared via Google Docs between the authors. Finally, in a fourth group meeting following the final wave of coding, the list of codes was grouped into themes which reflected the main issues across the data. These themes were workplace culture, relationships, sense of meaning and purpose, language teacher status and physical wellbeing, which served as the basis of the structure in the findings section.

Limitations

To recruit participants for the study, the call was shared on the authors' social media and via personal contacts. Naturally, this means volunteers stem from those who are active on social media and are possibly familiar with the work of the authors. This may suggest a bias to those with an interest in language teacher wellbeing. As such, they may be more aware and conscious of factors affecting wellbeing than other teachers, although there was no particular evidence of this in the data. In this study, participants self-identified as working in educational institutions that they felt contributed positively to their wellbeing. Typically, this perspective is subjective and the same workplace would perhaps be described differently by other employees in these institutions. However, it is the subjective perspective on the institutional factors which enhance their wellbeing which is at the heart of this study. Finally, we are aware that the paper could have gotten a different perspective had we included teachers who find their work conditions stressful and negative for their wellbeing. However, our research design encompassed and focused solely on teachers who identified their institutions as having a positive influence on their job satisfaction and overall wellbeing.

Findings

The findings section comprises five major themes that emerged from the data as affecting participants' wellbeing either positively or negatively: Workplace culture, social relationships, sense of meaning and purpose, language teacher status, and physical wellbeing.

Workplace Culture

Workplace culture typically "refers to the deep structure of organizations, which is rooted in the values, beliefs, attitudes, practices, norms, customs, and assumptions held by organizational members and that characterize a workplace environment" (Wentling & Thomas, 2009, p. 27). In this study, the most prominent aspects of the workplace culture mentioned by participants comprised the workplace atmosphere, support systems, autonomy, hierarchical structures, and the workplace as a physical space.

The majority of the participants (n = 14) described their workplace atmosphere as generally positive, open, friendly, and welcoming, which they felt

was important for their wellbeing. Thirteen participants particularly highlighted colleagues who are motivated, eager to learn, and help each other. For example, Thomas felt that staff members of the university were professional and positive and were, together, "rowing in the right direction." Niko, a teacher from Japan, is a particularly interesting example of how the workplace atmosphere can be influenced by the behavior and attitude of the staff members. She explained that, in the past, she had worked with male teachers who "had the power" and had created a toxic work environment for her; however, since these teachers retired, the work atmosphere in Niko's institution became pleasant and enjoyable—when "hard ones left the job at the university […], we created quite a better atmosphere."

Another feature of a positive work climate according to these teachers was when institutions encouraged staff members to be autonomous and to openly voice their beliefs. In particular, six participants stressed that the staff opinions "matter" to the institution and that they felt teachers were involved in making important decisions related to the institutional policies. For example, Darja mentioned that her institution valued her opinion, as well as the opinion of her colleagues and students. She explained that the dean arranges meetings once a month in which student representatives participate together with university teachers. Similarly, Thomas referred to the program development at his institution as a "bottom-up approach" where he could share his ideas and influence certain decision-making. However, Hana's portrayal of her institution was a unique example of a lack of open communication and ways in which it influences staff members:

> I think that [the dean] doesn't take into consideration that we are human beings, we are all academics and she cannot treat us like bags of something, push us around, right? People? I don't think they really want to express their opinions, some of them don't even dare.

She further mentioned that the hierarchical structure in her university has developed into a "talk-down management" in which the decision-making is only reserved for the dean. In contrast, other participants (n = 14), explained that the hierarchy in their respective institutions exists, but describe it as, "more of a flat structure" (Tatjana).

The physical space of the participants' workplace cultures also appeared to play a role in their wellbeing. Eleven participants mentioned that they were generally satisfied with the physical space, which they referred to in terms of green spaces, cleanliness, and sports areas, as well as resources, such as technology, teaching materials, comfortable furnishings and staff rooms with sufficient office equipment. For example, Bao explained how the physical space of her university affected her wellbeing positively:

> I feel the positive effect of the physical environment, for example, in my university […], there's a very beautiful small forest, it's very green and a lot of benches, so I like to bring my phone there and […] time spent there makes me refreshed.

Interestingly, four participants reported on poor work conditions in terms of physical space. Lili explained that her school "has not been renovated, maybe ever. […] It's in a very bad shape. Literally the walls are crumbling, the wires, the electric system, it sometimes just catches fire, so, it's really in very bad need of renovating." However, despite the negative aspects of the physical space, on balance, these participants still felt positive overall about their workplace, implying other factors played a more significant role.

Social Relationships

Relationships at the workplace were a prominent topic among all the participants as seen above in respect to workplace culture. Oksana explained that she enjoys working in her school because, "the team is wonderful, the people are very positive" and she added, "this is a positive place, we enjoy being here, together." In another example, Lili said she and her colleagues love spending time together and any issues that might arise they resolve as a team. Eight participants reported that they have formed true friendships with some of their colleagues. Abigail said that with two colleagues she became "friends outside the office." Having supportive relationships seems to have positively influenced the participants (n = 12) and appear to be crucial for their wellbeing.

However, three teachers explained that dysfunctional relationships with their colleagues can have a negative effect on their wellbeing. In particular, Mina said, "if the relationships in the workplace are not working well, that stresses me the most." Hana, for example, explained that the unity between teachers was lost and, consequently, the overall satisfaction among the colleagues significantly decreased. She recalled multiple incidents in which her colleagues showed a lack of collegiality. In this way, the quality of relationships between colleagues can either bolster or hinder their wellbeing.

Another important aspect for eleven participants was their relationship with their boss. The other four participants were school owners themselves and felt that they were creating a positive and supportive work climate for their employees. Ten participants referred to their bosses as being motivating (e.g., Ayla and Sari), open for discussion and suggestions (e.g., Lili and Niko), full of understanding (e.g., Abigail), and friendly and approachable (e.g., Ema and Niko). Bao explained that her boss was "very supportive" and "when you get something or achieve something, you can feel she is very happy for you, […],

this happiness is contagious, and you can feel like you are really supported." As mentioned in the previous section, only one participant, Hana, was dissatisfied with her superiors who she felt were mistreating their employees and did not value their opinions.

An equally important aspect of the workplace was participants' relationships with their students. Overall, all participants were especially proud of this relationship. They described it as friendly and trusting (e.g., Lili and Oksana), understanding and respectful (e.g., Abigail), caring (e.g., Elena and Bao), close (e.g., Sari and Oliwia), and humorous (e.g., Ema and Tatjana). Thomas depicted the reciprocal relationship between teachers and students in his institution:

> Most of the students are really, really passionate, we challenge them a lot but they kind of rise to the challenge each time, and I think that […] they feed a lot off the teachers' professional enthusiasm for the course, and then the teachers feed so much more off the students, so you have this kind of virtuous cycle of positivity.

Sense of Meaning and Purpose

Eleven teachers in this study specifically reported seeing a purpose and meaning in their jobs, especially in educating the next generation. For example, Elena explained:

> This is all for me. This is the idea. We work for education and for young people, especially for young people. We want to open young people's minds, to open their minds to fly […] and to be happy, to be full of dreams.

Building learners' knowledge and seeing them grow gave these eleven teachers sense of meaning. Tatjana said:

> What brings me satisfaction is when I see progress in students and when they are enjoying what they are doing here at school […], these things make you really satisfied and […] when I close the door at night and I go home, I really feel it was a good day.

Elena reported on a student who won a scholarship and thanked Elena for believing in her, "'I owe this especially to you because you believed in me and my dreams.' I started crying. Imagine. Because I say 'well, this was my aim when I opened this school.'" This filled Elena with pride and joy, knowing that she has made a difference in this student's life. Eleven participants in this study reported on feeling positive about the fact that they are making

a difference "leaving a mark on someone's life, a positive one" (Hana). In line with a PERMA perspective on wellbeing (Seligman, 2011), it can be seen that having a sense of meaning and purpose appeared fundamental for these teachers' wellbeing in the workplace and in life more broadly.

Language Teacher Status

In this study, we specifically asked participants to comment on their perception of the status of language teachers in their resident countries. Twelve teachers reported that the teaching profession in their respective countries is low and underappreciated. Ayla expressed that teachers in Turkey "are not respected" and felt that this was generally "the problem of education." She further added: "The education issue is the only thing that every person in my country can comment on. Everyone is an expert in that issue, everybody has an idea, but nobody asks the real teachers, the real workers of that sector."

Ayla believed that teachers are "not well paid"; a fact which was also highlighted by eight other participants. Lili said that, in Hungary, the media typically report that teachers get "a huge raise every year. I mean, that's what's in the media, but that's not true, but somehow they manage to communicate it in a way that half the country thinks that teachers are overpaid." In addition, Elena said that she had "suffered a lot" and that being underpaid caused other teachers in Argentina to become demotivated: "They don't want to do anything because they say there is no sense in anything because we are not well-paid."

Interestingly, only three participants, who came from Japan and China and were employed at universities, reported on high teacher status in their respective countries and institutions. One participant, Bao, explained that the status of teachers in China is very high, although this is not reflected in their salaries. Similarly, Niko said that

> [s]ocial appreciation of university faculty in Japan is quite high, so it is really appreciated, especially at top universities […]. So, I feel very much appreciated in social settings and money-wise, of course, I would like to get more, but I really have enough, and I appreciate what I have.

Two participants explained that although they are not content with their salaries, they see their work as meaningful beyond the financial issues. For example, Hana said that, "salaries are not enough, but we have other advantages" and Bao mentioned that "money is just an amount, one of the dimensions, that's it."

Physical Wellbeing

Five participants explicitly mentioned that the absence of physical health affected their work and overall wellbeing. For example, Niko explained that, in the past, when she lived far from her campus, she "wasn't able to manage to cook […] and it caused a lot of health problems." However, she moved "closer to the campus" and added: "I am cooking for myself, so I think I have a good balance." Hana mentioned "problems with [her] heart," and Oksana said that she suffered from back pain. Perhaps the most eye-opening example of how overall wellbeing can be influenced by physical health and contextual factors was portrayed by Abigail. She explained that she experienced burnout which she attributed to too many deadlines, her former boss who was unsupportive, frequent travelling, and a generally hectic lifestyle:

> We had a social political crisis in the country that burst in April and hasn't ended yet […]. I had no idea what was going to be, I just went every step, one after the other, and personally, I got a lot of physical effects. You know the symptoms, you know your stressors, and you know how it shows up in your body, but I had something new. By the end of the year, like a month before finishing everything I just couldn't, I almost couldn't get out of bed one day. I just couldn't move my arms anymore […] it was a most terrible experience I had, and I ended up going to the hospital.

Other participants (n = 7) showed an awareness of the importance of remaining physically healthy and reported on employing several strategies such as taking care of nutrition, spending time in nature, making sure they were physically active and doing sports and meditation. For example, Bao said: "Exercise is really helpful both for physical and mental health, especially when you feel tired, exhausted and you are just running." Thomas explained that he goes "for a little bike ride round the forest. […] This is just a very relaxing place to go." Finally, Oliwia said: "When I'm stressed, […] I do yoga and I do meditation and without it, I would be a different person."

Discussion

This study examined the perspectives of language teachers on the factors that seemed to have supported their wellbeing in relation to the workplace. In particular, we looked for EFL teachers across institutional contexts and countries, who felt that their workplaces were generally positive for their wellbeing.

Five main themes perceived as relevant for professional wellbeing dominated the data across settings and participants: Workplace culture, social relationships, sense of meaning and purpose, language teacher status, and physical wellbeing. Interestingly, all participants talked about the relevance of their lives beyond the workplace, including most notably physical health. This highlights the importance of considering teachers' lives from a holistic perspective when examining their wellbeing. Day and Gu (2010) refer to the 'blurry boundaries' between different teachers' life domains and highlight that wellbeing does not stem just from one isolated area of a person's life. Indeed, as the spillover-crossover model (Bakker & Demerouti, 2013) makes clear, there are permeable boundaries across an individual's life domains and experiences in each domain influence each other. Our sense of wellbeing emerges from the interaction of experiences in all areas although these are potentially weighted differently by an individual according to the relative importance of the respective domain for their overall sense of wellbeing (Robertson & Cooper, 2011).

In this study, the analysis showed that although the participants rated their workplaces as positive for their wellbeing (hence their participation), it did not mean these places were perceived as being perfect. Instead, the teachers flourished despite potential problems in the workplace. Dodge, Daly, Huyton, and Sanders (2012) suggest that wellbeing can be conceived of as a 'see-saw' with individuals balancing challenges to their wellbeing with resources available (e.g., personal, social, and material). This implies that teacher wellbeing can fluctuate according to the relative weighting of positive resources to draw on, balanced against negative factors and demands challenging their wellbeing. In other words, if the sense of positivity outweighs the significance for the individual of the negative factors, the overall sense of wellbeing can be balanced as positive without denying potential negative influences.

As was evinced in these data, certain factors are important for everyone, but some factors more or less, important for each individual. For example, Abigail's experience of burnout makes clear that when one's health is severely damaged by stress in the workplace, then no matter how significant other factors may be, nothing else really matters; wellbeing is at zero. In other words, her experience highlights that basic physical needs must be met first and foremost, otherwise other factors affecting wellbeing remain merely a luxury. From the perspective of the researchers in the field of positive psychology, it was heartening to see, in this specific instance, that the majority of the teachers in this study were aware of the importance of physical health and took conscious steps to protect and enhance their wellbeing in physical terms such as by attending to sleep, nutrition, and exercise.

The factor which was consistently emphasized as important by all participants was the workplace culture. As evinced in these data, this can be conveyed through communication structures within an institution, the degree of autonomy

given to staff, the level of personalization, the shared values of staff, and the attention to the care and quality of physical space among others (Fernet et al., 2014). All of these participants tended to see their workplace culture as positive with a friendly atmosphere and motivated colleagues. However, different facets of the workplace culture were rated differently by individuals. For example, Hana reported on the lack of democratic procedures and autonomy, whereas Lili was dissatisfied with the physical space. Yet, for these participants, despite these issues, they retained an overall sense of positive workplace culture. The research and theory of "person-organizational fit" (O'Reilly et al., 1991) can help us to understand this finding. This model suggests that the degree of wellbeing/stress experienced by an individual depends on how well they as an individual 'fit' with the conditions in that organization. This is important to appreciate as it emphasizes that it is not only the objective conditions that matter, but it is also their relative importance for an individual and how the two fit together that is defining. In other words, two people working under the same conditions can experience their wellbeing in that organization differently. It is the interaction of the personal, individual, and subjective experience of objective conditions that lead to an emergent sense of wellbeing.

Related to the workplace culture is the importance of social relationships in the workplace including relationships with colleagues, students, and director or principal. While this was critically important for all the participants, not all participants enjoyed positive relationships, most notably, Hana who felt a lack of unity and community across colleagues and an especially challenging relationship with leadership. Relationships have been shown to be critically important for teacher wellbeing (Kinman et al., 2011). For example, Hargreaves (2001) examined emotional dynamics between teachers and their colleagues. He conducted 53 interviews with elementary and secondary school teachers in Canada and found that teachers who work together, are in good relations, and support and value each other typically try to avoid conflicts and have more harmony in the workplace. Indeed, Morrison (2004) shows that having a friend at work can enhance one's satisfaction with the workplace. Another dominant theme across all the data was having a sense of purpose and meaning in one's work; however, this may be a double-edged sword for teacher wellbeing. It could imply that some teachers are willing to tolerate poor working conditions, including low remuneration (e.g., Hana and Bao), as they rated that aspect of their job on balance as less important for them than the positivity gained from helping learners to succeed. There are studies at all levels of education (primary, secondary, and tertiary) which show that educators' intrinsic motivation and dedication to the profession often means they work beyond what is expected of them (Ballet & Kelchtermans, 2009). This is potentially one contributory factor to the high levels of burnout which characterize the profession. Interestingly, this finding cautions against the value of so-called objective list theories of wellbeing that

refer to a "catalog of goods required for a well lived life" (Jayawickreme et al., 2012, pp. 329–330). What is needed for an individual to flourish in the workplace may vary considerably and it emerges from the interaction of personal perspectives on contextual factors. However, a vital word of caution is needed here. It is imperative that basic existential needs must be met, and teachers should receive just and fair remuneration and working conditions. Their intrinsic motivation and dedication to the profession must never be exploited by policy makers and/or owners or heads of schools. The issue of precarity in the ELT profession suggests that, sadly, this is not always the case (Walsh, 2019).

Finally, an interesting factor was the role of teacher status in the respective countries. This factor came out as the least positive for the majority of participants—only the teachers in China and Japan felt that their work and profession were esteemed but yet also not especially well remunerated. It is interesting to note that research has shown that professional status can affect teacher wellbeing (Troman, 2000), and yet, overall, these teachers retained a positive sense of wellbeing despite the low status assigned to their chosen professions. Notably, most of them complained about poor salaries which they felt did not reflect their qualifications and investment of time and energy. Again, on balance, for this group of teachers, although this situation is somewhat dispiriting, the other positive factors outweighed on balance, and they felt that they were still flourishing in spite of these societal and institutional conditions.

Conclusion

This study has examined the perspectives and experiences of language teachers who feel that their respective institutions contribute positively to their wellbeing. They were interviewed to explore the kinds of factors in their organizations and their own experiences which they felt supported their wellbeing. It was found that a number of key factors not just restricted to the workplace in isolation emerged and these were not always reported on in positive terms, despite the participants' overall conclusion that their situation in their workplace was positive for their wellbeing. This led us to reflect on the relative importance of various factors for individuals and the personal wellbeing needs of individuals where the positivity of certain factors could outweigh the negative effects of others.

The study has a number of implications for practice. Investing in the quality of staff relationships and opportunities for staff to connect socially would seem prudent given the strong emphasis throughout the data on a positive climate in the workplace. Despite the tolerance displayed by these specific individuals,

there is also a need to address the issues they raise about payment and the quality of physical space and resources provided. While they were merely a source of dissatisfaction for these participants, we know from research that for other individuals, they could contribute to them leaving the profession (Borman & Dowling, 2008; Karsenti & Collin, 2013; Kelchtermans, 2017).

Teachers are qualified professionals who typically exhibit great dedication to their chosen career, and their expertise and professionalism should receive the recognition in esteem and financial terms that they deserve throughout educational systems across the globe. Language teaching specifically has in part been characterized by precarity and poor working conditions (Mercer et al., 2016; Wieczorek, 2016) and the field must collectively consider how to counter these conditions for educators (Mercer, 2021).

With the experiences of online and remote teaching during COVID-19, there has been a further awakening and acknowledgement of the incredible work done by teachers and the centrality of their wellbeing for good practice (Arvisais et al., 2020; MacIntyre et al., 2020). It is hoped that this public and state recognition will remain even once the crisis has passed as will the emergent research agenda on teacher wellbeing. A concern raised by this study is that teachers' intrinsic motivation, commitment, and sense of purpose drawn from their work should not be exploited and utilized as an excuse explicitly or implicitly not to directly support teacher wellbeing and work conditions in systemic practical terms. Wellbeing is never the sole responsibility of the individual, but it is very much the responsibility of institutions and communities to provide the best conditions to enable each individual educator to thrive in their professional roles. When teachers flourish in the workplace, everyone benefits including the learners.

References

Aldrup, K., Klusmann, U., & Lüdtke, O. (2017). Does basic need satisfaction mediate the link between stress exposure and well-being? A diary study among beginning teachers. *Learning and Instruction, 50*, 21–30. https://doi.org/10.1016/j.learninstruc.2016.11.005

Arvisais, O., Deslandes Martineau, M., & Charland, P. (2020, April 27). A shout-out to teachers: Why their expertise matters in the coronavirus pandemic, and always. *The Conversation*. Retrieved from: https://theconversation.com/a-shout-out-to-teachers-why-their-expertise-matters-in-the-coronavirus-pandemic-and-always-136575

Bajorek, Z., Gulliford, J., & Taskila, T. (2014). Healthy teachers, higher marks? Establishing a link between teacher health and wellbeing, and student outcomes. *The Work Foundation*. Retrieved from: http://bit.ly/1rl9Wnm.

Bakker, A. B., & Demerouti, E. (2013). The spillover–crossover model. In J. Grzywacz & E. Demerouti (Eds.), *New frontiers in work and family research*. Psychology Press.

Ballet, K., & Kelchtermans, G. (2009). Struggling with workload: Primary teachers' experience of intensification. *Teaching and Teacher Education, 25*(8), 1150–1157. https://doi.org/10.1016/j.tate.2009.02.012

Borman, G. D., & Dowling, N. M. (2008). Teacher attrition and retention: A meta-analytic and narrative review of the research. *Review of Educational Research, 78*(3), 367–409. https://doi.org/10.3102/0034654308321455

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Briner, R., & Dewberry, C. (2007). *Staff wellbeing is key to school success: A research study into the links between staff wellbeing and school performance.* University of London.

Butt, R., & Retallick, J. (2002). Professional well-being and learning: A study of administrator-teacher workplace relationships. *Journal of Educational Enquiry*, *3*(1), 17–34.

Calvo, E., Rojas, M., & Martínez, I. (2012). Measurement, research and inclusion in public policy of subjective wellbeing: Latin America. *Foro Consultivo y Tecnológico*. Retrieved from: https://ssrn.com/abstract=2528528

Cameron, K. S., & Caza, A. (2004). Contributions to the discipline of positive organizational scholarship. *American Behavioral Scientist, 47*(6), 731–739. https://doi.org/10.1177%2F0002764203260207

Cameron, K. S., & Spreitzer, G. M. (2012). *The Oxford Handbook of Positive Organizational Scholarship.* Oxford University Press.

Caza, B., & Caza, A. (2008). Positive organizational scholarship: A critical theory perspective. *Journal of Management Inquiry, 17*(1), 21–33.

Chang, M.-L. (2009). An appraisal perspective of teacher burnout: Examining the emotional work of teachers. *Educational Psychology Review*, *21*, 193–218. https://doi.org/10.1007/s10648-009-9106-y

Cho, E. Y. N. (2014). Children's wellbeing in east and southeast Asia: A preliminary comparison. *Social Indicators Research, 123*, 183–201. https://doi.org/10.1007/s11205-014-0731-6

Day, C., & Gu, Q. (2010). *The new lives of teachers*. Routledge.

Day, C., Sammons, P., Stobard, G., Kington, A., & Gu, Q. (2007). *Teachers matter: Connecting work, lives and effectiveness*. Open University Press.

Diener, E. & Seligman, M. E. P. (2004). Beyond money: Toward an economy of well-being. *Psychological Science in the Public Interest, 5*(1), 1–31. https://doi.org/10.1111%2Fj.0963-7214.2004.00501001.x

DeVries, R., & Zan, B. (1995). Creating a constructivist classroom atmosphere. *Young Children, 51*(1), 4–13.

Dodge, R., Daly, A., Huyton, J., & Sanders, L. (2012). The challenge of defining wellbeing. *International Journal of Wellbeing, 2*(3), 222–235.

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Fernet, C., Lavigne, G. L., Vallerand, R. J., & Austin, S. (2014). Fired up with passion: Investigating how job autonomy and passion predict burnout at career start in teachers. *Work & Stress, 28*(3), 270–288. https://doi.org/10.1080/02678373.2014.935524

Golombek, P., & Doran, M. (2014). Unifying cognition, emotion and activity in language teacher and professional development. *Teaching and Teacher Education, 39,* 102–111. https://doi.org/10.1016/j.tate.2014.01.002

Goodman, L. A. (1961). Snowball sampling. *The Annals of Mathematical Statistics, 32*(1), 148–170.

Government of Finland. Finnish National Agency for Education. (2018). *Act on Early Childhood Education and Care*. Retrieved from https://www.finlex.fi/fi/laki/ajantasa/2018/20180540

Government of Ireland. (2019). *Wellbeing Policy Statement and Framework for Practice*. Retrieved from: https://www.education.ie/en/Publications/Policy-Reports/wellbeing-policy-statement-and-framework-for-practice-2018%E2%80%932023.pdf

Gu, Q., & Day, C. (2007). Teacher resilience: A necessary condition for effectiveness. *Teaching and Teacher Education, 23*(8), 1302–1316. https://doi.org/10.1016/j.tate.2006.06.006

Hargreaves, A. (2001). The emotional geographies of teachers' relations with colleagues. *International Journal of Educational Research, 35*(5), 503–527. https://doi.org/10.1016/S0883-0355(02)00006-X

Harvey, F. (2020). Britons want quality of life indicators to take priority over economy. *The Guardian, Business*. Retrieved from: https://www.theguardian.com/society/2020/may/10/britons-want-quality-of-life-indicators-priority-over-economy-coronavirus

Hiver, P., & Dörnyei, Z. (2017). Language teacher immunity: A double-edged sword. *Applied Linguistics, 38*(3), 405–423. https://doi.org/10.1093/applin/amv034

Horwitz, E. K. (1996). Even teachers get the blues: Recognizing and alleviating non-native teachers' feelings of foreign language anxiety. *Foreign Language Annals, 29*, 365–372. https://doi.org/10.1111/j.1944-9720.1996.tb01248.x

Ingersoll, R. (2011). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*, 499–534. https://doi.org/10.3102%2F00028312038003499

Jayawickreme, E., Forgeard M. J. C., & Seligman, M. E. P. (2012). The engine of well-being. *Review of General Psychology, 16*(4), 327–342. https://doi.org/10.1037%2Fa0027990

Karsenti, T., & Collin, S. (2013). Why are new teachers leaving the profession? Results of a Canada-wide survey. *Education, 3*(3), 141–149.

Kelchtermans, G. (2017). 'Should I stay or should I go?': Unpacking teacher attrition/retention as an educational issue. *Teachers and Teaching: Theory and Practice, 23*(8), 961–977. https://doi.org/10.1080/13540602.2017.1379793

Kern, M. L., Waters, L., Adler, A., & White, M. (2014). Assessing employee wellbeing in schools using a multifaceted approach: Associations with physical health, life satisfaction, and professional thriving. *Psychology, 5*, 500–513. http://dx.doi.org/10.4236/psych.2014.56060

King, J., & Ng, K.-Y. S. (2018). Teacher emotions and the emotional labour of second language teaching. In S. Mercer & A. Kostoulas (Eds.), *Language teacher psychology* (pp. 141–157). Multilingual Matters.

Kinman, G., Wray, S., & Strange, C. (2011). Emotional labour, burnout and job satisfaction in UK teachers: The role of workplace social support. *Educational Psychology, 31*(7), 843–856. https://doi.org/10.1080/01443410.2011.608650

Krekel, C., Ward, G., De Neve, J. M. (2019). Employee wellbeing, productivity, and firm performance. *Saïd Business School WP 2019-04*. Saïd Business School.

Kyriacou, C. (2001). Teacher Stress: Directions for future research, *Educational Review, 53*1, 27–35.

Luthans, F., Youssef-Morgan, C., & Avolio, B. J. (2015). *Psychological Capital and beyond*. Oxford University Press.

Lyubomirsky, S., King, L., & Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin, 131*(6), 803–855. https://doi.org/10.1037/0033-2909.131.6.803

MacIntyre, P., Gregersen, T., & Mercer, S. (2020). Language teachers' coping strategies during the Covid-19 conversion to online teaching: Correlations with stress, wellbeing and negative emotions. *System, 94*. https://doi.org/10.1016/j.system.2020.102352

MacIntyre, P., Ross, J., Talbot, K., Mercer, S., Gregersen, T., & Banga, C.-A. (2019). Stressors, personality and wellbeing among language teachers. *System*, *82*, 26–38. https://doi.org/10.1016/j.system.2019.02.013

Mason, S. (2017). Foreign language teacher attrition and retention research: A meta-analysis. *NECTFL Review, 80*, 47–68.

McCallum, F., & Price, D. (2010). Well teachers, well students. *Journal of Student Wellbeing, 4*(1), 19–34. https://doi.org/10.21913/JSW.v4i1.599

Mercer, S., Oberdorfer, P. & Saleem, M. (2016). Helping language teachers to thrive: Using positive psychology to promote teachers' professional well-being. In D. Gabryś-Barker & D. Gałajda (Eds.), *Positive Psychology perspectives on foreign language learning and teaching, Second language learning and teaching* (pp. 213–229). Springer.

Mercer, S. & Gregersen, T. (2020). *Teacher wellbeing*. Oxford University Press.

Mercer, S. (2021). An agenda for well-being in ELT: An ecological perspective. *English Language Teaching Journal*, 1–8. https://doi.org/10.1093/elt/ccaa062

Morrison, R. (2004). Informal relationships in the workplace: Association with job satisfaction, organizational commitment and turnover intentions. *New Zealand Journal of Psychology, 33*, 114–128.

Moskowitz, S., & Dewaele, J.-M. (2019). Is teacher happiness contagious? A study of the link between perceptions of language teacher happiness and student attitudes. *Innovation in Language Learning and Teaching*. Retrieved from: https://www.tandfonline.com/doi/full/10.1080/17501229.2019.1707205?scroll=top&needAccess=true

OECD (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing. https://doi.org/10.1787/38a34353-en

Ončevska Ager, E., & Mercer, S. (2019). Positivity for teacher wellbeing: A training course to help language teachers flourish. *The Teacher Trainer Journal, 33*(1), 20–22.

O'Reilly, C., Caldwell, D. F., & Chatman J. A. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *The Academy of Management Journal, 34*(3), 487–516. https://doi.org/10.5465/256404

Peterson, C., Park, N., & Sweeney, P. (2008). Group well-being: Morale from a positive psychology perspective. *Applied Psychology: An International Review, 57*, 19–36. https://doi.org/10.1111/j.1464-0597.2008.00352.x

Piechurska-Kuciel, E. (2011). Foreign language teacher burnout: A research proposal. In M. Pawlak (Ed.), *Extending the boundaries of research on second language learning and teaching* (pp. 211–223). Springer.

PISA (2015). *PISA 2015: Results in Focus*. Paris: OECD publishing. Retrieved from: https://www.oecd.org/pisa/pisa-2015-results-in-focus.pdf

Robertson, I., & Cooper, G. (2011). *Well-being. Productivity and happiness at work*. Palgrave MacMillan.

Roffey, S. (2012). Pupil wellbeing – teacher wellbeing: Two sides of the same coin? *Educational and Child Psychology, 29*(4), 8–17.

Rogers, B. (2012). *The essential guide to managing teacher stress*. Pearson.

Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. Simon & Schusters.

Seligman, M. E. P., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *American Psychologist, 55*, 5–14. https://doi.org/10.1037/0003-066X.55.1.5

Skaalvik, E. M., & Skaalvik, S. (2007). Dimensions of teacher self-efficacy and relations with strain factors, perceived collective teacher efficacy and burnout. *Journal of Educational Psychology, 99,* 611–625. https://doi.org/10.1037/0022-0663.99.3.611

Skaalvik, E. M., & Skaalvik, S. (2011). Teachers' feeling of belonging, exhaustion, and job satisfaction: The role of school goal structure and value consonance. *Anxiety, Stress, & Coping, 24*, 369–385. https://doi.org/10.1080/10615806.2010.544300

Smithers, A., & Robinson, P. (2008). *Good teacher training guide*. CEER.

Sulis, G., Mercer, S., Babic, S., Mairitsch, A., & Shin, S. (Eds.) (in preparation). *Language teacher wellbeing across the career lifespan*. Multilingual Matters.

Swanson, P. (2008). Efficacy and interest profile of foreign language teachers during a time of critical shortage. *NECTFL Review, 62*, 55–74.

Travers, C. J. & Cooper, C. L. (1996). *Teachers under pressure: Stress in the teaching profession*. Routledge.

Troman, G. (2000). Teacher stress in the low-trust society. *British Journal of Sociology of Education, 21*(3), 331–353. https://doi.org/10.1080/713655357

Walsh, P. (2019). Precarity. *ELT Journal, 73*(4), 459–462.

Wentling, R. M., & Thomas, S. (2009). Workplace culture that hinders and assists the career development of women in information technology. *Information Technology, Learning and Performance Journal, 25*(1), 25–42.

White, S. C. (2010). Analysing wellbeing: A framework for development practice. *Development in Practice, 20*(2), 158–172. https://doi.org/10.1080/09614520903564199

Wieczorek, A. (2016). High inhibitions and low self-esteem as factors contributing to foreign language teacher stress. In D. Gabrys-Barker & D. Galajda (Eds.), *Positive psychology perspectives on foreign language learning and teaching* (pp. 231–247). Springer.

Worth, J., De Lazzari, G., & Hillary, J. (2017). *Teacher retention and turnover research: Interim report*. NFER.

Wyatt, M. (2018). Language teachers' self-efficacy beliefs: An introduction. In S. Mercer & A. Kostoulas (Eds.), *Language teacher psychology* (pp. 122–140). Multilingual Matters.

Zacher, H., & Rudolph, C. W. (2020). Individual differences and changes in subjective wellbeing during the early stages of the COVID-19 pandemic. *American Psychologist*. Advance online publication. http://dx.doi.org/10.1037/amp0000702

Sonja Babic, Sarah Mercer, Astrid Mairitsch, Johanna Gruber, Kirsten Hempkin

Wohlbefinden der Sprachlehrer am Arbeitsplatz: Bedürfnisausgleich

Zusammenfassung

Lehrkräfte, die sich an ihrem Arbeitsplatz wohlfühlen, unterrichten effektiver, haben bessere Beziehungen zu Lernenden und erzielen höhere Leistungen bei ihren Studenten (Mason, 2017). Um zu verstehen, was einen Einfluss auf das Wohlbefinden der Sprachlehrer hat, wurden drei Säulen der positiven Psychologie (Seligman, 2011) untersucht. Dabei stützten sich die Autoren insbesondere auf das Konzept von Positive Organizational Scholarship (Cameron & Spreitzer, 2012), um institutionelle und persönliche Faktoren zu erforschen, die von den Lehrkräften als maßgebend für ihr Wohlbefinden wahrgenommen werden. In der Studie werden Erfahrungen von 15 Sprachlehrern aus 13 verschiedenen Ländern präsentiert. Die gewählte Stichprobenmethode bietet eine umfassende Betrachtung des Themas. Die Daten wurden in ausführlichen, halbstrukturierten Interviews erhoben und mittels der Thematischen Analyse (Braun & Clarke, 2006) ausgewertet. Die Analyse ergab fünf Hauptaspekte, die von den Lehrkräften als relevant für ihr Wohlbefinden erachtet werden, darunter Arbeitsplatzkultur, soziale Beziehungen, Sinnstiftung, Status als Sprachlehrer und körperliches Wohlbefinden.

Die Ergebnisse machen deutlich, dass Wohlbefinden nicht nur ein persönliches und subjektives Phänomen ist, sondern auch kollektiv und gesellschaftlich bestimmt wird. Die Studie schließt mit Überlegungen zu den Auswirkungen auf die Praxis, politische Entscheidungsträger und Schulleiter sowie zu den Fragen der Individualität, die in der künftigen Forschung zu berücksichtigen sind.

Schlüsselwörter: Wohlbefinden der Sprachlehrer, Arbeitsplatz, Positive Organisational Scholarship, Arbeitszufriedenheit

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 35–52
https://doi.org/10.31261/TAPSLA.10482

Xavier Martin-Rubió
Universitat de Lleida, Spain
https://orcid.org/0000-0002-9643-9530

Irati Diert-Boté
Universitat de Lleida, Spain
https://orcid.org/0000-0002-6964-5996

Catalan Law and Business Students in Italy: The Impact of a Stay Abroad on Fluency and Accuracy

Abstract

The growing number of English-Medium Instruction courses offered across Europe is increasing the opportunities for student exchanges. This study follows the progress of three students from Universitat de Lleida after their Erasmus experience at three different European universities, two in Milano and one in Macerata. The students took a monological English oral test before and after their stay abroad, and fluency and accuracy measures have been calculated from it. The students were also interviewed and participated in focus-group discussions. The measures from the two students who went to Milano show an improvement in their English level, whereas the student who went to Macerata performs even worse on his return. However, the experience from the two students in Milano was substantially different from the one who went to Macerata. Using the ethnographic information and the qualitative data available, we bring forward arguments that can help to account for these different outcomes.

Keywords: fluency, accuracy, study abroad, additional language learning, Englishization, ELF

Student Mobility in Europe

A growing number of universities across Europe offer English-medium bachelor degrees. According to Maiworm and Wächter (2014), the number of English-taught programs in Europe drastically increased between 2007 and 2014 in percentages that range between 112% for Central West Europe (the lowest increase) and 866% for South West Europe (the highest increase).

The advantage of this process of Englishization of Higher Education is that it removes linguistic barriers that made student exchanges more difficult in the past. The disadvantages include a domain loss for the local languages (Salö, 2014) and the obstacles for integration for incoming students and staff who might be considering a longer stay (Martin-Rubió & Cots, 2016). In this study, we follow the progress of three students from Universitat de Lleida (UdL henceforth), a higher education institution (HEI henceforth) situated in Catalonia, who spent one or two semesters at another European HEI. The three host HEIs are situated in Italy (Milano and Macerata). The students took an English grammar test and carried out two tasks, one written and one oral, before and after their stay abroad. They also participated in focus group discussions before and after their stay, and were interviewed during their exchange. Thus, the study focuses, on the one hand, on the analysis of the measures of accuracy and of the temporal measures of fluency of the oral task; and on the other, on the contextual circumstances of their stay, as self-reported in the group discussions and interviews.

Fluency and Accuracy Measurements

There is a long tradition in Second Language Acquisition to examine the evolution of fluency, accuracy and complexity in learners' production as a way to keep track of their learning process. For the present article, only fluency and accuracy are considered. Fluency has been defined as "the ability to talk with normal levels of continuity, rate and effort" (Starkweather, 1987, p. 12). When speech is constantly interrupted by pauses, continuity is affected; when only a small number of syllables are uttered in a given second, the rate is low; and when the speaker struggles to find the suitable word, an unusual amount of effort is required. Accuracy, on the other hand, has been described as "the ability to produce error-free language" (Czwenar, 2014, p. 82), although what can count as an error is rather debatable and a subject of controversy. Determining the levels of fluency and accuracy depends on, at least, two elements: the task to be developed and the level of language proficiency. Apart from these, the level of accuracy might be influenced by the socio-educational context of the speaker, and the level of fluency by the speaker's personality, such as self-confidence and extroversion (MacIntyre, Dörnyei, Clément, & Noels, 1998; Arnold, 1999), and by personal traits such as speech impairments, like stuttering (Fortunato-Tavares, Howell, Schwartz, & Furquim de Andrade, 2017). Nevertheless, for the present study, only the task at hand and the language proficiency are the foci of interest.

Different tasks present different demands on the people that engage in them. This can be easily gathered from a task comparison. By way of illustration, we could take the three following tasks: reading a text, describing a comic strip, and participating in a debate. While reading a text, the reader must only process the language written on the page; there is no need to improvise, even though the reader must still make sense of the words in the page. The person reading might produce pronunciation errors, but they would not be responsible for the lexical and grammatical elements in the text; the author of the text would be. When asked to describe a comic strip, one needs to choose what to say, and although the task itself narrows the vocabulary to be employed, one has a larger range of lexical choices than while reading. Furthermore, in a monological comic-strip description, the participant has to solve the task in a self-sufficient way. In debates with other people, conversely, one can pick and reuse lexical and grammatical elements from the other participants. Additionally, one needs to process the information conveyed by the other participants and decide what to say and how to say it, while simultaneously dealing with issues of taking, holding and yielding the floor. Any person, thus, will speak with different levels of fluency and accuracy across these different tasks.

Language learners at the early stages of their learning processes will evidently struggle to formulate long meaningful chunks, subsequently producing a higher number of pauses in their speech, a portion of which will be situated in the middle of phrases, thus giving the impression to the listeners that they are not fluent (Wennerstrom, 2000). As Chambers (1997, p. 540) points out, "becoming fluent therefore is […] about pausing less often and pausing at the appropriate junctures in an utterance." At a phonological level, Hieke (1985) argued that fluent speech equals connected speech. When the number of uninterrupted syllables is measured, what is actually being measured is how capable the speaker is of connecting syllables without pausing; this is a great indicator of the progress in the language. This is further exacerbated in English, considering it is a stress-timed language; therefore, the learner needs to learn to connect syllables with an alternation of stressed and unstressed syllables that is different in syllable-timed languages like Spanish (Leal, 1995).

Accuracy is intimately-related to the notion of error and mistake. According to Brown (2004, p. 216), "a mistake refers to a performance error in that it is a failure to utilize a known system correctly; while an error is a noticeable deviation from the adult grammar of a native speaker, reflecting the interlanguage competence of the learner." In this case, the adult grammar of a native speaker is taken as the norm, and noticeable deviations are considered errors, whereas performance errors are described as mistakes. In the same vein, Canagarajah (2015) argues that errors, unlike mistakes, are systematic and indicate the personal "grammar" of the language user. Errors are an integral part in the learning of an additional language, and one good indicator of progress

is precisely the learner's decrease in the number of errors. Errors can be classified as lexical (word choice), phonological (pronunciation), semantic (meaning), syntactic (grammar), and pragmatic (content) errors (Jiménez Arias, 2004, p. 177). A special focus is placed in this article on pronunciation errors. Figure 1 shows one of the two comic strips that our participants had to describe for their oral task. One of the participants used the expression "sequence" three times in the description, but pronounced it as /ˈse.kwens/ rather than the standard /ˈsiː.kwəns/ in all three occasions. It is, thus, not a mistake but a pronunciation error in the sense that the participant probably is unaware of the standard pronunciation of this word.

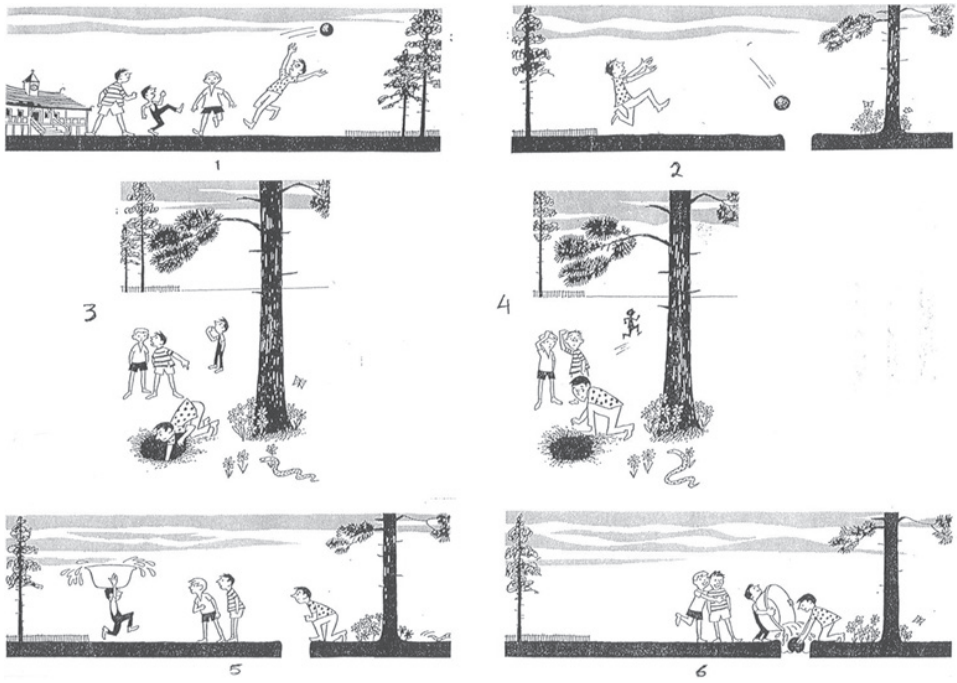


Figure 1. Comic 1: *Picnic task* (Heaton, 1966).

While describing the third picture of the strip, the participant says "the boys search the ball/ (0.6) /in.aːˈprɒk.siˌmeɪt.liˌesˈneɪk/." In the picture, we can see a boy looking for the ball inside the hole, and a snake approaching the hole, so we can imagine he wants to say something like "and a snake is approaching," but the fact is that the sentence he produces is ungrammatical and that we can only make sense of it all with the help of the picture. Focusing on pronunciation accuracy, if we decide he was producing the expressions "approximately" and "snake," we could identify two pronunciation mistakes. Rather than /əˈprɒk·sɪ·mət·li/, he generates a secondary stress in the fourth syllables (/aːˈprɒk.

si ˌmeɪt.li/); he also utters "snake" in three rather than two syllables. Moreover, he makes a lexical mistake, since he uses the adverb "approximately" rather that the verb "(to) approach," and a grammar mistake, because he is using the wrong order and structure. Needless to say, significant differences are bound to be found amongst different raters, which is why the two authors of this paper rated the cases independently and then discussed the different positions until a common ground was found.

Additional Language Learning at Home and Abroad

The status of English in international communication has progressively gained momentum and its impact upon many societies (including the Catalan and the Spanish) is much more far-reaching nowadays than a few years ago. The necessity to learn the English language is thus evident, being English one of the central subjects in the Catalan curriculum of both primary and secondary education. As an example, Catalan students are gradually obtaining higher results in the assessments of basic competences at the end of primary and secondary education (Generalitat de Catalunya, 2017); nevertheless, those evaluations do not assess oral production, which students believe to be the most problematic area for them (Diert-Boté, 2016).

In order to improve the competence in the English language during compulsory education, it is not uncommon for English learners to attend private language schools, to receive content subjects in English in many high schools—referred to as Content and Language Integrated Learning (CLIL)—and to go abroad, particularly during summer breaks. Spending some time abroad is frequently regarded as one of the most effective ways of learning the language of the country where the stay takes place. Although not as intensely as in other countries (Park, 1997), sending one's child(ren) to countries like the UK or Ireland for two to five weeks during the summer break is generally perceived in Catalonia as an effective way to boost the chances to learn English (Tragant et al., 2017).

When students start a university degree, HEIs normally offer similar options to help students to keep progressing in the English language. On the one hand, HEIs usually run language courses through their language centers addressed to two collectives: local students, who study additional languages (mostly—but not only—English); and international students, who can study the host university's local language(s). For instance, UdL offers Catalan and Spanish courses to international students, and there are several courses of English and other languages at different levels.

On the other hand, the equivalent to CLIL subjects at university are English for Specific Purposes subjects (ESP) and content subjects taught in English, commonly referred to as English Medium Instruction (EMI). Maiworm and Wächter (2014, p. 48) analyze the evolution of the number of programs taught in English (ETPs henceforth) in HEIs across Europe and argue that the Nordic Region and Central West Europe are "the pioneers in the implementations and running of ETPs"; the authors also state that although the growth rates in these regions are below the average, this is only because "further growth becomes more and more difficult" given their leading role in this process. In turn, "the growth rates were highest in South West Europe (866%)," although "most South European countries are still at the bottom of the ranking list—despite the impressive growth rates." Spain and Italy belong to this South West Europe region. Several studies (Cots, 2013; Dafouz & Camacho-Miñano, 2016; Doiz et al., 2014; Mancho-Barés & Arnó-Macià, 2017; Salaberri-Ramiro & Sánchez-Pérez, 2015) testify to the intensification of the Englishization process in Spanish HEIs. Particularly at UdL, the teaching in English in the academic year 2012–2013 accounted for 4.6%, whereas in the year 2016–2017 it rose to 6.1%. Notwithstanding the increase, the percentages of programs in English, and more specifically of students enrolled in such programs, are still very low.

Ultimately, the third option that HEIs offer is study abroad programs, most notably the Erasmus exchange program, which provides some students with the opportunity to spend one or two semesters at a foreign HEI. Unlike the summer-break stays discussed above, which normally take place in the UK or Ireland, university exchanges cover many different countries. The top destination for UdL students, for instance, is Italy (UdL, 2014). Research has shown that study abroad (SA henceforth) benefits second/foreign language acquisition, especially oral proficiency (Freed, 1990; Ginsberg & Miller, 2000). SA programs are naturalistic settings in which students can meet native speakers and thus participate in different communicative situations from those provided in educational settings (i.e., formal classroom learning) (Manchón & Murphy, 2002); yet, many of these students also enroll in language courses, so the combination of both settings seems to be the most appropriate one to learn a foreign language (Llanes & Serrano, 2011).

Although Manchón and Murphy (2002) mention contact with native speakers, on many occasions, and specially for the present study, native speakers of English are not what students (expect to) encounter in their SA experiences. Lecturers, students, and staff generally coexist in an environment in which several languages are employed, but where English is often used as the only available lingua franca. When UdL students decide to spend one or two semesters abroad, one of their goals is to improve their English in one of these English as a Lingua Franca (ELF henceforth) settings, which is a feasible goal according to research. Kaypak and Ortaçtepe (2014) found out that students

in an SA ELF setting showed a shift in focus from accuracy to intelligibility and a higher level of self-efficacy, which led them to engage more frequently in interaction in the target language. In turn, Borghetti and Beaven (2015) point out that in an ELF setting students not only experience lower levels of embarrassment, fear of being judged, and concern about participating in interaction; they also perceive a greater level of accommodation, negotiation, and cooperation strategies on the part of non-native speakers of English. Therefore, these non-Anglophone countries appear to be suitable places to practice English.

Koylu (2016) confirms these positive findings, adding the analysis of oral and written fluency, accuracy, and syntactic and lexical complexity to the picture. The participants in the study are Turkish students in three different contexts: at home internationalization (AH), stays in countries where the target language is the national language, in this case England (SA) and stays in countries where the target language is not the national language (ELFSA), in this case universities in Austria, Czech Republic, Denmark, Finland, Germany, Greece, the Netherlands, Italy, Poland, and Portugal. Koylu (2016, p. 173) identifies different accounts from students studying in countries such as Denmark, the Netherlands, Germany, and Finland in relation to those from students in Italy or Greece, with the different English level of the local residents as an important factor. Figures from the Eurobarometer (European Commission, 2012, p. 23) in relation to the percentage of people in the different EU countries who claim they can hold a conversation in English show a very stark contrast between countries like the Netherlands (90%) or Denmark (86%), and countries like Italy (34%) or Spain (22%). There are, however, many other factors, such as the length of stay, the teaching practices of the host HEI, the place of residence in the host city, or the network of friends sojourners establish, that also play a big role in the learning opportunities of the exchange students.

Research Questions and Methodological Aspects

The data used for this project come from a mixed methods research project which aimed to analyze the impact of a stay abroad on students from UdL. The study consisted of a quantitative part that targeted all the students from UdL selected to participate in the Erasmus program in the 2013–2014 academic year, and a qualitative part that focused only on those travelling to Denmark, Italy, and the United Kingdom. These three countries were selected because English plays very different roles in them. The 109 students who had been selected to participate in an Erasmus exchange during the 2013–2014 academic year were asked to complete a questionnaire, a grammar test, a written task, and an oral

task. Twenty-five of those 109 students completed all of it, and three of those 25 students had chosen an HEI in Italy for the SA. All the students provided their written consent to participate in the research project and the name of the three participants of this study has been replaced to preserve anonymity.

Beatriu, a Business student, went to Milano (HEI code IMILANO16) for ten months. She spent a month in Venice taking an Italian course before travelling to Milano. She was placed in a hall of residence on the outskirts of Milano, where a great number of Spanish Erasmus students resided, although she shared a room with a Slovakian student until February.

Alma, a Law student, went to a different university in Milano (IMILANO01) and stayed there for five months, from February 2014. She travelled to Milano with a friend from Lleida (that we will call Gertrudis), another Law student with whom she shared a room. They stayed in a different hall of residence, this one situated in the center of Milano. At first, they mingled with Italian students, but ended up hanging out with other mostly Spanish Erasmus students. They argued it was just easier to do that because they went to the same parties and had the same schedules, whereas many Italian students worked and partied less often. They had many German, Dutch, and Romanian students in the hall of residence with whom they often talked.

Finally, Josep Maria, another Law student, went to Macerata (IMACERATA01) for six months, although in this case from September 2013. Macerata is a much smaller place than Milano. His best friends in Macerata were two Italians, a girl from Córdoba (Spain) and a Greek student who had spent several years in Italy and with whom he always spoke in Italian.

The oral task mentioned above consisted of a comic strip of six different frames that the participants had to describe in less than five minutes after spending up to a minute preparing what to say. Two different comics were employed, comic 1 in the pre-tests (see Figure 1) and comic 2 in the post-tests (see Figure 2). A mistake while implementing the test to Josep Maria meant that he actually described comic 1 in both the pre- and post-tests. Given Josep Maria's performance in the post-test, one can only wonder what the result would have been if comic 2 had been used as originally planned.

The six oral stories produced were subjected to a two-phase manual analysis. The first phase analyzed temporal measurements of fluency and measurements of quantity, and the second included disfluency and accuracy measurements. The first phase started with the identification of three types of chunks: silent pauses, filled pauses, and between-pauses units (henceforth bp-units). Silent pauses are pauses of 0.25 seconds or longer, and filled pauses include hesitations, laughter or coughing. Bp-units are stretches of talk found between silent/filled pauses and are measured in syllables. A spreadsheet was used to indicate the type of chunk, the length in milliseconds for all chunks, the number of syllables, and the rates of each bp-unit.

Figure 2. Comic 2: *Football task* (Heaton, 1966).

Five measures of quantity have been employed. Speech Time (ST) results from adding the time of all the bp-units. Ginther et al. (2010, p. 387) define speech time as "speaking time, excluding silent and filled pauses." In turn, Pause Time (PT) is calculated by adding the time of all pauses (both silent and filled). Total Response Time (TRT) is the number of seconds the participant takes to complete the task and results from adding ST and PT. Speech Time Ratio (STR) gives us the percentage of time that the speaker spent speaking in relation to the TRT. The last measure of production is the number of meaningful syllables uttered.

As for fluency, three temporal measures have been used, two devised for this project and the third (MSR) adopted from Ginther et al. (2010). The first two measure the speech rate, that is, the speed at which the syllables are delivered: the Rate of Speech Time (ROST henceforth) results from dividing the total number of

syllables uttered by the ST; the Average Rate per bp-unit (ARbpu) results from calculating the speech rate of every bp-unit, and then establishing the average rate of all bp-units. The third measure is the Mean Syllables per Run (MSR), which results from dividing the total number of syllables by the number of runs (or bp-units).

For the second phase of the analysis, the bp-units in this spreadsheet were copied and pasted into another tab, and four new columns were added. Two disfluency and two accuracy measures were chosen: on the one hand, repetitions and false-starts/self-corrections (FS/SC) on the one hand (disfluency); and pronunciation, and lexical and grammatical errors on the other (accuracy). These measures were determined separately by the two authors of this paper, who then compared and discussed their results. We started from a number of rules and guiding criteria, but even so several minor issues emerged during the meetings. For example, a rule to measure repetitions was that all repeated consecutive syllables would be counted. In the case of the utterance "in the_ in the house," for instance, two repeated syllables were counted because the participant was undoubtedly repeating these syllables. However, in "ha_ had," the two syllables are different so it was agreed that this would be considered a false-start, as we do not surely know what the speaker intended to say in the first syllable. Pronunciation errors were particularly complex to determine. On several occasions, isolating the specific expression was what allowed us to come to a decision. The next two sections deal with the two different kinds of findings. An analysis of the results of the fluency and accuracy measures will be presented in Results.

Results

Table 1 contains the figures for the six stories produced by the participants before and after their stays in Italy. Josep Maria and Beatriu start from very low MSRs (3.88 and 3.15 respectively), but whereas Beatriu increases to 5.26 MSR (a 67% increase), Josep Maria actually decreases to 3.09 MSR (a 20% decrease). Alma's MSR increases 120%, from 5.45 to 12.18. Alma and Beatriu produce many more syllables in their post-stories (90% and 59% increases, respectively), whereas Josep Maria produces just one more syllable. Also remarkably different are the pausing behaviors. STR for Alma and Beatriu remain similar (small increases of 4.2% and 8.9%), whereas Josep Maria's STR decreases 33.2%. Josep Maria pauses more than he produces syllables, although we must indicate that there is a pause of more than eight seconds (chunk 62) that is partly responsible for this very unusual measure. Speech rates also increase for Alma and Beatriu, whereas Josep Maria's rate is slower in his post-story.

Table 1.

Quantity, temporal fluency, disfluency and accuracy measures

	Josep Maria		Alma		Beatriu	
	Pre-story	Post-story	Pre-story	Post-story	Pre-story	Post-story
Syllables	101	102	109	207	63	100
ST	37.08	40.68	39.43	54.75	24.70	37.54
PT	23.72	59.13	14.32	16.88	14.90	17.76
TRT	60.80	99.81	53.74	71.63	39.60	55.30
STR	60.98%	40.75%	73.36%	76.43%	62.37%	67.89%
ROST	2.72	2.51	2.76	3.78	2.55	2.66
ARbpu	2.58	2.46	2.63	3.66	2.44	2.65
MSR	3.88	3.09	5.45	12.18	3.15	5.26
Rep	9.90	6.86	0	0	1.59	7.00
FS/SC	2.97	3.92	0.92	0.97	3.17	6.00
Pr E	6.93	0.98	1.83	0	3.17	7.00
L&G E	10.89	17.65	6.42	5.80	14.29	10.00

Disfluency measures indicate very different patterns: Alma never repeats syllables and produces very few FS/SC; Beatriu clearly increases her repetitions and FS/SC in the post-story; and Josep Maria slightly decreases the number of repeated syllables per 100 syllables and slightly increases in FS/SCs. Combined with the fluency measures, we see that Alma improves a great deal, Josep Maria performs worse, and Beatriu is the one who improves in fluency measures but with more disfluency, which may be indicating that she is more willing to take risks.

In pronunciation accuracy terms, Josep Maria only produces one error in his post-story, but this is due to the very limited number of words he uses. He spends 43 of the 102 syllables he produces in his post-story with the words "and," "the," "boy(s)," and "(foot)ball." He even utters at least one word in Italian (*acqua*), and possibly three more (*per*, *te* and *i*). We are not sure about these last three because they are also words in Catalan. In his pre-story, he makes seven pronunciation errors, but three of them correspond to the way he pronounces "sequence," a word he does not use in the post-story. In Josep Maria's case, and due to a mistake in the protocol, he described the same comic on both occasions, so that a very unusual situation arises: bp09 of his pre-story and bp12 of his post-story are almost identical. In the first he says "in approximately snake" and in the second "and approximately the snake." We have counted two pronunciation errors in the pre-story (wrong stress in "approximately" and wrong vocalic sound before "s" in "snake") and one in the post-story (still wrong stress, but this time the vowel has a reason to be there because of the definite article "the" before "snake").

Beatriu produced more pronunciation errors in her post-story, but this is mostly because of the vocabulary in the two stories. Four of the seven pronunciation mistakes have to do with voiced final consonants, which she pronounces with total voicelessness ("dog"—three times and "bag"—once); two errors correspond to "mountain," which she pronounces without the diphthong in the first syllable; and the last corresponds to the wrong presence of an "l" sound in "walking." In the pre-story, in which she only uttered 63 syllables, there were no words used ending in final "g." In the pre-story she mispronounced "hole" in two different ways, in fact (first without the diphthong and then without final /l/). So we can actually see some slight improvement in general, and the increase in pronunciation errors could be due to the lack of chances to make the errors.

Finally, Alma makes no pronunciation errors in her post-story, and only two in the pre-story. Wrong pronunciation of "hole" (no final /l/) and wrong vocalic sound in "third." Alma produces seven lexical & grammatical errors in her pre-story and 12 in her post-story, but since her post-story is almost twice as long, there is actually an improvement in accuracy (6.42 to 5.80 errors per 100 syllables). Beatriu also improves (nine errors in her pre-story and ten in her post-story mean a 30% decrease in the number of errors per 100 syllables). Josep Maria produces 18 errors in his post-story, coming from 11 in the pre-story. Given both stories are just one syllable apart in length, accuracy worsens.

In the next section, details from the contextual circumstances of the three students' stay, as self-reported in the group discussions and interviews, will complement the fluency and accuracy analysis.

Insights from the Qualitative Data of the Study

The qualitative data for the Italy data sub-set are, on the one hand, focus group discussions, and, on the other, interviews. Table 2 provides details about the three group discussions. The reason why there are two post-focus group discussions is that the students who went to Macerata did so during the first semester, which means they were back from Italy in January already, so the focus group in June would have taken place too long after their return. The three students in the other focus group had just returned from Italy, although whereas Beatriu had spent a whole academic year in Milano, Alma and Gertrudis had been there during the second semester.

The interviews were conducted while the students were in the host HEIs. Two researchers travelled to Milano and Macerata, and conducted three interviews in Milano and one in Macerata. Alma and Gertrudis were interviewed

together, and the audio recording lasted 40' 47'' (INT_MI_1). Beatriu was interviewed separately, and it lasted 53' 39'' (INT_MI_2). Josep Maria was interviewed in Macerata, and the audio recording lasts 56' 58'' (INT_MA_1).

Table 2

Group discussions

CODE	Date	Participants	Time
pre-FG	18/06/2013	4 students (including Beatriu) + 2 researchers	123 minutes
post-FG1S	20/03/2014	3 students (including Josep Maria) + 2 researchers	53 minutes
post-FG2S	17/06/2014	Alma, Beatriu, and Gertrudis + 2 researchers	44 minutes

Three factors emerge from these groups discussions and interviews. A first factor concerns the starting self-confidence and proficiency level of the students. Beatriu could not apply for an Erasmus grant in the North of France, as she wanted to, because she did not have the required B1 level of English. She passed the test in June 2013, right before going to Italy. However, throughout the interview she expresses her lack of confidence when speaking English. Josep Maria also starts the stay without the B1 level, whereas Alma had a strong B1 and she is the one who benefits the most from the stay.

The second element is where they lived and the people they interacted with the most (and the languages in which they did that). Beatriu was accommodated in a hall of residence located on the outskirts of Milano, which was full of Spaniards (something she complains about). However, she shared a room with a girl from Slovakia, with whom she had to use English. Beatriu saw the need to move to another room to practice more Italian and English. In fact, the three participants complained about a great number of Spanish students in the program, and they realized that Italy was not the best place to practice English, at least for two reasons: the low English level of Italians and an excessive number of Spaniards. Gertrudis even said she practiced Spanish (her L1 is Catalan, after all). However, Alma (and Beatriu to a lesser extent) did improve their English, which might be due to the fact that they end up using more English than at home: Beatriu with her Slovakian room-mate, and Alma (and Gertrudis) with the German, Dutch, and Romanian people in their hall of residence.

The third element is the nature of the expectations the three students had in relation to their stay. The three students in post-FG1S indicate they expected a higher presence of English in general, and admit they would change their destination if they could go back. Josep Maria mentions the Czech Republic and Finland as places he would now choose, since at least there he would have an opportunity to practice English and to experience a truly different culture. However, learning Italian was also part of the plan. The lessons at the three

universities were conducted in Italian, and they had to take oral exams in that language, whereas non-Spanish Erasmus students took the lessons in English, as most did not have learning Italian amongst their goals. Alma, for example, improved her English but also learned Italian. In the post-focus group she claimed that she would even write "a bit of Italian" in her résumé. From Josep Maria's post-story, it is obvious he also learnt Italian.

Discussion and Concluding Remarks

Most research conducted in SA contexts seem to indicate that oral fluency is benefited from SA, whereas this does not seem to be the case for accuracy, where findings are less conclusive (Juan-Garau & Pérez-Vidal, 2007). Freed et al. (2004) compared the progress of 28 American students of French studying in three different learning contexts. The authors discovered that determining the amount of time students actually spent using the target language was essential in order to accurately interpret the findings. Apparently, the students in the immersion program, who were the ones with the greatest improvements, reported "devoting significantly more time to using French in out-of-class activities compared not only to the students in AH [at home] context but also to those in the SA context" (Freed et al., 2004, p. 294). In this line, Llanes et al. (2018) suggest that certain individual differences such as motivation and foreign language use in free time (among others) play a role in the language gains students will experience in the different contexts to the point that they can account for the different outcomes.

In our case, the two students who went to Milano learnt some Italian and simultaneously improved their English proficiency. Although their immersion in the Italian language was high, they also mixed enough with students from different countries and linguistic backgrounds, which may have made an impact on their progress in English. In contrast, the student who went to Macerata does not show an improvement in the language. The stay has certainly done little for his English, and his feeling at the end is that his destination was a place similar to Lleida (similar weather, food, culture, etc.), and although he has learnt some Italian, he would have probably benefited more from a stay in a country further North within Europe. The results highlight the many factors contributing to the context of stay abroad (which in turn affect the process of additional language learning) and resonate with findings from other studies like Borghetti and Beaven (2015), Kaypak and Ortaçtepe (2014), Kalocsai (2009) and Koylu (2016). These factors can be the role that English plays in the specific campus and in the town/city/country, the language repertoires of the network

of friends the students establish, and the attitudes which the different subjects adopt towards the different local languages.

Some limitations of the study must be pointed out. The oral tests were monological productions in which students had a minute to prepare themselves and then they improvised a short descriptive text. However, what they practiced the most during their stay were dialogical improvisations. In this sense, it would have been interesting to analyze the students' progress in dialogical productions in English. Another limitation of the study is that Josep Maria was given the same comic on his return, and even though several months had elapsed, this makes the comparison with the other two participants problematic. Certainly, it would have been more suitable to follow the same strategy for all the participants. Another issue is the small number of participants; this entails that the results must be taken with caution. A final limitation has to do with the monolingual bias of this study. The project as a whole aimed at identifying the impact of the stay on three main areas: language level, intercultural awareness, and feeling of Europeanness. Although the students participating in the qualitative part of the study were asked about their relationship with the local languages, it would have been very enriching to include tests in the local languages of their host HEIs to see their progress in those languages.

I believe this paper points towards at least two different lines for further research. The first would be new studies comparing pre- and post-oral productions of not just monological but also dialogical tasks from students engaged in different kinds of methods to improve their English. The second would be a more ethnographic approach to the many factors influencing additional language learning while studying abroad.

References

Arnold, J. (Ed.) (1999). *Affect in language learning*. Cambridge University Press.

Borghetti, C., & Beaven, A. (2015). Lingua francas and learning mobility: Reflections on students' attitudes and beliefs towards language learning and use. *International Journal of Applied Linguistics, 27*(1), 221–241.

Brown, H. D. (2004). *Principles of language learning and teaching* (4th ed.). Longman.

Canagarajah, S. (2015). Clarifying the relationship between translingual practice and L2 writing: Addressing learner identities. *Applied Linguistics Review, 6*(4), 415–440.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535–544 .

Cots, J. M. (2013). Introducing English-medium instruction at the University of Lleida, Spain: Intervention, beliefs and practices. In A. Doiz, D. Lasagabaster, & J. M. Sierra (Eds.), *English-medium instruction at universities: Global challenges* (pp. 106–130). Multilingual Matters.

Czwenar, I. (2014). Analysing spoken language for complexity, accuracy and fluency: Some methodological considerations. In W. Szubko-Sitarek, Ł. Salski, & P. Stalmaszczyk (Eds.), *Language learning, discourse and communication* (pp. 81–92). Springer.

Dafouz, E., & Camacho-Miñano, M. M. (2016). Exploring the impact of English-medium instruction on university student academic achievement: The case of accounting. *English for Specific Purposes, 44*, 57–67.

Diert-Boté, I. (2016). "I didn't learn English at school": An exploration of English language learners' beliefs and emotions." Unpublished Master's Thesis, Universitat de Lleida.

Doiz, A., Lasagabaster, D., & Sierra, J. M. (2014). Language friction and multilingual policies in higher education: The stakeholders' view. *Journal of Multilingual and Multicultural Development, 35*(4), 345–360.

European Commission (2012). Europeans and their languages. *Special Eurobarometer*, 386. http://ec.europa.eu/commfrontoffice/publicopinion/archives/ebs/ebs_243_en.pdf

Fortunato-Tavares, T., Howell, P., Schwartz, R. G., & Furquim de Andrade, C. R. (2017). Children who stutter exchange linguistic accuracy for processing speed in sentence comprehension. *Applied Psycholinguistics, 38*(2), 263–287.

Freed, B. F. (1990). Current realities and future prospects in foreign language acquisition research. In B. F. Freed (Ed.), *Foreign language acquisition research and the classroom* (pp. 3–27). DC Heath.

Freed, B. F., Segalowitz, M., & Dewey, D. P. (2004). Context of learning and second language fluency in French: Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition, 26*(2), 275–301.

Generalitat de Catalunya (2017). Avaluacions censals de competències bàsiques. Consell Superior d'Avaluació del Sistema Educatiu, July 6. Accessed 5 March 2018, http://premsa.gencat.cat/pres_fsvp/docs/2017/07/06/17/17/5365f81b-cb55-45ba-b015-a070b412178e.pdf

Ginsberg, R. B., & Miller, L. (2000). What do they do? Activities of students during study abroad. In R. D. Lambert & E. Shohamy (Eds.), *Language policy and pedagogy: Essays in honor of A. Ronald Walton* (pp. 237–260). John Benjamins.

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing,* 27(3), 379–399.

Heaton, J. B. (1966). *Composition through pictures*. Longman.

Hieke, A. E. (1985). A componential approach to oral fluency evaluation. *Modern Language Journal, 69*(2), 135–142.

Jiménez Arias, I. (2004). Treating students' errors in oral production. *Letras, 36*, 175–188.

Juan-Garau, M., & Pérez-Vidal, C. (2007). The effect of context and contact on oral performance in students who go on a stay abroad. *Vigo International Journal of Applied Linguistics, 4*, 117–134.

Kalocsai, K. (2009). Erasmus exchange students: A behind-the-scenes view into an ELF community of practice. *Apples – Journal of Applied Language Studies, 3*(1), 25–49.

Kaypak, E., & Ortaçtepe, D. (2014). Language learner belief and study abroad: A study on English as a lingua franca (ELF). *System, 42*, 355–367.

Koylu, Z. (2016). *The influence of context on L2 Development: The case of Turkish undergrtaduates at home and abroad*. (Unpublished doctoral dissertation). University of South Florida, USA.

Leal, C. F. (1995). A pause as a cause of change. *ERIC document ED* 379935. Educational Resources Information Center.

Llanes Baró, À., & Serrano Serrano, R. (2011). Length of stay and study abroad: Language gains in two versus three months abroad. *Revista Española de Lingüística Aplicada, 24*, 95–110.

Llanes, À., Tragant, E., & Serrano, R. (2018). Examining the role of learning context and individual differences in gains in L2 writing performance: The case of teenagers on an intensive study-abroad programme. *The Learning Language Journal, 46*(2), 201–216.

MacIntyre, P. D., Dörnyei, Z., Clément, R., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545–562.

Maiworm, F., & Wächter, B. (2014). The big picture. In B. Wächter & F. Maiworm (Eds.), *English-taught programmes in European higher education the state of play in 2014* (pp. 25–62). Lemmens.

Mancho-Barés, G., & Arnó-Macià, E. (2017). EMI lecturer training programmes and academic literacies: A critical insight from ESP. *ESP Today, 5*(2), 266–290.

Manchón, R., & Murphy, L. (2002). *An introduction to second language acquisition research*. UMU/Diego Marín.

Martin-Rubió, X., & Cots, J. M. (2016). Englishisation at a global space: Students and staff making sense of language choices. *Language and Intercultural Communication, 16*(3), 402–417.

Park, J.-K. (1997). 'English-fever' in South Korea: Its history and symptoms. *English Today, 25*(1), 50–57.

Salaberri-Ramiro, M. S., & Sánchez-Pérez, M. M. (2015). Analyzing writing in English-medium instruction at university. *Linguarum Arena. Revista de Estudos em Didática de Línguas da Universidade do Porto*, *6,* 45–58.

Salö, L. (2014). Language ideology and shifting representations of linguistic threats: A Bourdieusian re-reading of the conceptual history of domain loss in Sweden's field of language planning. In A. K. Hultgren, F. Gregersen, & J. Thøgersen (Eds.), *English in Nordic universities: Ideologies and practices* (pp. 83–110). John Benjamins.

Starkweather, C. W. (1987). *Fluency and stuttering*. Prentice Hall.

Tragant, E., Serrano, R. & Llanes, A. (2017). Learning English during the summer: A comparison of two domestic programs for pre-adolescents. *Language Teaching Research, 21*(5), 546–567.

Universitat de Lleida, 2014. *Memòria Acadèmica del curs 2013–2014.* Lleida: Edicions de la Universitat de Lleida. http://www.udl.cat/export/sites/universitat-lleida/ca/organs/secretaria/.galleries/docs/Memoria-Academica/Memoria_general_de_la_UdL_13_14.pdf

Wennerstrom, A. (2000). The role of intonation in second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 102–127). University of Michigan Press.

Xavier Martin-Rubió, Irati Diert-Boté

Katalanische Jura- und Wirtschaftsstudenten in Italien: Die Auswirkung eines Auslandsaufenthalts auf Sprachkompetenzen und Ausdrucksfähigkeit

Zusammenfassung

Die wachsende Zahl der in ganz Europa angebotenen englischsprachigen Lehrveranstaltungen verbessert die Möglichkeiten des Studentenaustauschs. In der vorliegenden Studie werden die Fortschritte von drei Studenten der Universitat de Lleida nach ihrem Erasmus-Aufenthalt an drei verschiedenen europäischen Universitäten – zwei in Mailand und

einer in Macerata – analysiert. Vor und nach ihrem Auslandsaufenthalt nahmen die Studierenden an einer monologischen mündlichen Englischprüfung teil, bei der ihre Sprachkompetenzen und Ausdrucksfähigkeit beurteilt wurden. Darüber hinaus wurden sie interviewt und beteiligten sich an Fokusgruppendiskussionen. Die Ergebnisse der beiden Studenten, die nach Mailand gegangen sind, zeigten eine Verbesserung ihrer Englischkenntnisse, während der Student, der nach Macerata gegangen ist, nach seiner Rückkehr noch schlechter abschnitt. Allerdings unterschieden sich die Erfahrungen der beiden Studenten aus Mailand erheblich von denen des Studenten aus Macerata. Anhand von ethnographischen Informationen und verfügbaren qualitativen Daten wird ein Versuch unternommen, die voneinander abweichenden Ergebnisse nachzuvollziehen.

Schlüsselwörter: Sprachkompetenzen, Ausdrucksfähigkeit, Auslandsstudium, Fremdsprachenlernen, Engischisierung, ELF

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 53–70
https://doi.org/10.31261/TAPSLA.11162

Hyang-Il Kim
Sahmyook University, South Korea
https://orcid.org/0000-0002-4340-0173

The Impact of Individual Interest and Proficiency on Self-efficacy Beliefs in Foreign Language Listening

Abstract

Interest and self-efficacy beliefs are important components of motivational constructs that share some common characteristics and influence learning. In this regard, several studies have explored how these two variables relate to each other in various domains and have offered meaningful implications. However, in the field of language learning, it appears that related research is relatively insufficient, and information offered through prior studies is very limited. Therefore, the current study aims to investigate how individual interest in learning English and proficiency influence self-efficacy in listening to expand the current knowledge. With a pre-/post-test design, it analyzed the data collected from 107 EFL Korean university students with a low-intermediate English proficiency, using hierarchical regression analyses. The results reveal that, unlike at the beginning of the course, individual interest turned to be a significant factor in the development of self-efficacy in advanced English listening around the end of the semester, even predicting it more than proficiency did. On the other hand, individual interest influenced self-efficacy in basic English listening much more than proficiency did in the pre-test and maintained this pattern in the post-test. This study provides insightful information into the relationship between interest, self-efficacy, and proficiency, highlighting the essential role of individual interest in learning English. It follows that the maintenance and nurturing of interest is crucial for the development of self-efficacy beliefs which in turn can contribute to advances in language learning.

Keywords: individual interest, listening, L2 proficiency, self-efficacy beliefs

Developing listening skills in a foreign language can be a difficult process that may lead students to experience frequent frustrations and negative feelings before attaining a level of satisfactory skill (Graham, 2011). It can be even more challenging for students with low proficiency (Wang & Fan, 2015).

Individuals who perceive a given task as difficult, recognize that they do not have the capabilities to perform such a task in a specific field or have repeatedly experienced failure are more likely to have relatively low motivation (Graham, 2007), or may not have much interest in learning a foreign language. Bandura (1997) maintains that an individual's motivation, affective states, and behavior tend to be formed based on their beliefs or self-perceptions rather than on their objective and actual state. This indicates that an individual's positive perceptions such as interest and self-efficacy beliefs are likely to play pivotal roles in academic performance and students may need urgent aid to facilitate them having positive experiences and thus build their positive perception toward language learning.

Interest and self-efficacy are two motivational constructs that come from one's judgments and beliefs accumulated overtime during past learning efforts, which are domain-specific (Hidi et al., 2002). Focusing on the similar characteristics and the influential roles in academic fields that the two variables have in common, a few studies offer some information about the relationship between these two constructs in various educational fields. However, it seems that the details and complexity of these relationships remain relatively little known in language learning circles. The study by Hidi et al. (2002) called for more research on this issue to understand its nature almost two decades ago, yet relevant information on the relationship between them still appears insufficient to result in useful implications. Recognizing the current gap, this study attempts to explore the complexity and interactive characteristics of the two motivational structures, along with the proficiency of Korean university students with a low-intermediate English proficiency level. More specifically, it aims to broaden the scope of information presented by the literature by exploring how students' interest and proficiency influence the formation of self-efficacy in English listening. This may allow us a better understanding of how these motivational variables are related.

Literature Review

Interest in Language Learning

Interest in language learning is considered one of the motivational components that help students stay engaged and committed during the learning process (Renniger & Hidi, 2002). A few researchers claim that interest involves an individual's cognitive and affective aspects (Ainley et al., 2002; Hidi & Renninger, 2006). More specifically, Hidi and Renninger (2006) claim

that the positive emotions that accompany engagement can be represented as the affective component of interest, while the perceptual and representational activities in relation to engagement are described as the cognitive counterpart of interest. It appears to be a general consensus that interest may strongly impact the cognitive performance and affective experience of an individual's learning path, leading to positive outcomes (Hidi et al., 2002).

Hidi (1990) differentiates interest into two types: individual interest and situational interest. She defines individual interest as an individual's persistent tendency to repeatedly engage in certain content over time. Since individual interest is related to the previous knowledge and experiences that an individual possesses, it is considered topic-specific and as having more potential for holding personal value for a relatively long period of time. On the other hand, situational interest is defined as the focused attention and affective response initiated while in a particular situation and by stimuli found in the immediate environment and is likely short-lived. This type of interest appears to be context-specific and without long-term value. Previous research indicates that situational interest is likely part of the fundamental constructs of individual interest as suggested in a developmental model of interest (Hidi & Renninger, 2006; Krapp, 2002; Silvia, 2001).

It should be noted that individual interest appears to have some traits in common with intrinsic motivation, but care should be taken with this interpretation since intrinsic motivation encompasses both individual interest and situational interest (Lee et al., 2014). Individual interest in a specific area is shaped through positive experiences—both in a cognitive and affective manner—during long-term participation in a task or activity related to that specific domain; reflecting a relatively well-developed and sustained form of interest. This serves as a basis for the probability of individual interest working as a concrete construct that can reveal part of an individual's characteristics. For example, a few studies delineate its facilitative role in several fields of learning, such as self-regulation processes and academic performance (Lee et al., 2014; Hidi & Renninger, 2006; Sansone & Thoman, 2005; Schiefele et al., 1992). The study by Schiefele et al. (1992) based on meta-analysis shows that individual interest significantly predicts academic achievements, although some differences exist depending on gender. Lee et al. (2014) collected data from 132 Korean adolescents and examined the roles of individual interest in the domains of Korean, English, Mathematics, and Science with other variables using structural equation modeling. The study found remarkably similar patterns in the four subject areas. Individual interest best predicts self-regulation, with self-regulation being the most influential mediator of individual interest and academic performance. These results support the idea that individual interest may also have a relationship with the development of self-efficacy, which is a strong predictor of academic achievement and can be seen to persistently

help individuals engage with and put effort into their learning processes, in a similar way that interest does (Renninger & Hidi, 2002).

Self-efficacy Beliefs

Learning a foreign language involves understanding and becoming familiar with a whole new intricate system of a language—the cultural, social, and linguistic aspects (e.g., syntax, lexis, etc.) of language and language use. In other words, it poses a significant challenge for a student to learn another language that has a completely different system from their mother tongue. In this respect, it is vital for students to have the ability to predict and manage potential challenges or threats appearing in their learning processes, or to have a strong sense of self-efficacy. As individuals exercise their ability to overcome such difficulties as they encounter them, they build a solid sense of efficacy that helps tackle and handle new challenges (Bandura, 1997).

Bandura's (1997) self-efficacy falls within the framework of the social cognitive theory of human functioning that comes from the supposition that humans have the ability to control their behavior, with a self-belief system serving as a possible guide to control what they think, feel, and act. What he defines that as "perceived self-efficacy" refers to beliefs in one's capabilities to organize and execute the course of action required to produce given attainments" (p. 3). In this point of view, a number of scholars maintain that a robust sense of efficacy beliefs equip students with a willingness to face the challenge of given tasks, to make an effort to fulfill them, and to stay on track persistently in order to perform the tasks despite there being obstacles along the way in the learning processes (Bandura, 1997; Kim & Cha, 2017; Stevens et al., 2004; Vuong et al., 2010).

Not surprisingly, there are a lot of studies that reveal the influential role of self-efficacy in academic achievements or self-regulation, which in turn are related to learning listening skills in language (e.g., Graham, 2007, 2011; Mills et al., 2006; Rahimi & Abedini, 2009, etc.). The study by Graham (2007, 2011) emphasizes the importance of student self-efficacy in developing listening skills, showing that students benefit from the formation of self-efficacy through listening strategy instruction. Recognizing that self-efficacy is a motivational construct essential for learning, it is very important to look at the relationship of variables that can reinforce students' self-efficacy.

The Relationship between Interest and Self-efficacy in Learning

Interest and self-efficacy are reported to have some association with each other in several educational fields such as writing (Hidi et al., 2002), Mathematics (Bandura & Shunk, 1981), and Physical Education (Zimmerman & Kitsantas, 1997). Hidi et al. (2002) show that interest and self-efficacy in the same domain or content area can be correlated by sharing the same knowledge base. They contribute to similar behavioral outcomes such as positive emotional reaction, persistence, making an effort, and focused attention during learning processes. This strongly indicates that these two motivational constructs may influence each other's development. However, the nature and method of their interaction still appear to be unclear (Renninger, 2010), and this holds true for the field of L2 learning. It is at this point that the current study finds its relevance and attempts to investigate how these motivational variables are related in the field of L2 listening, which may lead to further consolidate pedagogical and theoretical conjectures. The research questions are as follows.

1. What are the underlying factors found in listening self-efficacy beliefs that beginner to low-intermediate students have?
2. Is there any relationship between interest in learning English, listening proficiency, and self-efficacy beliefs in listening?
3. How do these relationships change over time?

Method

Participants

This study was conducted at a university located in eastern Korea, approximately two-hours away from Seoul, by car. This university requires students to take at least two mandatory English courses regardless of major and is offered by the general education center. The 107 participants—who agreed to join the study voluntarily—chose Listening Comprehension (LC), which aims to develop students' listening comprehension skills based on TOEIC.

Most participants' English proficiency varied between beginner and low-intermediate as seen by their TOEIC LC scores taken at the beginning of the course; the average score was 217.3 (*S.D.*[1] = 51.0) out of a total of 495 (see Table 2). Most of them were freshmen from various majors with ages rang-

[1] *S.D.* = Standard Deviation

ing from 18 to 24 (M[2] = 19.2, *S.D.* =1.6) and had spent an average of about eight years learning English. The 107 students were from the three classes taught by the author. Among them, one class was asked whether they study English apart from in the English class in the third week from the first data collection and 29 participants responded to the question. It was found that only three students (10.3%) answered that they were studying apart from the English course they were taking and almost 90% of them (n = 26) responded that they got English input only through the English class taught by the author. The data was collected through convenience sampling because the population targeted was considered sufficiently diverse for the aim of the study.

Instruments

Listening self-efficacy. For this study, eight items of self-efficacy in listening from the English Self-efficacy Questionnaire (Wang et al., 2014) were used. Some of them were modified and some added to suit the educational environment and purpose for which the research was being conducted. Firstly, three original items were modified or eliminated. For example, the item "3. Can you understand American English TV programs?" was modified to "Can you understand English TV programs (e.g., dramas which were produced in the U.S.A, the U.K. or Korea, etc.)?". By deleting the word 'American,' it was intended to expand the concept of English used in many places, not English used only in one region. With this change, item "10. Can you understand English TV programs made in China?" was removed to avoid redundancy. In addition, considering the listening contents covered in class, "15. If your English instructor gives you a tape-recorded English dialogue about everyday school matters, can you understand it?" was changed to "If your English instructor gives you a recorded English dialogue between two or three people, can you understand it?", and then, the following three items were added to measure self-efficacy that can be developed through English listening classes; "Can you concentrate on the content to which you listen?", "Can you understand important information of conversations in English?" "Can you do the tasks and assignments you have to do well, to improve your listening skills?".

In this study, ten items were translated into Korean, considering the level of language proficiency of the participants, and used to measure the participants' self-efficacy beliefs in listening and the overall Cronbach's alpha coefficient was .89, which is almost the same as that shown in Wang et al.'s (α = .88).

Listening proficiency. The results obtained from the two TOEIC LC tests were used as the participants' listening proficiency data. One of the publishers

[2] M = mean

that regularly make TOEIC related books provided the tests and they were administered in the first and the twelfth week. The publisher calculated the results. Since the process took about two weeks, the test was taken on the twelfth week so that the students were able to check their grades within the 15-week course. With a total score of 495 in the listening section, the means of their test results were 217.34 (*S.D.* = 50.96) and 259.35 (*S.D.* = 65.82), respectively.

Interest in English learning (IEL). In this study, IEL represents "individual interest," showing relatively stable affective orientations (Hidi, 1990). In order to elicit the participants' level of individual interest toward learning English, a question, "Do you have interest in learning English?", was asked in a similar way to the study by Hidi et al. (2002). They were instructed to use their own judgment to mark their answer on a six-point Likert scale ranging from 1 and 6.

Data Collection

A brief outline of this study was described, and data were collected only from those who agreed to participate. The students were aware that their responses would be released for analysis after the semester is over, and that they would remain anonymous. Data were collected twice in the 15-week course of the English class; in the second and fourteenth weeks. For the analysis, a dataset of 107 participants who responded to both pre-/post-data collection were used.

Data Analysis

To address the first research question, an exploratory factor analysis was carried out using the ten-item questionnaire to explore the underlying factors of listening self-efficacy beliefs in the current population. For the second and third research questions, two multiple hierarchical regression analyses were run for each set of pre- and post-tests. The scores of interest and proficiency were used as independent variables and the identified underlying self-efficacy factors were used as dependent variables for analysis.

Hierarchical regression analysis enables examination of whether independent variables significantly predict the identified underlying self-efficacy factors, and further examines which predictors have a greater proportion in explaining the total variance of the dependent variable. In addition, they allow examination of the changes occurring in the patterns of models by comparing the results of pre- and post-tests.

Results

To answer research question 1, an exploratory factor analysis was computed, employing the Principal Axis Factoring Extraction method and the Oblique Direct Oblimin Rotation method. Items were eliminated if they were not greater than a cutoff value of 1.0 based on the eigenvalue.

The first trial resulted in a two-factor solution, accounting for 55.1% of the total variance. All ten items of self-efficacy beliefs in listening yielded a two-factor structure. The six items were loaded on Factor 1 (F1) (α = .88), accounting for 47.5% of the total variance and was labeled as Self-efficacy in advanced English listening (SAEL) since these items indicate capabilities in advanced levels of listening, such as asking if they can understand mostly authentic texts in English through media such as speeches or TV programs. Factor 2 (F2) included the rest of the four items (α = .78), accounting for 7.6% of the total variance, and was labeled as Self-efficacy in basic English listening when learning (SBEL). These items are related to activities and tasks required to improve listening competence which are normally conducted during listening classes. A display of the factor loadings gained from factor analysis is shown in Table 1.

Table 1

Factor loadings for self-efficacy beliefs in listening

SAEL, SBEL	Factor	
	1	2
Factor 1: Self-efficacy in advanced English listening (SAEL)		
2. Can you understand English TV programs (e.g., dramas which were produced in the U.S.A, the U.K. or Korea, etc.)?	0.853	
5. Can you understand English movies without subtitles?	0.848	
4. If your English instructor gives you an English dialogue between two or three people, can you understand it?	0.747	
3. Can you understand radio programs in English speaking countries?	0.586	
6. Can you understand English songs?	0.582	
1. Can you understand stories told in English?	0.567	
Factor 2: Self-efficacy in basic English listening when learning (SBEL)		
8. Can you concentrate on the content to which you listen?		0.745
10. Can you do well the tasks and assignments you have to do to improve your listening skills?		0.658
7. Can you understand numbers spoken in English?		0.529
9. Can you get important information from English conversations heard in the class?		0.522

Note. Extraction method: Principal axis factoring / Rotation method: Direct oblimin

With the second question, this study aimed to uncover any relationship existing between interest in English learning (IEL), listening proficiency, and self-efficacy beliefs in listening when joining the English listening class. Table 2 presents the descriptive statistics for the variables that were collected in pre- and post-data collection. First, when examining the values of the pre-variables, the means for pre-SAEL and pre-SBEL were 2.7 and 3.7, respectively, on a 6-point Likert scale. This may partly reflect the characteristics of the students who have proficiency levels between beginner and intermediate, as evidenced by their listening scores in Table 2. However, their IEL was 4.3 and 4.2; higher than the mid-point.

Table 2

Descriptive statistics for the variables in pre- and post-test

Variables	*n.*	Min.	Max.	*M*	*S.D.*
Pre total self-efficacy	107	1.2	4.6	3.1	0.7
Post total self-efficacy	107	1.3	5.3	3.5	0.8
Pre-SAEL	107	1.0	4.3	2.7	0.8
Pre-SBEL	107	1.5	5.3	3.7	0.8
Post-SAEL	107	1.0	5.5	3.2	0.9
Post-SBEL	107	1.5	5.8	4.0	0.8
Pre-IEL	107	1.0	6.0	4.3	1.3
Post-IEL	107	1.0	6.0	4.2	1.3
Pre-TOEIC score (Listening proficiency)	107	65.0	350.0	217.3	51.0
Post-TOEIC score (Listening proficiency)	107	130.0	460.0	259.4	65.8

Note. *n* = number; *M* = mean; *S.D.* = Standard Deviation; SAEL = Self-efficacy in advanced English listening; SBEL = Self-efficacy in basic English listening when learning; IEL = Interest in English learning

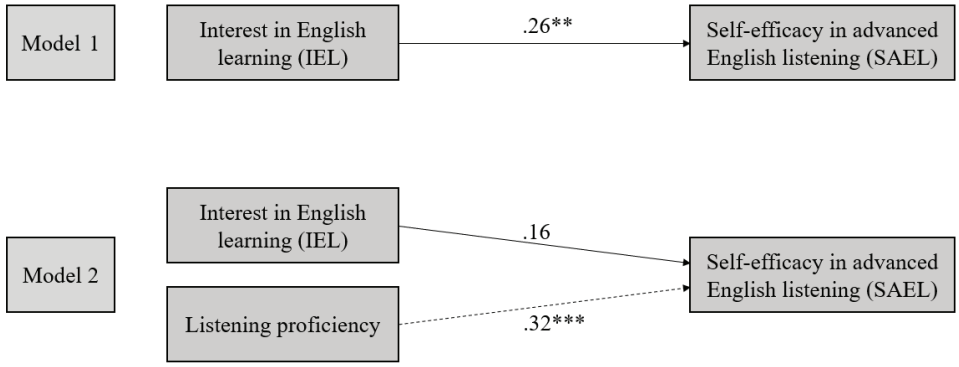
For the pre-test, the relationship between IEL, listening proficiency, and self-efficacy beliefs in listening when joining the English listening class was explored. Hierarchical multiple regression was employed for the analyses. The dependent variable was SAEL and the independent variables were IEL and listening proficiency. As Model 1 shows in Table 3, when IEL alone was used as an independent variable, it significantly predicted SAEL, $F(1, 105) = 7.89$, $p = .01$, accounting for 7% of the total variance ($R = .26$). However, when proficiency was put along with IEL as independent variables, IEL failed to retain its significance as a variable that influences SAEL ($p = .11$). In other words, while Model 2 was significant, $F(2, 104) = 9.88$, $p = .001$, accounting for 16% of the total variance, proficiency was the only variable that significantly predicts SAEL ($p = .001$). The summary of Table 3 is illustrated in Figure 1.

Table 3

The regression models with SAEL as a dependent variable for pre-test

Independent variable(s)		Unstandardized Coefficients		Standardized Coefficients			
		B	*SE*	beta	*t*-value	*p*-value	R^2
Model 1	(Constant)	1.98	.26		7.61	.00	.07
	Pre-IEL	.16	.06	.26	2.81	.01	
Model 2	(Constant)	1.21	.34		3.55	.00	.16
	Pre-IEL	.10	.06	.16	1.63	.11	
	Pre-listening proficiency	.00	.00	.32	3.33	.00	

Note. SAEL = Self-efficacy in advanced English listening; IEL = Interest in English learning



* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 1. The relationship between the variables and SAEL for pre-test

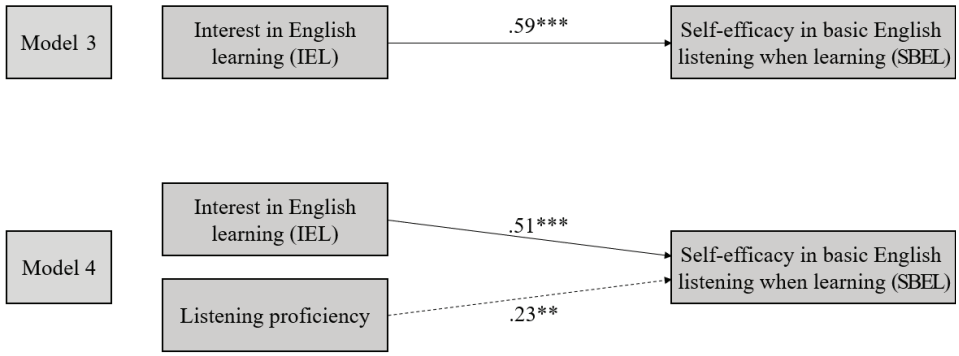
Another hierarchical regression was conducted to examine how interest and proficiency affect SBEL for the pre-test. SBEL was used as the dependent variable and IEL and listening proficiency were the independent variables. As Model 3 shows in Table 4, when IEL alone was used as an independent variable, it significantly predicted SBEL, $F(1, 105) = 55.28$, $p = .001$, accounting for 35% of the total variance ($R = .59$). When proficiency was put along with IEL as the independent variables in Model 4, this model was found to be significant, $F(2, 104) = 33.42$, $p = .001$, accounting for 39% of the total variance, predicting 4% more than that of Model 3. Examining the two predictor variables, IEL predicted SBEL far more strongly (beta = .51) than proficiency (beta = .23). The summary of Table 4 is illustrated in Figure 2.

Table 4

The regression models with SBEL as a dependent variable for pre-test

Independent variable(s)		Unstandardized Coefficients		Standardized Coefficients			
		B	*SE*	beta	*t*-value	*p*-value	*R*2
Model 3	(Constant)	2.18	.21		10.30	.00	.35
	Pre-IEL	.35	.05	.59	7.44	.00	
Model 4	(Constant)	1.64	.28		5.85	.00	.39
	Pre-IEL	.30	.05	.51	6.26	.00	
	Pre-listening proficiency	.00	.00	.23	2.81	.01	

Note. SBEL = self-efficacy in basic English listening when learning; IEL = Interest in English learning



* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 2. The relationship between the variables and SBEL for pre-test

The third research question was related to the changes in the relationship between the variables on each self-efficacy in listening over time. In the post-test, the results of the hierarchical multiple regression indicate that the explanatory values of the two predictor variables—IEL and proficiency on SAEL increased considerably, compared to those of the pre-test. For example, Model 5 with IEL as a single independent variable was significant, $F(1, 105) = 26.73$, $p = .001$, accounting for 20% of the total variance ($R = .45$), 13% more than that of the pre-test. Model 6 with both IEL and proficiency as independent variables was also found significant, $F(2, 104) = 16.13$, $p = .001$, accounting for 24% of total variance which is 8% more than that of the pre-test. It is worth noting that among the two variables, IEL was still a significant variable when proficiency was put as a combined variable on SAEL, a result different from the pre-test—the contribution of IEL to SAEL became insignificant when proficiency was

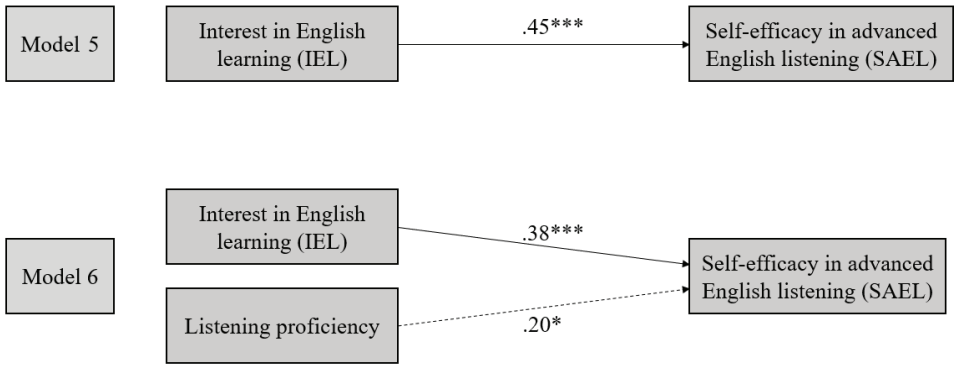
put in as a combined independent variable (see Table 3). Moreover, IEL had more predictor value (beta = .38) than that of proficiency (beta = .20) on SAEL this time. In other words, the contribution of IEL became almost twice as high as that of proficiency after a series of lessons aimed at learning to listen. The summary of Table 5 is illustrated in Figure 3.

Table 5

The regression models with SAEL as a dependent variable for post-test

Independent variable(s)		Unstandardized Coefficients		Standardized Coefficients			
		B	*SE*	beta	*t*-value	*p*-value	*R*2
Model 5	(Constant)	1.90	.26		7.19	.00	.2
	Post-IEL	.31	.06	.45	5.17	.00	
Model 6	(Constant)	1.43	.34		4.26	.00	.24
	Post-IEL	.26	.06	.38	4.05	.00	
	Post-listening proficiency	.00	.00	.20	2.15	.03	

Note. SAEL = Self-efficacy in advanced English listening; IEL = Interest in English learni



* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 3. The relationship between the variables and SAEL for post-test

To investigate the impact of the two variables on SBEL in the post-test, a hierarchical multiple regression was performed. When IEL was put in as a single independent variable, its contribution to SBEL was significant, $F(1, 105) = 45.12$, $p = .001$, accounting for 30% of the total variance. When proficiency was combined along with IEL, these two predictor variables significantly affected SBEL, $F(2, 104) = 26.81$, $p = .001$, accounting for 34% of the total variance. Compared to the results of the pre-test, the predicting values of the independent variables decreased moderately. However, these

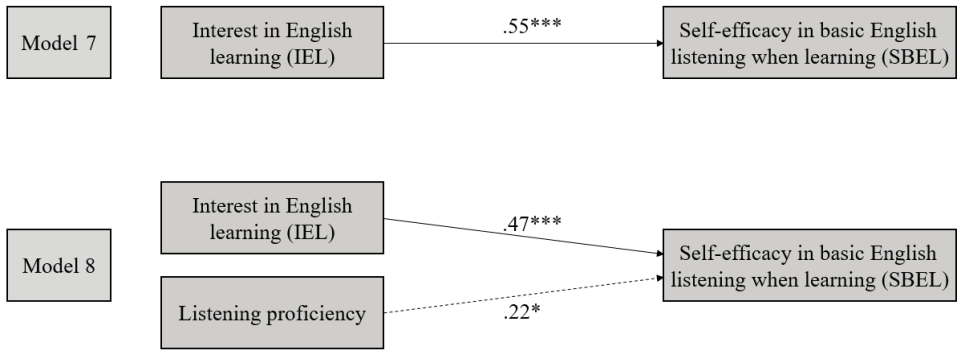
values remained in a similar pattern. In other words, IEL had a better explanatory value—more than double—with SBEL than proficiency did in both the pre- and post-test (see Figure 2 and 4). This illustrates that SBEL increases as both IEL and proficiency increase, but the degree of IEL has more influence on SBEL than proficiency.

Table 6

The regression models with SBEL as a dependent variable for post-test

Independent variable(s)		Unstandardized Coefficients		Standardized Coefficients			
		B	*SE*	beta	*t*-value	*p*-value	*R*2
Model 7	(Constant)	2.52	.23		11.01	.00	.3
	Post-Interests	.35	.05	.55	6.72	.00	
Model 8	(Constant)	2.06	.29		7.10	.00	.34
	Post-Interests	.30	.06	.47	5.42	.00	
	Post-listening proficiency	.00	.00	.22	2.50	.01	

Note. SBEL = self-efficacy in basic listening when learning; IEL = Interest in English learning



* $p < .05$, ** $p < .01$, *** $p < .001$

Figure 4. The relationship between the variables and SBEL for post-test

Discussion

This study aimed to examine the effects of IEL—individual interest—and proficiency on self-efficacy beliefs which are considered an important contributor to an individual's academic development in the domain of listening. As

shown by the average scores of the two TOEIC LC tests, most of the participants consisted of EFL university students who have an English proficiency ranging from beginner to low-intermediate. Notably, the scores of self-efficacy and listening performances increased by the end of the course as compared to those collected at the beginning of the course as shown in the descriptive statistics, but the IEL scores remained almost intact. This partly reflects the nature of individual interest in that it is considered to possess more potential for holding personal value for a relatively long period of time without much change (Hidi, 1990; Hidi & Harackiewicz, 2000).

When participants began taking this course, IEL as a single independent variable maintained its significance in predicting SAEL, but it is worth noting that its significance was lost when placed together with listening proficiency as another independent variable. In other words, the interest of participants with low proficiency in English did not play a role in predicting self-efficacy in performing advanced listening skills, mostly dealing with authentic texts, when considered alongside listening proficiency. This may be due to the possibility that engaging in difficult listening activities can be considered a daunting task for these participants to approach, even armed with their interest in learning English.

Conversely, it is considered a natural result to discover that listening proficiency is a predictor for self-efficacy beliefs especially in dealing with high-level listening tasks since several studies support the close relationship between them (e.g., Bong & Skaalvik, 2003; Cubillos & Ilvento, 2013; Graham, 2011; Mills et al., 2006, etc.)

Around the end of the course, however, the data analysis results show a markedly different pattern in the relationship between IEL and SAEL from those of the pre-test (see Table 3 and 5). For example, IEL as a single independent variable plays a significant role in predicting the formation of SAEL and explains SAEL approximately three times (20%) more in the post-test (Model 5) than the model in the pre-test (Model l). Moreover, even when English proficiency is put as a predictor along with IEL, IEL does not lose its significance and rather is shown to have more (about twice as much) influence on SAEL than English proficiency (Model 6). This means that interest did not play a role in predicting the development of SAEL in the beginning, but over time, interest turned into an influential variable, even more influential than English proficiency.

Although they had a relatively high level of interest, the pre-test result suggests that students with low proficiency in English did not describe themselves as capable of advanced listening skills. It is likely that such a result may have been caused by them cognitively recognizing the gap between their skills, based on their subjective judgments, and the level of English listening activities they knew they would face at the beginning of the course or the authentic materials they would have to listen to. The post-test result, however, indicates that the interest they possess in learning English may emerge as a significant

and influential facilitator of their efficacy beliefs in listening to and understanding difficult or authentic materials through a series of learnings such as taking classes, even for students who are considered low-intermediate. It is worth recalling that at the time of data collection, few students were learning English other than in the English course, despite the data being collected from only one of the three classes (see Participant section in Method). This possibly indicates that the majority of their input of English learning was through the English course that they were taking. It is plausible to infer that the four major sources of self-efficacy that Bandura (1997) indicates—personal success in performing listening activities (mastery experiences), observed indirect successful experiences by classmates (vicarious experiences), received encouragements and positive feedback from the teacher and classmates (social persuasion), and psychological and affective states (Bandura, 1997)—may have contributed to such a pattern of changes. However, such casual relationships could not be investigated further as they fall beyond the scope of the current study. The positive experiences students likely have through the classes can turn students' interest into a factor that actually affects self-efficacy.

Regarding the SBEL, at the beginning of the course, the participants who had interest in studying English seem to have the idea that they are capable of performing listening activities that require the basic skills needed to improve listening competence, which are normally conducted during listening classes (e.g., understanding numbers spoken in English, concentrating on listening content, performing related activities, etc.). This is seen in Models 3 and 4, accounting for 35% and 39% of total variance respectively and IEL had a greater effect on SBEL than actual English proficiency, unlike its effect on SAEL. This indicates that these participants' perception that they are interested in English learning affects their self-efficacy making them believe that they can perform listening activities at a relatively easy level. Even around the end of the semester, this pattern remained unchanged; although interest had a higher influence on SBEL in both models (models 4 and 8), the explanatory powers of the two post-models (Model 7 and 8) became moderately smaller than those of the pre-test, decreasing by 5% of the total variance each. This is possible because as time went by, most students seem to perceive that they have obtained the enhanced ability to deal with and get used to handling a basic level of activities with relative ease, through participation in the classes. In this respect, it can be inferred that the explanatory powers in the regression models decreased moderately in the post-test whereas the regression models maintained the same pattern—individual interest has a greater influence on the development of SBEL—in both pre- and post-test. Looking at the relationship between IEL and SBEL in pre- and post-results indicates that providing only an appropriate level of activities that are not demanding or challenging for the student's level does not help develop self-efficacy.

Conclusion

This study aimed to understand the complexities and the nature of interest and self-efficacy beliefs in university students with low English proficiency in Korea. The findings show that their interest in English learning was quite high despite the participants of this study having had little success in learning English in a relatively long period of time. In addition, these students were found to have their self-efficacy in a provisional or malleable state (Klassen, 2006)—a transitional state of self-efficacy that is still undergoing changes and has positive development even in students with a low level of proficiency. Moreover, it was revealed that the formation of self-efficacy beliefs in performing advanced listening skills can be facilitated by individual interest, one of the motivational variables in learning English. As Hidi and Harackiewicz (2000) highlight, students' interest is likely to influence the development of a strong sense of efficacy over time since interest plays a strong facilitating role in cognitive function. In this respect, helping students develop their interest in language learning is effective in developing their self-efficacy beliefs.

Individual interest and self-efficacy beliefs are particularly important in that they are motivational constructs that can influence learning processes. The findings of this study suggest that even for those with a low-proficiency and relatively little success in learning EFL, it is important to ensure the continuation of individual interest in English learning. If done correctly, the possibility of the student's progress is substantially higher than if their interests are not taken into account. When individual interest in one domain is met with several conducive learning conditions, it may play a role as a facilitator of self-efficacy. It would be worth investigating what role individual interest plays as a mediator in the formation of self-efficacy beliefs in future studies. Notably, the results were derived from convenience sampling and cannot be generalized. This suggests further research is needed with different samples. In addition, a valid and reliable instrument to measure individual interest should be considered to better support the findings.

References

Ainley, M., Hidi, S., & Berndorff, D. (2002). Interest, learning, and the psychological processes that mediate their relationship. *Journal of Educational Psychology, 94*(3), 545–561.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* W. H. Freeman and Company.

Bandura. A., & Schunk, D. H. (1981). Cultivating competence, self-efficacy, and intrinsic interest through proximal self-motivation. *Journal of Personality and Social Psychology, 41*(3), 586–598.

Bong., M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1–40.

Cubillos, J. H., & Ilvento, T. (2013). The impact of study abroad on students' self-efficacy perceptions. *Foreign Language Annals, 45*(4), 494–511.

Graham, S. (2007). Learner strategies and self-efficacy: Making the connection. *Language Learning Journal, 35*(1), 81–93.

Graham, S. (2011). Self-efficacy and academic listening. *Journal of English for Academic Purposes, 10*(2), 113–117.

Hidi, S. (1990). Interest and its contribution as a mental resource for learning. *Review of Educational Research, 60*(4), 549–571.

Hidi, S., Berndorff, D., & Ainley, M. (2002). Children's argument writing, interest and self-efficacy: An intervention study. *Learning and Instruction, 12*(4), 429–446.

Hidi, S., & Harackiewicz, J. (2000). Motivating the academically unmotivated: A critical issue for the 21st century. *Review of Educational Research, 70*(2), 151–179.

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist, 41*(2), 111–127.

Kim, H. I., & Cha, K. A. (2017). Effects of experience abroad and language proficiency on self-efficacy beliefs in language learning. *Psychological Reports, 120*(4), 670–694.

Klassen, R. M. (2006). Too much confidence: The self-efficacy of adolescents with learning disabilities. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents* (pp. 181–200). Information Aga Publishing.

Krapp, A. (2002). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction, 12*(4), 383–409.

Lee, W., Lee, M., & Bong, M. (2014). Testing interest and self-efficacy as predictors of academic self-regulation and achievement. *Contemporary Educational Psychology, 39*(2), 86–99.

MacIntyre, P. D., & Mercer, S. (2014). Introducing positive psychology in SLA. *Studies in Second Language Learning and Teaching, 4*(2), 153–172.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276–295.

Rahimi, A., & Abedini, A. (2009). The interface between EFL learners' self-efficacy concerning listening comprehension and listening proficiency. *Novitas-ROYAL (Research on Youth and Language), 3*(1), 14–28.

Renninger, K. A., & Hidi, S. (2002). Student interest and achievement: Developmental issues raised by a case study. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 173–195). Academic.

Renninger, K. A. (2010). Working with and cultivating the development of interest, self-efficacy, and self-regulation. In D. D. Preiss & R. J. Sternberg (Eds.), *Innovations in educational psychology: Perspectives on learning, teaching, and human development* (pp. 107–138). Springer Publishing Company.

Sansone, C., & Thoman, D. B. (2005). Interest as the missing motivator in self-regulation. *European Psychologist, 10*(3), 175–186.

Schiefele, U., Krapp, A., & Winteler, A. (1992). Interest as a predictor of academic achievement: A meta-analysis of research. In K. A. Renninger, S. Hidi, & A. Krapp (Eds.), *The role of interest in learning and development* (pp. 183–212). Erlbaum.

Silvia, P. J. (2001). Interest and interests: The psychology of constructive capriciousness. *Review of General Psychology, 5*(3), 270–290.

Stevens, T., Olivarez, A. Lan, W. Y., & Tallent-Runnels, M. K. (2004). Role of mathematics self-efficacy and motivation in mathematics performance across ethnicity. *The Journal of Educational Research, 97*(4), 208–221.

Vuong, M., Brown-Welty, S., & Tracz, S. (2010). The effects of self-efficacy on academic success of first-generation college sophomore students. *Journal of College Student Development, 51*(1), 50–64.

Wang, L., & Fan, J. (2015). Listening difficulties of low-proficiency EFL learners: A comparison of teacher and learner perspectives. *The Asian EFL Journal, 17*(3), 85–110.

Wang, C., Kim, D. H., Bai, R., & Hu, J. (2014). Psychometric properties of a self-efficacy scale for English language learners in China. *System, 44*(1), 24–33.

Zimmerman, B. J., & Kitsantas, A. (1997). Developmental phases in self-regulation: Shifting from progress to outcome goals. *Journal of Educational Psychology, 89*(1), 29–36.

Hyang-Il Kim

Der Einfluss von individuellen Interessen und Kenntnissen auf Selbstwirksamkeitsüberzeugungen im Verstehen von Fremdsprachen

Zusammenfassung

Interesse und Selbstwirksamkeitsüberzeugungen sind wichtige Bestandteile von Motivationskonstrukten, die einige gemeinsame Merkmale aufweisen und einen Einfluss auf den Lernprozess haben. In diesem Zusammenhang hat man in mehreren Studien untersucht, wie die beiden Variablen unter verschiedenen Aspekten miteinander in Beziehung stehen, und auf dieser Grundlage sinnvolle Schlussfolgerungen formuliert. Nichtsdestotrotz scheint die diesbezügliche Forschung im Bereich des Fremdsprachenlernens relativ unzureichend und die aus den früheren Studien gewonnenen Informationen sehr begrenzt zu sein. Aus diesem Grund ist die vorliegende Studie darauf abgezielt, zu untersuchen, wie das individuelle Interesse am Englischlernen und Sprachkenntnisse die Selbstwirksamkeit im Hörverstehen beeinflussen, um das vorhandene Wissen zu erweitern. Mittles eines Vortest-Nachtest-Designs wurden die Daten von 107 koreanischen EFL-Studenten mit niedrigeren mittleren Englischkenntnissen unter Verwendung hierarchischer Regressionsanalysen ausgewertet. Die Ergebnisse zeigen, dass, anders als zu Beginn des Kurses, das individuelle Interesse ein signifikanter Faktor bei der Entwicklung von Selbstwirksamkeit in Bezug auf die fortgeschrittene Kompetenz am Ende des Semesters war und sie sogar mehr voraussagte als Sprachkenntnisse. Von der anderen Seite hatte das individuelle Interesse einen viel stärkeren Einfluss auf die grundsätzliche Selbstwirksamkeit im Hörverstehen als die Sprachkenntnisse im Vortest und dieses Modell wurde auch im Nachtest beibehalten. Die Studie bietet auch aufschlussreiche Informationen über die Wechselbeziehung zwischen Interesse, Selbstwirksamkeit bzw. Leistung und hebt die bedeutende Rolle des individuellen Interesses beim Englischlernen hervor. Daraus folgt, dass die Aufrechterhaltung und Förderung des Interesses entscheidend für die Entwicklung von Selbstwirksamkeitsüberzeugungen ist, die wiederum zu Fortschritten beim Fremdsprachenlernen beitragen können.

Schlüsselwörter: individuelles Interesse, Hörverstehen, Fremdsprachenkenntnisse, Selbstwirksamkeitsüberzeugungen

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 71–84
https://doi.org/10.31261/TAPSLA.9501

Bogusława Maria Gosiewska-Turek https://orcid.org/0000-0002-8438-1618
University of Opole, Poland

Dyslexia, Self-efficacy, and Language Instruction in Foreign Language Learning— A Mixed Quantitative-qualitative Study

Abstract

The aim of the research is to investigate the interdependence between dyslexia, self-efficacy, and foreign language instruction. The author of the study applied a mixed-method: quantitative and qualitative. The quantitative data were collected through self-efficacy questionnaires filled out by dyslexic and non-dyslexic students, and the qualitative data were collected during a case study conducted with a dyslexic student. The research findings in the first quantitative part of the study reveal that self-efficacy in dyslexic students is substantially lower than in non-dyslexic students. According to the results of the second mixed, quantitative-qualitative part of the study encompassing a case study, foreign language instruction has an impact on dyslexic students' self-efficacy and appropriate language instruction rises self-efficacy in students with dyslexia.

Keywords: dyslexia, self-efficacy, foreign language instruction, language learning, language teaching

Dyslexia and Dyslexia Related Foreign Language Learning Problems

In each school environment teachers deal with students of at least average intellectual level but displaying symptoms of learning difficulties, primarily in reading and writing. According to the data revealed by the Polish Dyslexia Association 10–15% children worldwide are affected by dyslexia, 4% by deep dyslexia. Based on various learning difficulties Bogdanowicz (1993) distinguished four dyslexia types:

- visual dyslexia: learning problems cover mainly perception, aural memory combined with difficulties related to language function;
- integrative dyslexia: difficulties concerning motor skills-integration of perception;
- mixed dyslexia: difficulties related to aural perception and memory, visual memory of words, visual sequencing memory as well as spatial imagination;
- aural dyslexia: related to perception and auditory memory disorders frequently combined with language function disorder.

Dyslexic students experience educational difficulties not only learning their native language but also foreign language (Crombie, 2000; Kormos et al., 2009). These difficulties concern orthographic and phonological processing of sentences, reciting or weak utterance fluency; poor phonemic awareness characterized by lack of proper spelling which in turn results in problems with vocabulary learning due to poor working memory; lack of metacognitive abilities or individual language analysis, self-assessment or weak concentration.

Foreign language learning difficulty is determined by interdependence between graphic signs and sounds, which in science is called deep orthography (Frost, 2005). Deep orthographies can be characterized by inconsistent and unpredictable interdependence. English is a language with deep orthography in which this relationship is unstable, and the same graphic signs can be read in different letters in various ways which affects and increases learning difficulties (Spencer, 2007). Polish language, on the other hand, is a transparent language where graphic signs are always read in the same way. As a result, foreign language learning proves to be exceptionally difficult for students with dyslexia.

Self-efficacy

Affective factors have an impact on learners' foreign language success. According to some authors affectivity is even more important than students' aptitude (Chastain, 1988), and self-efficacy is believed to be crucial in academic performance.

Self-efficacy theory has been introduced by Albert Bandura (1977, 1997), who based his model on socio-cognitive theory according to which people are capable of undertaking the intentional activity. This activity is prioritized in accordance with trifold relations. It points out that individuals' acts have an impact on their future behavior being a result of trifold interrelated associations which cover: the environment in which individuals are brought up, their behavior, and biological factors. The trifold model determines people's opinions about themselves and affects their acts (Bandura, 1977).

Albert Bandura came to the conclusion that people's opinions about their opportunities to act and the effects of their efforts have an impact on the way they behave. In his book *Social Foundations of Thought and Action*, Bandura claims that people's personal development is based on "what people think has an impact on their behaviour" (Bandura, 1983, p. 8).

Bandura (1977) in his cognitive theory enumerated the main sources of self-efficacy. Among them he listed: mastery experiences, modelling, verbal-social persuasion, and physiological and affective states. Mastery experiences are the most essential sources of self-efficacy and they relate to people's perception of their own previous success or failure. Thus, after the task is done by individuals, they assess the result of their effort. When they come to the conclusion that their effort led to success, self-efficacy when dealing with similar task, increases. Conversely, when they perceive their effort unsuccessful, they become unsecure during similar task.

Moreover, students built their self-efficacy through modelling, that is, observing the others and interpreting their experiences (Usher & Pajeras, 2006). The more students identify themselves with role models, the more probably it is that their self-efficacy will increase (Bandura, 1986). For instance, when the learner gets 50 out of 100 points in the test and discovers that the other classmates obtained fewer points, it is likely that student's self-efficacy will increase. On the other hand, when the student learns that the other classmates received more points, the learner's self-efficacy will decrease.

The subsequent source of self-efficacy is the verbal-social persuasion of significant others. In verbal-social persuasion students' self-evaluation is realized based on students' assessment of significant others. The students' attitude towards those people depends on how they trust them and their professionalism. Moreover, Bandura (1986) underlined that verbal persuasion is more efficient when it causes a long-lasting increase in self-efficacy. Thus, it is advisable to find a person, role-model who is capable of doing the task sufficiently. It could be a teacher, a parent, or symbolic model such as a person known from media (Bandura, 1997). In foreign language learning self-efficacy relates to self-regulation noticeable in metacognitive strategies. Academic self-efficacy describes individuals' assurance in realizing educational activities on a given level (Shunk, 1991). Thus, learners with a higher level of aptitude are more successful learners than those with lower aptitude. However, academic performance is a process covering various factors so high aptitude does not determine success.

Multisensory Instruction

Contemporary foreign language teaching theories suggest language learning which resembles native language learning: firstly, by listening, then speaking and reading, followed by writing. The widely employed teaching method is a communicative method in which meaning is inferred from the context, grammar rules are taught directly and the interdependence between a sound and a symbol has been limited (Ganschaw et al., 2000).

Nevertheless, researches confirm that the traditional teaching approach proves to be more efficient in teaching students with learning difficulties (Ganschaw et al., 1998). Reading comprehension with verbal processing combined with processing through recognition of words, symbol identification was efficient in teaching students with learning difficulties. Very important is early contact with orthography (relation between sound and its graphic sign) (Sparks et al., 2006).

Since the communicative method turned out to be inefficient in teaching dyslexic students, multisensory teaching instruction appeared the more effective method in teaching these students. In this method all senses are activated in accordance with the rule: seeing, hearing, and doing. It was presented for the first time by Gillingham and Stillman in 1960, and subsequently introduced in some schools around the world as a teaching method for dyslexic students. Teaching is realized with the employment of direct instruction, the meaning should not be inferred by students, but rather provided by a teacher. Dyslexics benefit primarily when they receive systematic and structured grammar and phonetic rules which are followed by relevant exercises and consolidation of knowledge they obtained.

Moreover, activation of various sensory channels is essential. Therefore, children make letter models from plasticine, trace words written on paper, carpet or sand. When their perception channel has been opened, they find the relationship between graphic and phonological aspects of words. As a result, information is acquired in complex activation of different senses.

In multisensory teaching, students learn vocabulary using pictures. They are asked to look at the picture with a written word, they say the word and then write. They also learn vocabulary at the board by repeating the word after the teacher, who then writes the word on the board. Finally, students repeat the word again and write in their notebooks. The teacher can also separate syllables and sounds in words (Sparks et al., 1991).

Grammar concepts are taught directly. For instance, the teacher can introduce color code according to which red color is used to match the subject, blue to match the object and yellow, the verb. When such code is

used systematically, the learners will understand grammar better (Kormos & Smith, 2012).

Communicative exercises can also be conducted with the employment of multisensory teaching. This activity begins with natural communication through listening and speaking. Nevertheless, students should activate more senses while learning. That is why they can be asked to prepare cards with personal pronouns *my*, *your*, *his*, etc. Later they write a few sentences with these pronouns and finally the learners are paired and one student is asked to read the sentence with personal pronoun, and the other student choses the second card with matching picture and repeats the sentence (Sparks et al., 1991).

However, opponents of multisensory instruction claim that in this method it is impossible to develop metacognitive processes and it does not take into account various learning and cognitive skills (Wearmouth & Reid, 2008).

Empirical Research

Methodology

This part of the paper outlines the empirical study. The primary objective of a study was to explore the relationship between self-efficacy and dyslexia, and the secondary objective, to investigate the impact of teaching methods on self-efficacy in dyslexic student.

Even though there are many studies concerning self-efficacy in a school context, only a few of them have been investigating the relationship between self-efficacy and foreign language learning. However, none of the researches examine self-efficacy in connection with teaching instruction employed in teaching students with dyslexia. Therefore, the author of the study formulated the following research questions:

RQ1: What is the relationship between self-efficacy and foreign language learning?

RQ2: How does foreign language learning instruction affect dyslexic students' self-efficacy?

The study's data collection method was a mixed study. The quantitative study relates to the objective analysis of the data achieved from self-efficacy questionnaires conducted among dyslexic and non-dyslexic primary school students. The qualitative method is realized in the form of a case study conducted with a dyslexic student.

Participants, the Instrument, and the Procedure

The informants in the quantitative part of the study were 20 dyslexic and 50 non-dyslexic primary school students who attend a primary school in Lędziny, Poland. There were 22 non-dyslexic girls, 28 non-dyslexic boys, and 20 dyslexic boys. The average age of the students was 13. All students studied English as a part of their school curriculum and had three hours a week of English class. As a result, they had at least 6-year English exposure at school.

In the first part of the study conducted with the application of the quantitative method, the employed instrument was a questionnaire consisting of 15 questions and distributed among the participants. Prepared by the researcher the Foreign Language Self-efficacy questionnaire measured the level of English language learning related self-efficacy. The researcher applied a scale evaluated according to a 4-point Likert scale ranging from 1 to 5 with the possible responses: strongly disagree—1, disagree—2, agree—3, and strongly agree—4, where 1 indicted very low, and 4—very high self-efficacy. The data were collected within two weeks. The headteacher's permission to conduct research was obtained in the primary school.

The second part covered the qualitative case study without generalizing to the whole population, combined with the quantitative one. The participant of a qualitative case study was a 13-year-old boy with dyslexia who is a sixth-year student of a primary school in Mysłowice. The qualitative study helped to understand the situation which is unique, and has not been analyzed before. The case study lasted six weeks during which the researcher conducted six 45-minute classes with the student. Also, the researcher applied Foreign Language Self-Efficacy questionnaire, the same which was filled out by 70 primary school students. The questionnaire was employed to a dyslexic student before and after a case study.

Research Results

In the first part of the study the author attempted to answer the first research question. Hence, the quantitative methodology was applied. The data were collected within one week and was computed by means of the statistical software RStudio with the primary calculations being descriptive statistics: mean (M), and standard deviation (SD), as well as interferential statistics operations: The Mann Whitney Test, and the Shapiro-Wilk T-parametric Test conducted in order to investigate whether there was a statistical difference between self-efficacy means scores in dyslexic and non-dyslexic students.

Table 1

Research results obtained from dyslexic and non-dyslexic students' self-efficacy foreign language questionnaires

Variable	All (70)	Non-dyslexics (50)	Dyslexics (20)	Dyslexics and non-dyslexics
	M SD	M SD	M SD	P-value T-value
Self-efficacy	39.84 6.13	41.80 5.59	34.95 4.53	3.466e-06 –5.32461

The research results reveal that there is a substantial difference in dyslexic and non-dyslexic learners of English with dyslexics experiencing lower levels of self-efficacy (M 34.95), than non-dyslexic students (M 41.80). Furthermore, the score of 6.13 obtained from SD calculation in all (dyslexic and non-dyslexic students) revealed a normal distribution of scores in all students where students' lowest score was 28 points in the Self-efficacy Foreign Language questionnaire, and the highest was 53 points in non-dyslexic students. The highest score which could have been obtained by the participants was 60 points, and the lowest—15. The result 5.59 achieved from non-dyslexic students' questionnaires also show normal distribution of scores with the lowest score amounted 30 points and the highest—53. Nevertheless, dyslexic students with a score 4.53, displayed non-normal distribution, and their lowest score in the questionnaires was 28, and the highest 45.

Following the procedure, as the sample size is not large, in order to compare the two groups of participants (dyslexics and non-dyslexics), the researcher applied The Mann Whitney interferential (Willcoxon Rank-sum Test). It is a non-parametric test which tests whether there is a significant difference in the two groups of participants. The value above 0.05 shows that there is no statistical difference. In this study, the P-value which amounted 3.466e-06 indicates that there is a substantial difference in self-efficacy between dyslexic and non-dyslexic students. Finally, in order to examine the result obtained from the Willcoxon Rank the researcher applied the Shapiro-Wilk t-parametric Test. The score –5.32461 confirmed the result achieved in the Willcoxon Rank-sum.

In order to answer the second research question, the author of the study administered a mixed quantitative-qualitative study. The participant of a qualitative study was a 13-year old dyslexic boy who holds an opinion from pedagogical-psychological suspensory. In the statement it has been stipulated that the boy has got decreased aural-vision-kinaesthetic integration as well as graphomotor abilities. In writing the boy commits many orthographic mistakes and distorts longer words. The boy reads slowly word by word. Moreover, his aural and phonological memory is limited, which results in weaker auditory channels.

To see the whole picture of the conundrum, the author began the research with the Foreign Language Self-efficacy questionnaire filled out by a student. The mean was 2.3 which clearly indicates that the boy had low self-efficacy regarding foreign language learning. For instance, to the statement "English is one of the easiest subjects," he responded "strongly disagree." He also gave the same answer to another statement "Teachers take care of everything to ease me to understand what they explain." Moreover, the learner assessed himself as a weak English language student.

The case study was performed by the researcher, who is simultaneously an active English teacher and was conducted from 9 November until 20 December 2019. For the purpose of the study the researcher changed the boy's name into Adam. Adam is the student in a state primary school in Mysłowice. The author decided to work with Adam to examine what can be done to help a dyslexic student in foreign language learning. The study was based on the learner's student book *Flash 6* (2019).

The first week. The author began the study examining the student's English language skills. Thus, she tested Adam's speaking, writing, listening, and reading skills. Adam did the language tasks as well as he could. At first, he seemed to be reluctant so as "an ice breaker" she applied Taboo game. In the game, the student had to describe the words he saw on the card without using enumerated words. Subsequently, the researcher checked Adam's writing skills to which she used vocabulary from student's book *Flash*. To encourage Adam to work, for listening the researcher selected a video from YouTube entitled "What are computers for kids. Intro to computers."

After the first session with Adam, it has been noticeable that the learner who is willing to communicate in English, struggles while speaking making a number of language mistakes. Moreover, he commits many grammar mistakes and pronounces incorrectly most of the English words. Regarding writing, Adam wrote correctly 4 words out of 20. However, mistakes in his writing were not substantial for instance, instead of "ear infection," he wrote "er infection," or instead of "stomach ache," he wrote "stomachkage" and Adam's handwriting was illegible. While testing listening skill Adam watched the video with interest. However, it was problematic for Adam to report what he heard, either in Polish or in English. The last examined skill was reading the text *Escape room* from Adam's student's book. During reading Adam mispronounced many words or skipped some lines. Hence, initial diagnosis revealed the necessity of working on all the language skills.

The second week. During the second session the study was based on the fourth chapter *In the News* from Adam's student's book. The researcher focused on learning the vocabulary from the introduction to the chapter. She prepared flashcards with words such as "customer" or "virtual reality." Firstly, the tutor gave Adam flashcards with Polish and English meanings and his task was to

join the meanings in both languages. Then she read the words in English and the student repeated after her. The next stage was to tell words in English. Finally, the student traced the words in English on the sand, and the last task was to say and write words in a notebook. Towards the end, to make the session more attractive the researcher asked Adam to draw eight words and to make a story with these words.

The training based on multisensory instruction, activating various perception channels proved to be successful. Adam was able to recall from his memory 18 out of 20 learned words. The task in which Adam was asked to make a story was done with enthusiasm. As Adam is interested in history, he created the story set during World War II. His story was incorrect and while creating it Adam asked about a few words but the aim of this part was rather to convince him that he can break the language barrier.

The third week. In the third session the researcher, following what student's book suggested focused on grammar, that is teaching modal verbs such as *can*, *can't*, *could*, *couldn't*, *must*, *mustn't*, *have to*, *should*, *shouldn't*. Firstly, the author of the study explained to Adam the usage of modal verbs referring to learner's experiences, for instance, "when you want to go out with your friends to play football you ask 'Can I go to the football pitch?'" The next step was to teach Adam how to use modal verbs. Here, the researcher used earlier prepared cards where the person, modal verb, and object were written in different colors. With these cards Adam learnt how to build affirmative sentences, questions, and negations. In the last task, the student did the exercises from his book.

Learning grammar with the reference to the boy's personal experiences helped him to understand the usage of modal verbs easier. The boy enjoyed learning grammar through cards, and later examination of acquired skills showed that the boy understood the topic and was able to build sentences with modal verbs himself. The only tasks which were done incorrectly were the ones where Adam did not understand the meaning of sentences.

Fourth week. This time the teacher focused on teaching reading comprehension. Adam worked with a text *Droids at work* from his student's book. Initially, Adam had to underline words he perceived difficult. Then, the trainer taught him these words and a few others with the multisensory method used earlier. The next point of the session was "shared reading" (Cimermanova, 2015) in which at first the researcher and then the student together with the teacher read the text in turns. Adam received a bookmark from the researcher, thanks to which he did not skip the lines. Subsequently, when it proved to be impossible for Adam to answer the questions to the text, they translated it together into a student's mother tongue. The following activity was to create his own comic with the words learnt during the session in which the webpage www.storyboardthat.com was valuable.

Fifth week. This class involved listening comprehension. The text *Internet safety tips* served as a basis. In order to rise student's interest, the researcher prepared cards with vocabulary from the listening that is shopping online, logging in, antivirus software, and a stranger in the Internet. Firstly, the teacher and the student discussed the problems from the cards, and then Adam listened to the recording with the book closed. Finally, the boy did the task in which he filled in the gaps in the sentences placed under the text. Adam did not manage to do the whole exercise, thus he listened to the recording again to finish the task. Subsequently, the researcher focused on vocabulary which she taught Adam with cards and educational sand, after which the boy wrote the words in his notebook. The last, final stage of the fifth session was to check whether the student learnt the vocabulary doing a quiz prepared by the teacher with the assistance of the web application www.classtime.com.

Sixth week. The aim of the last session was to examine the effectiveness of earlier applied strategies and whether these methods had a considerable impact on Adam's self-efficacy. To achieve this goal the researcher examined all taught skills. She began with speaking in which the student participated in the conversation about his family, interests, and friends. These were not the topics discussed earlier. However, the teacher noticed that although Adam still made grammar and lexical mistakes, he willingly took part in the speaking practice. Then the researcher suggested the so-called controlled conversation that is based on one of the topics discussed during the sessions. This time, Adam had to select cards with vocabulary related to the text *Internet safety*, then he discussed the topics suggested in the cards. The following task was to write ten words which Adam had learnt during our case study and the learner wrote correctly six out of ten. Compared to the knowledge of vocabulary learnt before our sessions it was a good result. Subsequently, Adam read loudly the text he had never read before. The first thing he did was to reach for a bookmark which helped him read without skipping lines. He still made a lot of mistakes but he tried to read slowly and carefully. However, the last grammar task examined whether Adam was able to learn how to use modal verbs proved successful. Although the session related to modal verbs took place a few weeks before, the exercises were done perfectly with only a few mistakes.

In the last stage of the study Adam filled out the self-efficacy questionnaire, the same which he received before the case study. This time the average score was higher and it amounted for 2.9 which showed that after 6-week sessions with the trainer the boy's self-efficacy increased and he began to believe that he could make decisions related to foreign language learning independently. For instance, to the statement "I could get better grades if I worked more," he responded "yes," and not "no" like before the case study.

Likewise, he agreed with the statement "Teachers care for us to understand everything."

Discussion and Conclusion

The presented study aimed to answer two research questions. The objective of the first research question was to examine foreign language self-efficacy in dyslexic and non-dyslexic students. Not surprisingly, the research results show that dyslexic students display a significantly lower levels of English language self-efficacy. The study results confirm the findings obtained from the research conducted by Stagg and Eaton (2018) who investigated self-efficacy in dyslexic, and non-dyslexic university students. Nevertheless, according to the researcher's knowledge, the present study is the first one investigating dyslexia, self-efficacy, and foreign language instruction.

The second research question examined the relationship between self-efficacy and English language learning in the dyslexic student. The results received from the Foreign Language Self-efficacy questionnaires filled out by a participant in a case study reveal that the employment of multisensory instruction improved his English learning related self-efficacy. It increased from 2.3 to 2.9. The lowest score which could have been obtained was 1, and the highest—4.

Summing up, 6-week time spent with Adam, the boy primarily needed serious support in order to help him to believe that he is capable of achieving success in foreign language learning. Observing Adam's engagement, it can be concluded that he needed the teacher's attention, and positive relation with the trainer encouraged him to work. Each session began with small talk and the teacher complemented each student's success. Regaining self-confidence in dyslexic student is perceived as an essential factor conditioning the skill to read and write (Nijakowska, 2010).

The participant benefited from multisensory instruction based on the simultaneous activation of a few senses while doing the task, which eased him to learn vocabulary and understand grammar, primarily creating the relationship between visual and phonological aspects of words. The employment of modern technologies: applications for creating comic books and a website for preparing quizzes was beneficial and motivating in boosting participant's confidence. Individual approach during private sessions caused that the boy began to believe in his language capabilities which lifted his self-efficacy, in dyslexic students usually lower than in students without any special educational needs (Stagg & Eaton, 2018).

There is a wide range of pedagogical implications of the present study. Research findings indicate that it is relevant to introduce special educational accommodations in public schools in order to ensure dyslexia-friendly education. More individualized approach should be realized in the form of special correction-compensation classes conducted with the application of the method adopted for teaching dyslexics during which students with special educational needs will receive positive feedback and support. Also, forgoing study shows that the application of multisensory instruction combined with modern technologies can be beneficial.

However, the attempt to generalize the following study results to the whole population can be a subject of certain limitations. The participants of the study filled out the questionnaires without the presence of the researcher which resulted in a lack of control over the procedure of data collection in the quantitative part of the research. Moreover, qualitative data were analyzed solely by the author. In order to avoid bias, another independent researcher with the same qualifications is recommended. The number of participants can also be inadequate. The quantitative study was conducted on 20 dyslexic and 50 non-dyslexic students. Thus, the research findings can be considered as assumptions rather than the facts relevant to the majority of Polish students.

References

Bandura, A. (1977). *Social learning theory*. Prentice Hall, Inc.

Bandura, A. (1983). Self-efficacy determinants of anticipated fears and calamities. *Journal of Personality and Social Psychology, 45*, 464–469.

Bandura, A. (1986). *Social foundations of thought and action: A social-cognitive theory*. Prentice-Hall, Inc.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. Freeman.

Bogdanowicz, M. (1993). Specyficzne trudności w opanowaniu mowy pisanej: czytania i pisania [Specific Writing and Reading Difficulties]. In T. Gałkowski, Z. Tarkowski, & Z. Zaleski (Eds.), *Diagnoza i terapia zaburzeń mowy* [Diagnosis and Therapy of Speech Impairment], (pp. 203–229). UMCS.

Chastain, K. (1988). Developing second language skills: Theory and practice (3rd ed.). Harcourt Brace Jovanovich.

Cimermanova, I. (2015). *Teaching foreign languages to learners with special educational needs: E-textbooks for foreign language teachers.* Constantine The Philosopher University.

Crombie, M. A. (2000). Dyslexia and learning of a foreign language at school: Where are we going? *Dyslexia 6*, 112–123.

Dooley, J. (2019). *Flash. Student's book.* Express Publishing.

Frost, R. (2005). Orthographic systems and skills word recognition in reading. In M. J. Snowling & C. Hulme (Eds.), *The science of reading. A handbook* (pp. 272–295). Blackwell.

Ganschow, L., & Sparks, R. L. (2000). Reflection on language study for students with language learning problems: Research, issues and challenges. *Dyslexia, 6*, 87–100.

Ganschow, L., Sparks, R. L., & Javorsky, J. (1998). Foreign language learning difficulties: A historical perspective. *Journal of Learning Disabilities, 31*(3), 248–258.

Gillingham, A., & Stillman, B. W. (1960). *Remedial training for children with specific disability in reading, spelling and penmanship* (7th ed.). Educators Publishing Service.

Kormos, J., Sarkardi, A., & Csizer, K. (2009). The language learning experiences of students with dyslexia: Lessons from an interview study. *Innovation in Language Learning and Teaching, 3*(2), 115–130.

Kormos, J., & Smith, A. M. (2012). Teaching languages to students with specific learning difficulties. Multilingual Matters [Book review], *CEPS Journal, 2*, 181–183.

Nijakowska, J. (2010). *Dyslexia in foreign language classroom*. Multilingual Matters.

Sparks, R., Patton, J., Ganschow, L., Humbach, N., & Javorsky, J. (2006). Native language predictors of foreign language proficiency and foreign language aptitude. *Annals of Dyslexia*, *56*, 129–160.

Spencer, K. (2007). Predicting children's word-spelling difficulty for common English words from measures of orthographic transparency phonemic and graphemic length and word frequency. *British Journal of Psychology, 98*, 305–338.

Shunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26* (3–4), 207–231.

Sparks, R., Ganschow, L., Kenneweg, S., & Miller, K. (1991). Orthon Gillinham approach to teach a foreign language to dyslexic/learning disabled students: Explicit teaching of phonology in a second language. *Annals of Dyslexia,* 45, 187–214.

Stagg, S. D., & Eaton, E. (2018). Self-efficacy in undergraduate students with dyslexia: A mixed methods investigation. *British Journal of Special Education, 45*(2), 1–29.

Usher, E. L., & Pajeras, F. (2006). Sources of academic and self-regulatory efficacy beliefs of entering middle school students. *Contemporary Educational Psychology, 31*(2), 125–141.

Wermouth, J. P., & Reid, G. (2008). Ocena i planowanie nauczania oraz nauki. [Assessment and Planning of Teaching and Learning]. In G. Reid & J. Wermouth (Eds.), *Dysleksja. Teoria I praktyka* [Dyslexia. Theory and Practice] (pp. 211–234). Gdańskie Wydawnictwo Psychologiczne.

Bogusława Maria Gosiewska-Turek

Legasthenie, Selbstwirksamkeit und Sprachunterricht beim Fremdsprachenlernen – eine gemischte quantitativ-quantitative Studie

Zusammenfassung

Das Ziel der Arbeit ist es, die Wechselbeziehung zwischen Legasthenie, Selbstwirksamkeit und Fremdsprachenunterricht zu untersuchen. Die Autorin der Studie wandte eine gemischte (quantitative und qualitative) Methode an. Die quantitativen Daten wurden anhand der von

legasthenen und nicht legasthenen Studenten ausgefüllten Fragebögen zur Selbstwirksamkeit erhoben und die qualitativen Daten – im Rahmen einer Fallstudie mit einem legasthenen Studenten. Die Forschungsergebnisse im ersten quantitativen Teil der Studie zeigen, dass die Selbstwirksamkeit bei legasthenen Studenten wesentlich geringer sei als bei nicht legasthenen Studenten. Laut Ergebnissen des zweiten gemischten (quantitativ-qualitativen) Teils, der eine Fallstudie umfasste, habe der Fremdsprachenunterricht einen Einfluss auf die Selbstwirksamkeit legasthener Studenten und der entsprechende Sprachunterricht erhöhe die Selbstwirksamkeit von Studenten mit Legasthenie.

Schlüsselwörter: Legasthenie, Selbstwirksamkeit, Fremdsprachenunterricht, Fremdsprachenlernen, Fremdsprachenlehren

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 85–104
https://doi.org/10.31261/TAPSLA.9629

Agnieszka Ślęzak-Świat
University of Silesia in Katowice, Poland
https://orcid.org/0000-0002-0940-0532

Development of Digital Literacy—Translanguaging and Transmedia Note Taking Formats for Academic Reading

Abstract

Generative note taking, being one of the strategies applied to manage difficult texts, requires not only comprehension and selection of information but also production. The current study focuses on note taking formats for a text read with the intention to summarize. Its focal aim is to improve both practical and theoretical understanding of this activity. It involves the investigation into note taking behaviors of 103 second-year English Department students, how they, as readers of FL, engage with complex texts, how they were instructed in note taking and what note taking strategies they employ for comprehending academic texts.

The analysis of the collected data attempts to identify how readers' (n = 103) translanguaging and transmedia (n = 103) note taking formats help increase their engagement in and access to difficult texts in L2. It shows that the subjects have not transitioned from the paper interface to the digital one, since they still display the screen inferiority effect in their reading habits. The collected data shows that only some subjects (n = 42/103) received some form of instruction in paper note taking techniques or digital applications facilitating note taking. The students were not able to enumerate more than four note taking applications which would be conducive to their formation of a coherent interpretation of the digital text they read.

The author contends that overt note taking instruction in both paper and digital mode will create avenues for encouraging, interacting and engaging in reading. Instruction in that field needs to be modified with regard to digital note taking/annotating tools to make use of the note taking formats available for processing digitally interfaced texts.

Keywords: note taking, screen inferiority, reading strategies, note taking applications, translanguaging

Digital Literacy—Instruction

Literacy instruction is under challenge to change because the pedagogies have to be integrated with students' everyday technology practices. Students do not only need to excel in paper but also digitally interfaced texts, drawing information from a text and forming coherent interpretation of it (Grabe & Stoller, 2020). One of the strategies aiding the formation of coherent interpretation of the text is note taking (Muller & Oppenheimer, 2014; Morehead et al., 2019). An analysis of note taking has a potential to illustrate the process of forming coherent interpretation of texts. Long before having access to the digital interfaces, readers' process of paper-based interpretation text was accompanied by physical actions of a reader such as using bookmarks, tracing the text with a finger or pencil, making notes on the margins, that is, annotating, plain scribbling or doodling. Such interpretation facilitating tools appearing as comments on the margins (marginalia) can be traced back to 500 B.C. in the form of scholia (Dickey, 2007), which contained additional clues to the interpretation of the texts that they accompanied.

Now, with education shifting into the online realm, we have to take into consideration Lorenzo and Dzuiban's (2006, p. 2) claims that "students aren't as net savvy as we might have assumed." The problem is that students might not be savvy in interaction with paper text either. Both paper and digital note taking formats need to be overtly taught to foster digital literacy that will support learning and skills, allowing students to manage enormous amounts of information that they have to filter and organize to form coherent text interpretations.

The paper begins with a brief characteristic of the reading purposes and a discussion of the imprecise use of the terms note taking and annotating for reading in both paper and digital interface. Next, the research on the use of translanguaging and transmedia note taking formats as well as the increase in cognitive effort conducive to the engagement in reading a text is reported. Then the collected data is presented and discussed. The conclusions from the present study indicate that students have to be provided with environments in which they can both build knowledge and increase their skill-sets to manage difficult texts, thus teaching implications involve an overview of available digital note taking applications conducive to interpretation of digitally interfaced texts.

Note Taking Techniques in Reading to Integrate Information and Write

Undoubtedly, students must be equipped with strategies to cope with difficult texts (Afflerbach et al., 2008; Grabe, 2009; Chodkiewicz, 2015; Kiszczak & Chodkiewicz, 2019; Grabe & Stoller, 2020). Strategic readers begin with a purpose for reading and recognize that different goals require different types of reading (Schwanenflugel & Knapp, 2016, p. 219) and—consequently varied note-taking techniques. Grabe and Stoller (2020) enumerate the following purposes of reading: (i) reading to search for simple information; (ii) reading to skim quickly; (iii) reading to learn from texts; (iv) reading to integrate information; (v) reading to write (or search for information needed for writing); (vi) reading to critique texts and (vii) reading for general comprehension. All of the enumerated purposes will require generative note taking, however, for the purpose of the present study only reading to integrate and to write will be taken into consideration.

Generative note taking (summarizing, paraphrasing, concept mapping) requires three important activities: comprehension, selection of information and production (Piolat et al., 2004). The major function of taking notes is to gather and transmit information conveyed in a text that needs to be remembered (Armbruster, 2000; Piolat & Boch, 2004). In academic contexts, manipulating and anticipating relevant information are crucial because a note taker has to judge (Middendorf & Macan, 2002) and make decisions on what to prioritize (Castello & Monereo, 1999). Note taking in reading to integrate information and write requires the ability to select, critique, and compose information from a text. Thus, in the case of note taking with the purpose of summarizing the text, notes constitute the first step of the composition, as it requires additional decisions about the relative importance of complementary, mutually supporting, or conflicting information and the likely restructuring of a rhetorical frame to accommodate information from multiple sources (Grabe & Stroller, 2020).

The reader/note taker has to remember points of comparison or opposition, assess the relative importance of the information, construct a framework in which the information will be organized, and establish the main theme (Grabe & Stroller, 2020), thus note taking techniques may take the forms of substitutive techniques like mathematical (=) or iconic (→; ←; ↑; ↓, *), which are used not only to increase the speed of note taking (Piolat et al., 2004) but also to facilitate the hierarchy of items in lists; or to transform the physical formatting of a linear text into special organization of notes (Piolat, 2001). Comments referred to in literature as annotations (Marshal, 1997, p. 132) may take the forms of near or in the text markings, which record interpretive activity as the

result of careful reading. Marshal (1997, p. 134) views them as a visible trace of a reader's attention, a focus on the passing words, and a marker of all that has already been read.

As writers, note takers must select the information to record and format it in ways that differ from the source material (Pilat, 2001). Thus, notes may take the form of marginal jottings and interpolations—being the record of an interpretive activity; highlighting; underlining; circled words or phrases (Marshal, 1998) and help trace the progress through a difficult narrative.

Translanguaging and Transmedia Note Taking Formats

The use of mother tongue or other languages that one knows while taking notes has not been thoroughly investigated, however, Chaudron et al. (1994) and Clerehan (1995), conducted research, showing that some note taking formats are automatized sufficiently to be transferred from one language to another, which might be conducive to their effectiveness due to the fact that the more deeply information is processed during note taking, the greater the encoding benefits (Kiewra, 1985). Unfortunately, due to the omnipresence of foreign language immersion learning programs—deliberately discouraging learners from using languages other than the target language in any activity connected with learning—the potential of L1 in developing L2 competence has been overlooked. Using L1 for note taking purposes has many advantages as it serves as a sheer reference and a straightforward access to the concepts that are already well rooted in the brain, where the memory systems are intertwined to support the learning process. García et al. (2017) indicated the salient purposes for the strategic use of translanguaging in education in general. For the present study, the use of L1 (or other languages that one knows better than the target language—at least in the context of the text read) is of high importance as such a use of the other languages supports the students in comprehension of complex content of texts written in the target language. Second of all, translanguaging provides opportunities for students to develop linguistic practices for academic contexts, and finally it makes space for students' bilingualism and ways of knowing. Vogel et al. (2018) expanded the definition of translanguaging treating the concepts not only as encompassing the linguistic resources individuals draw upon to make meaning, but also as the unique social actions enabled by the use of technology like sharing ideas in social networking and gaming as well as video sharing.

The increasing uses of digital media for information seeking greatly expand the importance of both translanguaging and transmedia abilities needed

to integrate information (Van den Broek & Kendeou, 2017). Readers have to manage texts acquired from multiple both print and online sources adjusting their note taking formats accordingly. As far as research on longhand vs. typed notes is concerned, it still displays screen inferiority effect (Kong et al., 2018; Singer & Alexander, 2017; Ślęzak-Świat, 2019) as it indicates to the fact that that annotation on paper integrates more smoothly with reading than the on-line one (O'Hara & Sellens, 1997). Further neurolinguistic research (Vinci-Booher et al., 2016; James, 2017) corroborates such integration, proving that handwriting connects more visual and motor networks in the brain being conducive to memorization and retrieval of concepts that are written down. Such observations are also in line with the research of Fiorella and Mayer (2017) as well as Luo and colleagues (2018), reporting that there is a greater number of images in longhand notes than the laptop ones. Despite the screen inferiority effect, the issue of digital annotating will have to be attended as most of reading is done on screen generating greater cognitive effort.

Cognitive Effort in Note Taking

The digital culture has fostered immediacy expectation (Perez-Vega et al., 2016), which has led to general problems with focusing and sustaining attention in reading (Salmeron et al., 2018). Nevertheless, human cognitive capacity will have to adapt to accommodate to the rapid digitalization of educational context. Annotations can serve as a visible trace of the reader's attention (Marshal, 1997)—the reader can support their attention by means of note taking when the text is difficult; they can chunk the text into pieces which are easier to interpret— displaying negotiation for meaning strategy. What is more, students report that they prefer reading their own notes because of the change of/in the register of the text into less formal language (Marshal, 1998). Surprisingly the immediacy expectation makes students choose longhand, paper format of their note taking. As Kellogg and Mueller (1993) indicate, writing by longhand is less effortful than using a word processor even for skilled typists and Gérouit and collegues' (2001) research shows that taking notes from a digitally interfaced text is more effortful.

Note takers as readers have to interweave both comprehension and production processes (Piolat, 2005, p. 305). They first need to comprehend information and only then try to store it in the long-term memory by writing it down (Piolat, 2001). Thus, notes might be referred to as an external memory, whose content is more or less explicit (Piolat, 2005, p. 292)—facilitating inferencing, memorizing points of comparison or opposition and functioning as compre-

hension monitoring strategies. Yeung et al. (1997) indicated that note taking viewed as an external working memory is a means to decrease cognitive load during reading. Following the abovementioned view, it can be assumed that the major function of note taking is to capture and preserve information in a form that most conducive for the recall of ideas convey in the text. Siegel (2018, p. 86) defines effective notes as selective, organized and elaborating on ideas expressed by the text authors so that the note taker can learn in generative and constructive ways.

Salmeron and colleagues (2018) note that a new set of advanced reading skills emerges with digitally interfaced reading, including focused searching and navigating of hypertext and multimedia sources as well as integrating multiple sources of information. Skillful digital note taking (annotating and marking the text) would provide a scaffolding for the abovementioned skills and support working memory decreasing cognitive load during digitally interfaced texts.

Description of the Study

The present study attempts to address the following questions: what kinds of note taking training the students received and who delivered the instruction; in what language students recorded their notes; what motivated the students' choice of language of note taking; what the declared and recorded note taking habits of the students were.

The participants of the study were 103 University students aged majoring in English who were 20–23, taking the on-site/on-line course of Academic writing in the summer term of 2020. As regards their learning history, the majority of the subjects (39) had a long English learning history covering the period of 10–15 years. The remaining subjects were placed in two extremes labelled as "less than 10" and "more than 15 years," represented by 25 and 36 students, respectively.

The main areas investigated involve the subjects' note taking habits concerning texts that they need to summarize. The study involved collecting data from a computer-assisted questionnaire at the Moodle platform as well as collection of note taking samples that the subjects were supposed to prepare before writing a summary of an article they chose to work on during a term-long course of Academic writing. The articles of their choice were supposed to comply with the APA style sheets, they were of various length and content as they were chosen according to students' interests. The subjects were allowed to perform the task at their own pace so there was no time pressure involved. The summary was to

be submitted in a common (all study group had access to it) Google Document as one of the assignments required for obtaining the credit.

The subjects in the current study can be considered as experienced but not distinguished note takers as they declare to have been taking all sorts of notes since they were 12 (n = 56) and 16 (n = 29) years old, and those (n = 18) who do not take notes at all. Subjects declaring not to take notes were not really consistent in their statements, as this number declined with the answers provided for further, more detailed questions, for example, in the question about which language they chose for note taking, only eight (n = 8) persisted on the claim that they did not take any notes at all but then they (n = 18) declined to submit the assignment in which they were asked to take notes for the summary they were supposed to write during the next classes.

Out of the subjects who declare to take notes (n = 74), instruction on note taking was received by 42. Instruction was provided by an English language teacher (n = 20); a computer science teacher (n = 5), a YouTube tutorial (n = 5); a Polish language teacher (n = 3), a parent (n = 2), self-study (n = 2); a schoolmate (n = 2); a website (n = 1) and the remaining two indicated others, unfortunately without listing them. As far as instruction of note taking in electronic documents is concerned, only three subjects (n = 3) reported to have received some form of instruction, enumerating the following note taking applications: Evernote, Onenote, Google Keep and Simple Note. None of them mentioned the application allowing for electronic annotating and note taking in Google document that the group has worked on throughout the summer term of 2020, tools like, among others, Stoplight Annotator, Highlight Tool, which are free, easily accessible Google documents add-ons.

Results and Discussion

The results presented and discussed demonstrate two areas of note taking conditions involving translanuaging and transmedia practices. The analysis found evidence for the discrepancy between what was declared by subjects in their questionnaire and what they performed in their notes taken. It is worth discussing these interesting facts in the light of the note taking applications available to digital readers.

Translanguaging Note Taking Practices

As regards the language the notes were taken in, most of the subjects declared to use both L1 and L2 (n = 67). The group subdivides into those who conditionally use either L1 or L2 (n = 52) depending on their goals, and those who mix the two languages. Table 1 shows the categories of conditions of language choice depending on the goal of the note taker.

Table 1.

The conditions of language chosen for notes taken to summarize a text in L2

I take notes in L1 (Polish) if …	I take notes in L2 (English) if …	Number
I have to memorise a lot of information quickly	*I have ample of time to study*	21
I need deeper explanation	*I take general notes*	17
I want to understand the text better	*I want to see a particular phrase in context*	15
I can't find a word in English	*I need full, proper definition of a word*	6
There is a word I don't know	*I make general notes*	3
I find the word useful	*There is no good translation*	2

As far as the condition for L1 choice for note taking is concerned, the goal-oriented group can be categorized into representing three most salient choices: the selection of L1 is determined by effective memorization of information and shortage of time (n = 21) in contrast to the abundance of time for the use of L2; issues relating to detailed explanation (n = 17) and general note taking, and comprehension improving/monitoring (n = 15) in contrast to the need of contextualizing a given word/phrase. The collected results indicate the fact that the subjects strategically use translanguaging to take notes in a most efficient and time-saving way.

Those who opportunistically mix the two languages (n = 15) substantiate for their choice with the following arguments:

— opting for the language that is more conducive to memorization (n = 6): *Polish or English depending on which of them is easier to memorize*;
— their need for knowing the equivalents in both languages (n = 4): *I like to know equivalents in both languages*;
— linguistic economy (n = 2): *I mix English and Polish to make my notes shorter*;
— convenience (n = 1): *in whichever language it's easier*;
— immediacy of registering the idea (n = 1): *in language the idea comes to me first*;

— for thorough comprehension (n = 1): *in any language that allows me to understand it correctly.*

The group of subjects opting only for L2 notes when reading a text in L2 provided the following categories of arguments, which overlap with those given by the abovementioned groups. As far as the group declaring to use only L2 for their notes, they (n = 24) divided into the following categories:

— convenience (n = 8): *it is easier to take notes in English*;
— improved comprehension (n = 2): *because they are easier to comprehend in the same language*;
— consistency, being further subdivided into:
  - read in L2—think in L2 (n = 7): *the text is in English, so my thoughts are in English when reading it*;
  - L2 text—L2 notes (n = 4): *if a given text is in English then it's easier to make notes in the same language*;
  - L1 would be confusing (n = 2): *I think that taking notes in Polish would confuse me*;
  - L2 summary L2 notes (n = 1): *because it is easier to summarise the main topic if it is in the same language.*

Just like in the groups conditionally and opportunistically using translanguaging, in the group declaring to use Polish only, it is done mostly for facilitating purposes as the subjects (n = 4) claim that*: if I find the text difficult—polish notes help me understand the text* (original spelling); *to translate difficult vocabulary*; *the meaning is not clear enough*; *because it helps to focus on the main points.*

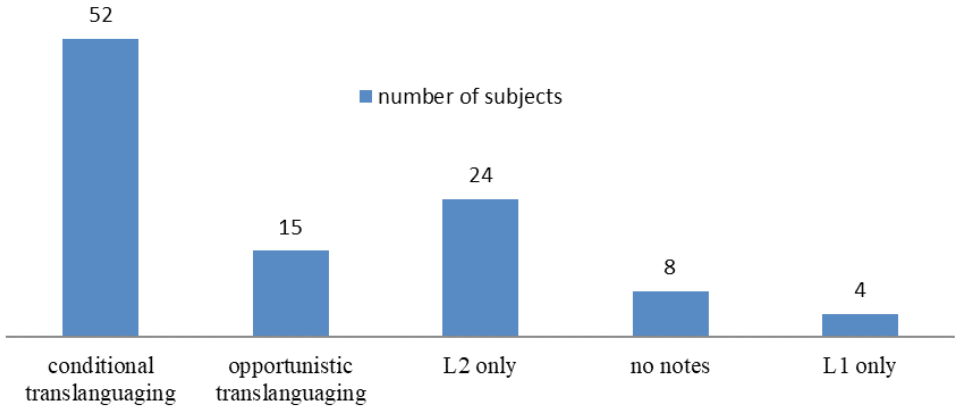


Figure 1. Language chosen for note taking (annotating)

Interestingly enough, out of those who initially declared not to take notes (n = 18) in the question concerning the use of L1 or L2 language for notes

only (*n* = 8) of them persisted on not taking notes at all (*n* = 4), claiming that it was unnecessary (*n* = 1), they did not see a point in taking notes (*n* = 1). Two of them (*n* = 2) declared that they do not take any notes apart from new vocabulary which allowed for determining the inconsistencies in their answers.

Transmedia ote Taking Practices

The choice of note taking strategy starts with the decision of printing the text to be read (*n* = 75) and only 28 subjects decided not to print the article they were required to summarize.

The group of subjects who decided to print the text emphasized first of all the physical aspect (i.e., eye fatigue, touch of paper, ease of navigating the text) of a paper copy that was important for them (*n* = 21): *I find it easier to work with a text when I hold it physically*, then general preference (*n* = 14), speed of taking notes (*n* = 13): *It's much faster to scribble something down than to open a program, choose a tool, and THEN scribble*; convenience (*n* = 12); conducive to concentration (*n* = 10): *It's easier for me to follow the text on paper*; improving comprehension (*n* = 7): *It helps me to understand better*; conducive to memorization (*n* = 4): I find it more effective to write right next to the tasks and writing helps me memorize; more organized (*n* = 1): *much less hassle.*

The answers provided by the group of subjects who did not decide to print (*n* = 28) can be categorized into those ecologically oriented (*n* = 6): *I'd rather go green*; *Paper is made from dead trees*; *Paper saving*; *I am eco-friendly* and those (*n* = 6) for whom the PDF document is enough. The next category with answers relating to being more technology oriented (*n* = 5) claims: *I think that it is high time to start working while using technology*; *I prefer digitalized version as my handwriting is bad and I write faster on a keyboard*, and there is a group of subjects who have an electronic device allowing for paper-like experience when taking notes (*n* = 4): *I have an electronic version on my iPad*. As in the previous groups, there is a group of subjects choosing an electronic note taking format out of convenience (*n* = 4) *Google Doc work became easier*, and general, not substantiated, preference (*n* = 3).

After having collected 85 note taking samples (18 subjects consistently declined taking notes, they even asked to do it as an assignment required for a credit) the notes can be divided into those generated on a separate piece of paper, printouts, e-text (pdf, print screen, referencing mode), and word-processing document (Microsoft Word). An interesting phenomenon is that only ten subjects (7.5%) out of those who decided to print it out (*n* = 75) decided to take notes on their printouts. It illustrates the fact

that note taking is strategic and purpose driven, as subjects adjusted the preferences to the purpose of writing a summary which had to be typed. Contrary to the declared reading preferences, most subjects (n = 33) decided to take notes in an electronic form. Figure 2 shows the distribution of the subjects' choices across the interface the subject chose for note taking. Neither the digitally-oriented nor the Word processing groups chose any of the enumerated note taking applications, neither of them chose the format of Google Document that is automatically supplemented with note taking applications such as Stoplight Annotator, Highlight Tool, MindMeister, and Lucidchart.

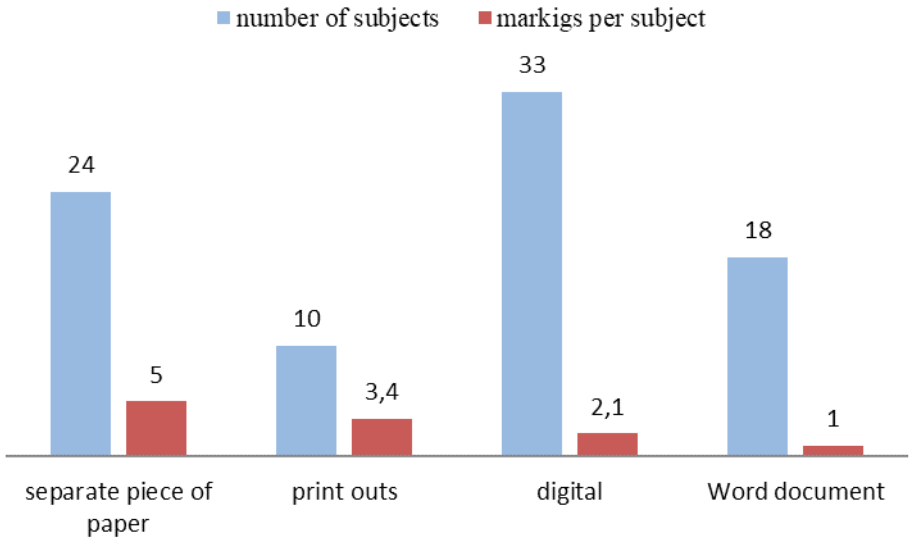


Figure 2. Interface chosen for note taking and average number of markings for a subject in a chosen interface.

As far as the diversity of the formats of markings in note taking is concerned, the greatest number (n = 119) was generated by the group who chose to take notes on a separate piece of paper then the digitally-oriented subjects followed with (n = 70), and those who chose to take notes on printouts (n = 34) and the least note-taking marking was registered in the group who took their notes on a word processed document (n = 18). It turns out that the most note-taking, flexible, and generative—as far as markings are concerned—were those who chose the paper interface for their note taking with average of four markings for a single note taker.

As for the formats of markings, they involve: paraphrase as a comment, character change (size, color, format of fonts), arrows (indicating the relationships between the concepts), lists, underlining, key words, color underlining,

highlighting, non-linear (e.g., change of writing from horizontal into vertical), graphic representation, mathematical symbols, exclamation marks and circling. Thus, when it comes to the registered formats of the marking (n = 13) used in the notes, the group who chose a separate piece of paper for their note taking medium was the largest (n = 13), which was followed by the printout (n = 9), word processing (n = 6) and with electronic (n = 4) as the last one. It adds up to the paper superiority phenomenon not only in reading but also in note taking, showing that it allows for grater creativity and less linearity of the note taking formats allowing to express concepts in radiant mind mapping fashion that allows for the categorization of the concepts presented in a linear text to become areas and allowing to prepare the conceptual map of the processed text Table 2 presents the formats of note taking marking for a given preference group.

Table 2.

Categories of note taking markings for a given note taking preference group

Markings	Separate piece of paper	Printouts	Digital	Word document	Total
paraphrase	24				
paraphrase as a comment		6	23	9	62
character change	17	3		2	22
arrows	14	4	2	2	22
lists	15	2		2	19
underlining	6	6	6		18
key word	12	3		2	17
underlining colour	10	3	2		15
highlighting	6	5		1	12
non-linear	6				6
graphic representation	6				6
mathematical symbol	1				1
exclamation mark	1				1
circling		1			1
Total	118	33	33	18	202

As Table 2 shows, a plain piece of paper allows for the greatest number of operations and manipulations of the ideas the note taker as a reader wants to present. As a result, it generates more engagement in the text. Paraphrasing, in all of the preference groups, is the most often used marking format of reference to the text. The next two in popularity are character changing and the use of arrows. They seem to be like the posts directing the note taker's

attention to the issues of interest, showing the relationships and connection between facts—they are also used together with lists (which were also very highly applied markings), allowing for hierarchical representation of the text. The next two important formats of note taking are key words and underlining with coded colors.

It is important to note here that when subjects were asked in the questionnaire about in what ways they indicated points of importance in the text, the answers they provided were not overlapping with the ones that were registered during their actual performance. The five categories that they enumerated were highlighting (n = 73), underlining (n = 48), paraphrasing (n = 33) and circling (n = 18). The abovementioned results highlight that students know little about note taking formats and their use is more intuitive than strategic. Thus, increasing awareness of the possible note taking formats will lead to more skillful and effective use of them, resulting in an improved ability to select, critique, and compose information from a text.

As far as the content of the comments provided near and within the text is concerned, most of the registered comments involved paraphrases of the text read (n = 25) identifying relationships to other concepts (n = 8) and key words (n = 5), which illustrate the interpretive activity as the result of careful reading and indeed show traces of a reader's attention.

Note Taking Applications

Once subjects craft commenting, the following note taking applications can be recommended: Evernote, Microsoft OneNote, Google Keep, and Simplenote. Their functionality and effectiveness shows best on a shared screen illustrating how it can be applied and tailored to the needs of paper-oriented subjects. Mastering their functionality may help in the transition from paper to digital note taking, preventing the screen inferiority effect. The four note taking applications enumerated by subjects, Evernote, Simplenote, Microsoft OneNote, and Google Keep are free applications whose functionality would cater for diverse needs and preferences of note takers.

When outlining the functionalities of the abovementioned note taking applications, it is Evernote which is the first and the most popular one. It is an application allowing for saving web pages for offline use as well as creating notes and tags. Notes can be accessed on laptops, mobile devices and through the web. It supports a wide variety of note types (text, images, audio memo, sketches, scanned documents, checklists, and clipped web pages). It also has tools for organizing and searching notes as well as its search text function in images. It can constitute a powerful note taking tool for subjects who rely on the use of key words which are a popular note taking tool in the present study

among paper-oriented subjects and hardly used among those taking notes digitally. Both groups would benefit from an overt training in tagging (key word) functionality, which is also available in Simplenote, which is accessible across devices for working on text only. It has simpler interface than Evernote, which makes it easy to keep track of notes and tags.

Then, having in mind that most of the subjects in the present study are paper-oriented in their note taking, Microsoft OneNote is a note-taking application that mimics paper and can help in transition to electronic note taking for those who are paper-oriented. Creating a new note involves clicking anywhere on the page and adding content to that spot, just as if working with paper. For the note takers relying on non-linear graphic representations of their track of thought as well as those who use a lot of arrows this is the application tailoring to their needs as sketches can be drawn. Note takers who use a lot of color codes will also benefit from it as a background for notes looks like textured or lined paper. A text can be typed and images and file attachments can be dragged and dropped into notes. For note takers basing their notes on highlighting, there is a digital highlighter; those who like to create lists (that was a very often used marking in paper oriented subjects) can easily create checklists in OneNote. As each note is meant to appear like a piece of paper, it can be moved around the page, placing a sketch memo next to a block of text. There is one feature which undoubtedly may be appealing to both paper and digitally oriented note takers, namely, optical character recognition (OCR) that can make all the writing searchable. It is important to highlight that for every note there is a record of its version history and there is an ink-to-text feature lets handwritten text be converted to type.

Finally, the simplest in use because of moderate functionalities is Google Keep. Its interface has a form of digitized Post-it Notes (there are 12 bright colors for each note that can be categorized). Note taking is done by typing, drawing, or adding an image. It is used as the Google Keep Chrome extension, URLs, text, and images can be saved while browsing the web. Everything that is saved in Google Keep stays synced across all platforms. The most outstanding feature of electronic note taking is the possibility of having them recorded and searched through in a systematic way.

Apart from applications working independently from a browser, there are also such that function as extensions to browsers (Liner, Weava Highlighter, Super Simple Highlighter, Multi-highlight, Yellow highlighter pen for web) or Google Document add-ons. They allow for highlighting the content of web pages and tagging them with the key words. Google Documents, which were used by the subject of the study, is accompanied with a number of applications that can be downloaded, for example, Stoplight Annotator (simple commenting tool), Highlight Tool, MindMeister (allows for mind map like note taking),

Lucidchart (for those who take notes in list-like fashion). Unfortunately, none of the subjects used them. Figuratively speaking, it is as if using only a pencil having a pen case full of other writing utensils and never using or just trying them out.

Conclusions and Teaching Implications

The results of the present study confirm that note taking formats are individual "writing signatures" (Van Waes & Schellen, 2003) and that "more heterogenous view of taking notes" would be beneficial for learners of English (Badger et al., 2001, p. 406). Notes are idiosyncratic signatures, however, signing requires knowing how to write. Knowing how to write requires training, which is planned and controlled. To develop handwriting, hours must be spent on tedious, repetitive exercises and, likewise, in the case of developing note taking, marking techniques would be conducive to the development of students' abilities to select, critique, and compose information from the text. Thus, students must be offered a range of opportunities to choose from so as to tailor it to their needs.

As far as the use of L1 in note taking is concerned, the obtained results show what potential it offers. The subjects' translanguaging practices provide fluid connections between the learned concepts without narrowing students' range of thought. In the case of note taking, the richness of information that a student is exposed to and the speed and reliability of the note that are being made are important. A particular piece of information that is being recorded in notes is to trigger memories in the form of words which facilitate recall. Translanguaging practices provide students with transitions they make between what they know and what they are yet to master. Unfortunately, these are only translation programs which focus on L1; however, the results obtained indicate to their encoding (improved memorization) and better comprehension benefits, confirming that the more information is processed and manipulated during note taking, the greater the encoding and organization benefits for the generated summaries in terms of integration of conceptual items expressed by specialized academic vocabulary.

Now, having so many applications available, the choice of them constitutes individual signature. Note taking while reading might be regarded as hyperlinking the text to the note taker's ways of knowing, which allows for constructing means by which new information is integrated with the existing knowledge and personalizing the text in a way that is meaningful to the reader. Lack of overt instruction on how to take notes presents possibilities of improving literacy in

general, be it paper of digitally interfaced. In the case of digital reading, such training will improve in general using, evaluating, and managing digital texts. To facilitate learning, technology needs to support it in authentic ways. Thus having collected data on the subjects note taking formats, building instruction on note taking can be built on what students already know without imposing on them solutions that would not be practical or feasible for them and for lecturers/teachers.

Instructing how to take notes, that is, decomposing texts into smaller components, has to be taught by providing a scaffolding for assigning significance to information processed. Such scaffolding can be provided by, for example, Annotation Studio (www.annotationstudio.org), which is an open source web application with commenting tools immediately accessible to students and lecturers. It facilitates the process of visualization of the readers' approach to texts in the context of commenting it on the screen as both the lecturer and students can see the comments made. The visualizations provided by the application show instructors which passages generate most interest or difficulties.

Yet another tool for social digital reading and commenting is *eComma*, which is a plug-in that works with most learning management systems such as Canvas, Blackboard or Moodle, it allows a group of users to annotate the same text together and to share their annotations with each other. Highlights can overlap. To distinguish which note corresponds to which highlighted passage, both light up when either is under the mouse cursor. If a passage of text corresponds to more than one note, both light up. Each annotation is associated with a specific username.

Considering a forced and accelerated transition of education into the digital realm, the collected data indicate that students' shift into digital note taking is not as rapid as could be expected. There was hardly any shift in students' note taking habits, which means that pedagogies must be focused on facilitating the transition to digital interface to model the effective use of different media platforms and teaching how to leave one's own trace on them to form a coherent interpretation. Thus, overt training on note taking on texts, images, and videos must be done to train students on dealing with performing reading tasks in digital media. Future research should consider the potential of note taking applications in developing not only digital but also general literacy.

References

Afflerbach, P., Pearson, P. D., & Paris, S. G. (2008). Clarifying differences between reading skills and reading strategies. *The Reading Teacher, 61*, 364–373.

Annotation Studio (2017). www.annotationstudio.org

Armbruster, B. B. (2000). Taking notes from lectures. In R. F. Flippo & D. C. Caverly (Eds.), *Handbook of college reading and study strategy research* (pp. 175–199). Lawrence Erlbaum Associates.

Badger, R., White, G., Sutherland, P., & Haggis, T. (2001). Note perfect: An investigation of how students view taking notes in lectures. *System, 29*, 405–417.

Castello, M., & Monereo, C. (1999). Strategic knowledge in note-taking: A study in high education. *Infancia y Aprendizaje, 88*, 25–42.

Chaudron, C., Loschky, L., & Cook, J. (1994). Second language listening comprehension and lecture note-taking. In J. Flowerdew (Ed.), *Academic listening: Research perspectives* (pp. 75–92). Cambridge University Press.

Chodkiewicz, H. (2015). Some insights into the academic reading experience of English philology students. *Konin Language Studies, 3*(1), 9–30.

Clerehan, R. (1995). Taking it down: Note taking practices of L1 and L2 students. *English for specific purposes*, *14*(2), 137–157.

Dickey, E. (2007). *Ancient Greek scholarship: A guide to finding, reading, and understanding scholia, commentaries, lexica, and grammatical treatises*. Oxford University Press.

Fiorella, L., & Mayer, R. E. (2017). Spontaneous spatial strategy use in learning from scientific text. *Contemporary Educational Psychology*, *49*, 66–79.

García, O., Johnson, S., & Seltzer, K. (2017). *The translanguaging classroom. Leveraging student bilingualism for learning.* Caslon.

Gérouit, C., Piolat, A., Roussey, J. Y., & Barbier, M. L. (2001). Coût attentionnel de la recherche d'informations par des adultes sur hypertexte et sur document papier. In M. Mojahid & J. Virbel (Eds.), *Actes du 4 Colloque International sur le Document Electronique* (pp. 201–215). Europia production.

Gross, M. D. (1996). The electronic cocktail napkin-computer support for working with diagrams. *Design Studies, 17*, (1), 53–69.

Grabe, W. (2009). *Reading in a second language: Moving from theory to practice.* Cambridge University Press.

Grabe, W. & Stroller, F. L. (2020). *Teaching and researching reading.* Routledge.

Igo, L. B., Bruning, R., & McCrudden, M. T. (2005). Exploring differences in students' copy-and-paste decision making and processing: A mixed-methods study. *Journal of Educational Psychology, 97*, 103–116.

James, K. H. (2017). The importance of handwriting experience on the development of the literate brain. *Current Directions in Psychological Science, 26*, 502–508.

Kellogg, R. T., & Mueller, S. (1993). Performance amplification and process restructuring in computer-based writing. *International Journal of Man-Machine Studies, 39*, 33–49.

Kiewra, K. A. (1989). A review of note-taking: The encoding storage paradigm and beyond. *Educational Psychology Review, 1*, 147–172.

Kiszczak, A., & Chodkiewicz, H. (2019). Text-based student questioning in EFL settings: Long-term strategy implementation in reciprocal reading tasks and its perception. *Theory and Practice of Second Language Acquisition, 5*(2), 39–57.

Kong, Y., Seo, Y. S., & Zhai, L. (2018). Comparison of reading performance on screen and on paper: A meta-analysis. *Computers & Education, 123*, 138–149.

Lorenzo, G., & Dziuban, C. (2006). Ensuring the net generation is net savvy. *Educause Learning Initiative*, ELI paper 2.

Luo, L., Kiewra, K. A., Flanigan, A. E., & Peteranetz, M. S. (2018). Laptop versus longhand note taking: Effects on lecture notes and achievement. *Instructional Science*, *46*, 1–25.

Marshall, C. (1997). Annotation: From paper books to the digital library. In *Proceedings of the ACM Digital Libraries '97 Conference* (pp. 131–140). Philadelphia, PA.

Marshall, C. (1998). Toward an ecology of hypertext annotation. In *Proceedings of ACM Hypertext '98* (pp. 40–49). Pittsburgh, PA.

Middendorf, C. H., & Macan, T. H. (2002). Note-taking in the employment interview: Effects on recall and judgments. *Journal of Applied Psychology,* 87(2), 293–303. https://doi.org/10.1037/0021-9010.87.2.293

Morehead, M., Dunlosky, J., & Rawson, K. A. (2019). How much mightier is the pen than the keyboard for note-taking? A replication and extension of Mueller and Oppenheimer 2014. *Educational Psychology Review, 31*, 753–780.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: advantages of longhand over laptop note taking. *Psychological Science, 25*, 1159–1168.

O'Hara, K., & Sellen, A. (1997). A comparison of reading paper and on-line documents. In *Proceedings of CHI'97* (pp. 335–342). Atlanta, Georgia (March 22–27, 1997).

Perez-Vega, R., Waite, K., O'Gorman, K. (2016). Social Impact Theory: An examination of how immediacy operates as an influence upon social media interaction. *The Marketing Review, 16*(3), 299–321.

Piolat, A. (2001). *La prise de notes* [Note Taking]. Presses Universitaires de France.

Piolat, A., & Boch, F. (2004). Apprendre en notant et apprendre à noter. In E. Gentaz & P. Dessus [Learning by Taking Notes and Learning to Take Notes] (Eds.), *Comprendre les apprentissages. Psychologie cognitive et éducation* [Understanding Learning. Cognitive Psychology and Education] (pp. 133–152). Dunod.

Piolat, A., Kellogg, R. T., & Farioli, F. (2001). The triple task technique for studying writing processes: On which task is attention focused? *Current Psychology Letters. Brain, Behavior and Cognition, 4*, 67–83.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291–312.

Salmerón, L., Strømsø, H. I., Kammerer, Y., Stadtler, M. & van den Broek, P. (2018). Comprehension processes in digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 91–210). Benjamins.

Sanchez, R. P., Lorch, E. P., & Lorch, R. F. (2001). Effects of headings on text processing strategies. *Contemporary Educational Psychology, 26*(3), 418–428.

Schwanenflugel, P., & Knapp, N. (2016). *The psychology of reading: Theory and applications*. Guilford Press.

Shulmann, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review, 57*(1), 1–22.

Siegel, J. (2018). Did you take "good" notes?: On methods for evaluating student note taking performance. *Journal of English for Academic Purposes, 35*, 85–92.

Singer, L. M., & Alexander, P.A. (2017). Reading on paper and digitally: What the past decades of empirical research reveal. *Review of Educational Research, 87*(6), 1007–1041.

Ślęzak-Świat, A. (2019). Complementarity of reading from paper and screen: Development of critical thinking. *Theory and Practice of Second Language Acquisition, 5*(2), 75–93.

Van den Broek, P., & Kendeou, P. (2017). Development of reading comprehension. In K. Cain, D. Compton, & R. Parilla (Eds.), *Theories of reading development* (pp. 283–305). Benjamins.

Van Waes, L., & Schellens, P. J. (2003). Writing profiles: The effect of the writing mode on pausing and revision patterns of experienced writers. *Journal of Pragmatics, 35*, 829–853.

Vinci-Booher, S., James, T. W., & James, K. H. (2016). Visual-motor functional connectivity in preschool children emerges after handwriting experience. *Trends in Neuroscience and Education, 5*, 107–120.

Vogel, S., Ascenzi-Moreno, L., & García, O. (2018). What counts as a literacy act? Examining pedagogical spaces created through digital translanguaging. In J. Choi & S. Ollerhead (Eds.), *Plurilingualism in teaching and learning: Complexities across contexts*. Taylor & Francis.

Wylie, J., Thomson, J., Leppanen, P., Ackerman, R., Kanniainen, L., & Prieler, T. (2018). Cognitive processes and digital reading. In M. Barzillai, J. Thomson, S. Schroeder, & P. van den Broek (Eds.), *Learning to read in a digital world* (pp. 57–90). Benjamins.

Agnieszka Ślęzak-Świat

Die Entwicklung von digitalen Kompetenzen – sprach- und medienübergreifende Notiztechniken für akademisches Lesen

Zusammenfassung

Generatives Notieren, eine der Strategien zur Bewältigung komplizierter Texte, erfordert nicht nur das Verstehen und die Selektion von Informationen, sondern auch die Produktion. Die vorliegende Studie befasst sich mit Notiztechniken für einen Text, der mit der Absicht gelesen wird, ihn zusammenzufassen. Ihr Hauptziel ist es, sowohl das praktische als auch das theoretische Verständnis der Tätigkeit zu verbessern. Sie umfasst die Untersuchung der Art und Weise, wie 103 Studierende des zweiten Studienjahres im Fachbereich Englisch Notizen erstellen, sich als fremdsprachige Leser mit komplexen Texten auseinandersetzen und in Notiztechniken unterrichtet worden sind sowie der Strategien, die sie zum Verstehen von akademischen Texten anwenden.

Ziel der Analyse der erfassten Daten ist herauszufinden, wie sprach- und medienübergreifende Notiztechniken den Lesern (n = 103) dabei helfen, ihr Engagement für bzw. Verständnis von komplizierten fremdsprachigen Texten zu verbessern. Es zeigt sich, dass die Probanden vom Papiernotizbuch zu digitalen Notizen nicht übergegangen sind, weil der Bildschirm in ihren Lesegewohnheiten immer noch eine inferiore Stellung hat. Die erhobenen Daten weisen darauf hin, dass nur ein Teil der Probanden (n = 42/103) in irgendeiner Form in Notiztechniken auf Papier oder digitalen Anwendungen zur Erleichterung des Notierens unterwiesen worden ist. Die Studenten waren nicht imstande,

mehr als vier Notizanwendungen zu nennen, welche für die Erstellung einer kohärenten Interpretation des von ihnen gelesenen, digitalen Textes förderlich wären.

Die Autorin der Arbeit behauptet, dass ein offener Unterricht in Notiztechniken, sowohl auf Papier als auch in digitaler Form, zur Förderung, Interaktion bzw. zum Engagement beim Lesen beitragen würde. Der Unterricht müsste im Hinblick auf digitale Notizanwendungen bzw. Anmerkungswerkzeuge entsprechend modifiziert werden, um die für Bearbeitung von digitalen Texten verfügbaren Notiztechniken in Anspruch zu nehmen.

Schlüsselwörter: Notieren, Unterlegenheit des Bildschirms, Lesetechniken, Notizanwendungen, Translanguaging

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 105–127
https://doi.org/10.31261/TAPSLA.11514

Ayalew Tilahun
Bahir Dar University, Ethiopia
https://orcid.org/0000-0003-1564-3044

Berhanu Simegn
Bahir Dar University, Ethiopia
https://orcid.org/0000-0001-6192-9858

Mulugeta Teka
Bahir Dar University, Ethiopia
https://orcid.org/0000-0002-9786-8871

Investigating Effects of Integrated Reading and Writing Skills Instruction on Enhancing Students' Critical Thinking Skills in EFL Classroom

Abstract

The primary concern of this study was to investigate the praxis of integrated reading and writing skills instruction on EFL learners' critical thinking development at Bahir Dar University. The study used a pre-test–post-test quasi-experimental design, and 96 English majoring students randomly assigned to (n = 48) experimental and (n = 48) control group took part in the study. Herein, the control group was instructed through a separated reading and writing approach and the experimental group learned through a newly designed integrated reading and writing way for 12 weeks concurrently with three sessions per week, and then, 25 pre-and post-tests of critical thinking questions were designed to assess students' critical thinking development. Here, Kappa inter-rater and split-half reliability tests were employed to compute the reliability and internal consistency of both tests, respectively. Finally, an independent t-test was employed to compute the data, and then the results revealed that both the control and experimental group were homogeneous regarding their level in the pre-tests of critical thinking skills. However, after the treatment, the study showed the supremacy of integrated reading and writing skills instruction over the conventional approach in enhancing students' critical thinking skills.

Keywords: integrated reading and writing instruction, language pedagogy, reading skills, writing skills, critical thinking skills

Introduction

English is functional as a foreign language in Ethiopian context. As a result, it is not working as a viable Lingua Franca in everyday lives of the people outside of the EFL classroom. Thus, in the country, the opportunity of learning and practicing English outside the classroom is less frequent, and for this reason, particularly at the university level, students more often learn the language through reading and writing in their academic careers. Therefore, the instructional methods that the EFL teachers use in reading and writing would play vital roles in students' language and reasoning development, for the reason that reading and writing skills are two of the major skills that need to be taught efficiently to enable students to comprehend texts in a critical way. Consequently, reading analytically, composing text coherently and critical thinking skills are supposed to be the basis for the success of any university student, for these are very vital cognitive fundamentals that the learners must develop to explore the sphere of knowledge.

The contemporary thinking about the nature of reading and writing views the two skills as interdependent and transactive (Carson, 1993; Spack, 1998; Rossenblatt, 2004). Furthermore, these researchers point out that meaning is created through the transaction between the reader, the writer, and the text. Therefore, to equip students with the required capabilities, the teaching and learning activities which are employed in the university setting, particularly in EFL classrooms, should focus on enabling students to interact with a text through the process of integrating reading and writing activities because when learners involved in both analytical reading and writing skills in integrated ways, they also build up their critical thinking ability in their academic career. In this respect, researchers claim that reading extensively can improve students' writing abilities, and in the process of writing, the logical thinking abilities will be enhanced (Cavdar & Doe, 2012). Here, instructors can no longer be information givers and students must learn thinking and reasoning skills to reach their fullest potential, and this can be done explicitly and directly in an integrated manner (Cobine, 1995; Fisher, 2001). Thus, involving students in the above activities encourages them to use the language for thinking rather than studying about the language. In a similar way, Facione (1990) asserts that "involving students in critical thinking activities helps them to become more mentally dynamic and decisive rather than passive receivers of contents in the classroom" (p. 3).

When reading and writing are integrated into classroom instruction, it enables students to think about effective comprehension strategies and engage in constructing meaning from a given genre. In line with this, researchers also claim that skill training and critical thinking can be reached simultaneously

by integrating reading and writing skills in ESL/EFL classrooms (Richards & Rodgers, 2001; Rosenblatt, 2004; Li & Yang, 2014). In doing so, a reader tries to use higher cognitive skills to describe, respond, or interpret a reading text, and then a new interpreted text is produced in the process. Under these strategies, learners would be trained to organize their thoughts through writing. Likewise, the integration of reading into writing enables students to develop both critical thinking and critical literacy because it augments their ability to transform information for their own purposes in reading and to blend their prior awareness with another text in writing. Furthermore, put differently, both reading and writing skills are so closely linked, they mutually reinforce each other, and therefore, promote learning when they are integrated into classroom activities (Cobine, 1995). By the same token, Atkins et al. (1996), Cavdar and Doe (2012) all confirm that integrating the skills not only develops students' ability to express themselves, but it also develops students' thinking power—they get mental training.

In the same vein, students become better readers, writers, and thinkers when they learn reading and writing together (Carson, 1993; Rosenblatt, 2004). Similarly, Fisher (2001) asserts that "students must develop thinking and reasoning skills to reach their fullest potential, and this can be done explicitly and directly in an integrated manner" (p. 17). Put briefly, making students engage in critical thinking helps them to maximize their involvement with the activities because critical thinking is an intellectually disciplined process of actively and skillfully conceptualizing, applying, inferring, interpreting, synthesizing, evaluating, reasoning activity during reading or writing communication (Paul & Scriven, 2008, p. 25). With this in mind, involvement in the above activities allows students use the language for thinking rather than studying about the language. Therefore, to make students fully engage in critical thinking, learners should get involved in both analytical reading and composing skills based on their reading. As students' knowledge and critical thinking develop through reading, their internal motivation to produce their own ideas through writing will also be enhanced. Furthermore, various scholars suggest that reading and writing skills are so closely linked that they mutually reinforce each other, and therefore promote learning when they are integrated into classroom activities (Alghonaim, 2018; Pysarchyk & Yamshynska, 2015; Al-Dosari, 2016; Eun-Hee & Myeong-Hee, 2020).

Thinking critically, as an essential part of the higher-order cognitive skills and the most important component of undergraduate EFL education, includes many other skills of higher-order mental processes like reasoning, reflective thinking, analysis, synthesis, inference, evaluation and decision making, creative thinking, and self-regulation. Thus, in relation to this, Elder and Paul (2009) recommend that every classroom activity with the aim to develop learners' critical thinking abilities should lead to an environment that enhances

these basic skills in students' learning. Subsequently, in the light of the above view, integrated reading and writing instruction is where developmental reading and writing are taught in one course within a reduced period of time. In brief, all the above research claims imply that separating reading and writing skills in EFL classroom instruction has its own detrimental effects on enabling learners to make connections between these two skills in comprehending texts and critical thinking development.

Therefore, researchers and theorists recommend examining assumptions and reviewing educational practices, as current pedagogical thinking seems to be shifting away from traditional behaviorist model of teaching to constructivist views of learning whereby teaching is seen as transformative (Brandis & Ginnis 1986, as cited in Richards & Rodgers, 2001, p. 197). Furthermore, these researchers also suggest that skill training and critical thinking can be reached simultaneously by integrating reading and writing skills in ESL/EFL classrooms. Similarly, Graham and Hebert (2010) argue that "understanding a text encompasses writing about it; writing about a text should augment comprehension, for it provides learners with a tool for visibly and permanently recording, connecting, analyzing, personalizing and manipulating key ideas in text" (p. 13). These pedagogical positions range from the traditional skill-based, text-driven models to more holistic, process-oriented approaches associated with integrating the language arts. In the same way, Ferire (1984) and DuBrowa (2011) state that transformative and real literacy learning takes shape when the learner takes part in reading and writing activities and participates in the real-life insightful process. When the two language skills are integrated, it augments deep and careful thinking in meaningful ways. Furthermore, the integration of reading into writing enables students to develop both critical thinking and critical literacy that promote students' ability to transform information for their own purposes in reading and to develop text in writing.

With all this in mind, integrated reading and writing instruction is widely discussed in the ESL and EFL classrooms in a foreign context. However, in Ethiopia, though theoretically discussed, it is not practically implemented in the EFL classroom. For instance, in Bahir Dar University English courses like Basic writing skills and Communicative language skills are designed in an isolated skill approach. Furthermore, in secondary and preparatory English textbooks, reading and writing skills activities are presented on a separate basis. Why is that so? Is integrated reading and writing skills instruction not effective in the Ethiopian context? Can't it be practiced in the Ethiopian context? Therefore, testing integrated reading and writing skills instruction empirically plays a vital role to see its effectiveness for building critical thinking skills in the context of the university (undergraduate) education in Ethiopia.

Statement of the Problem

In the context of EFL in Ethiopia, the curriculum and pedagogy of writing and reading courses have traditionally been designed on a separate basis. As a result, more emphasis is given to skill-based language activities. Moreover, tasks and activities have been designed usually focusing on one specific skill. For instance, in reading instruction, more emphasis is given to answering reading comprehension questions, vocabulary meaning, references, and so on. Indeed, these sorts of activities are important particularly for language practice, but their contribution to enable students to go beyond the literal meanings of the text is reduced. As a result, most learners neither understand the text fully nor give critical explanations about the text. Furthermore, in writing instruction classrooms, most often, writing lessons focus on writing activities on prescribed topics, and attention is too often paid to modeling correct grammatical and essayist forms instead of creating conducive environment for students to interact with language actively for authentic communicative purposes. Moreover, after students' writings, EFL instructors' main focus would be on correcting different kinds of writing mechanics like spelling errors, punctuation, wrong use of words, phrases, tenses and other related problems. Likewise, hand-written texts are evaluated on the accuracy of grammar rather than on content, style or creative expressions of ideas. As a result, students cannot get critical mental exercises through writing.

The above instructional approaches are very much emphasized in the traditional EFL teaching and curriculum provisions. Researchers like Taizad and Namaghi (2014) state that although segregated skills teaching may help students develop their knowledge of the language, its impact on enabling learners to use the knowledge in actual communication is less likely. Tsai (2006) also claims that the separate teaching of reading and writing skills approach is an obstacle in fluent learning of a foreign language. Likewise, Oxford (2001) states that segregated teaching emanates from the philosophy that sees successful L2 learning as a process departing from content learning. Furthermore, she explains that this philosophy is easier in practice, yet ineffective to warrant later whole language development. Similarly, Squire (1983) and Cavdar and Doe (2012) argue that failing to practice composing and comprehending, which are basically interrelated thinking oriented skills, impedes the efforts not only to teach students to read and write but also how to think.

With all these in mind, in the Ethiopian EFL context, as some past and recent local research findings and my close observation evidenced, the quality of critical thinking has been deteriorating in the EFL classroom, and then university students do not seem to meet the required competence. In his doctoral dissertation, Geremew (2009) reported that students' writing in

different faculties of Addis Ababa University is weak in treating a given topic both in content and form. Furthermore, he noted that most professors require their students to critique books, term/research papers, articles, academic essays, films, and formal reports related to the content of their courses. These activities require students to think critically about how they approach problems, questions, issues, and how they apply affective dispositions such as open-mindedness and diligence in seeking relevant information, being systematic in analyzing information, and inferring that can reasonably be drawn from facts. However, students are not capable of doing these sorts of activities.

In the same way, Dawit (2008) claims that even though the Ministry of Education (1994) asserts that active learning methods or student-centered teaching should govern educational practices in schools and universities, Ethiopian educational system continues to provide students with the traditional model of instruction. Furthermore, he adds that the faulty everyday reasoning and poor argumentation skills used by most students (both orally and in writing) indicate that even a college/university education appears to have a limited effect on graduates' critical thinking abilities, including making reasonable interpretations of texts. Adege (2009) conducted research on "Critical Thinking Pedagogy in EFL classrooms at Jimma University" and came to the conclusion that the majority of EFL instructors (71%) agreed that critical thinking is an important goal of their instructional objectives and/or practices, but only 2% of the total number of university EFL teachers bring explicit modeling of critical thinking in their classroom instruction and 5% of them bring critical thinking assessment into their assignments and examinations. Therefore, the study shows that critical thinking is given less emphasis, or it is a neglected aspect in EFL classrooms in Ethiopia.

Wondifraw (2018) also conducted a study to analyze the effect of infusing intellectual standards of critical thinking on students' critical reading performance. In his study, assessing EFL students' attitudes and practices in reading academic texts critically and exploring students' level of critical reading performance were aimed to supplement the analysis. Finally, it was found that most English language students had inadequate insights on critical reading tasks and inquisitive strategies, and the practices of critical reading activities were not engaging. Likewise, the students' critical reading performance was found to be unsatisfactory. In sum, all the above research results show that there is a gap in instructional practices that need to be addressed particularly in building critical thinking skills.

Concerning reading and writing integrated instruction, there are many recent research works that have been conducted in a foreign context. For instance, Li and Yang (2014) conducted a study on the effects of reading-to-write on critical thinking skills and concluded that practicing reading to write in the

classroom helps students to shift from the passive reception of knowledge to an active seeking for knowledge and also to move from the rote learning of the text to the practical use of the knowledge in solving problems. However, the research conducted in a Chinese educational setting which differs from the Ethiopian context. Similarly, Al-Dosari (2016) carried out a study on the effects of integrated reading and writing on the quality of writing. The results of the study showed a statistically significant improvement in writing, but the researcher did not address critical thinking skills. Likewise, Hailah (2020) conducted a study on the effectiveness of integrating reading and writing pedagogy in the EFL setting and teachers' perception, and he came to the conclusion that integrated reading and writing had a significant impact on students' reading abilities and writing proficiency over a short period. However, he did not address critical thinking skills.

In fact, in the Ethiopian context, there are not many pieces of research that have been conducted on the effects of integrated reading and writing instruction in the EFL classroom. However, Desta (2019) investigated the effects of integrated reading-and-writing practice on EFL learners' performance and self-efficacy of reading comprehension and summary writing with grade eight students. The researcher applied tests and interviews to collect the required data. Finally, he came up with the conclusion that integrated reading-and-writing skill instruction has a positive influence on students' reading comprehension, summary writing, and self-efficacy. However, the researcher did not investigate its effects on students' critical thinking skills achievement.

Likewise, Alemu (2015) conducted a study on the integrated content-based instructions in teaching English reading skills to grade 11 students. He used an interview, classroom observation, and document analysis to collect the data. Finally, he came to the conclusion that teachers had high theoretical orientations about integrated instruction, but because of a number of impediments in the environment related to school, integrated content based instruction could not be implemented in language teaching in reading classrooms. However, the researcher's focus was only on assessing the implementation of integrated content-based instruction on teaching reading skills, not on students' critical thinking.

Therefore, as the literature review and different research evidence imply, in the practice of integrating reading and writing skills, readers must be able to interact with text to construct meaning and in the process of gaining meaning, they should get involved in writing in response to comprehension exercises. Learners should interact with a text to comprehend and compose a text in various forms based on the text they have read. For instance, in the process of reading comprehension instruction, students are supposed to respond to comprehension exercises by writing their predictions, identifying the main idea of paragraphs, paraphrasing, summarizing, interpreting texts, analyzing activities,

explaining, inferring information, writing the gist of the text, and so on. Finally, students should be involved in connecting activities like text-to-self, text-to-text and text-to-world, connecting what students read to what they are familiar with, which is a part of the transaction readers use to comprehend and analyze the text.

Challenging students to think critically about academic subjects and to develop the reasoning abilities they need to deal successfully with real-world reasoning tasks in life is rarely practiced in EFL context in Ethiopia. This observation has, therefore, further motivated this researcher to empirically examine if the university undergraduate students in Ethiopia would be better able to demonstrate critical thinking abilities, and use these same abilities to do better academically in their learning, and in everyday reasoning tasks after having received a semester-long specific integrated reading and writing skills instruction.

Hence, the major concern of this study is to investigate the effectiveness of an integrated reading and writing instructional approach on undergraduate students' critical thinking skills development in the EFL classroom. As far as the researcher's knowledge is concerned, no research of this sort has been conducted in Ethiopia so far. Therefore, this research would be the first and new in its kind, and it would attempt to add to the knowledge in the field and try to fill the gap in this respect.

Objectives of the Study

Teaching and learning reading and writing skills in an integrated way is basic in contemporary language theory and pedagogy. Furthermore, the theory claims that readers must be able to transact with the text to construct the meaning and, in the process of gaining meaning, the students would get involved in writing in response to open comprehension exercises. Therefore, the main objective of this study is checking empirically whether the integrated reading and writing skills instructional approach has an effect on students' critical thinking development in the Ethiopian EFL context. Thus, in this study, the following research questions were formulated:

1. Is there a statistically significant difference between the achievement of the experimental group and that of the control group in their critical thinking skills before the treatment?
2. Is there a statistically significant difference between the achievement of the experimental group and that of the control group in their critical thinking skills after the treatment?

The Null Hypotheses

1. There is no statistical difference between the achievement of the experimental and the control group in their critical thinking skills achievement before the treatment.
2. There is no statistical difference between the achievement of the experimental and the control group in their critical thinking skills achievement after the treatment.

Research Design

In order to address the objectives of this research, experimental research design was employed, and the kind of experiment was a quasi-experiment. This type of experimentation is helpful to examine the effects of an independent variable on dependent variables. Moreover, the quasi-experiment plays a crucial role in the environment where it is not possible to control all other factors that might affect the results. In this experimentation, two intact groups of freshmen who take Communicative English skills courses I participated. They were assigned randomly to control experimental groups in order to avoid any experimental bias.

Scope of the Study

This study was delimited to first-year undergraduate students of Bahir Dar University. The reason for delimiting this study to the first-year undergraduate students was mainly that these students were supposed to study the course Communicative language skills. The assumptions of the course are that it would help them to maximize their communicative and reasoning competence in their oral and written communication. Furthermore, in this stage, students are expected to work at a higher level (i.e., to read more critically, to write more analytically, to think more conceptually with higher-order thinking skills than ever before).

On the other hand, the researcher has chosen these groups of students for some other basic reasons, too. The first one is that the researcher had tangible corpus evidence at hand on the existence of the problems in comprehending texts, composing and critical thinking performance that had been collected through assignments and examinations at various times from the same level of

students in the first semester. The corpus documents that the researcher collected from the groups of students were the main trigger that instigated him to conduct this experimental research with these groups of students.

Data Collection Instruments

In this study, the test was employed as the main data collection instrument. Pre- and post-tests were employed to collect the required data. The tests were adapted from the TOEFL tests by the researcher in collaboration with curriculum experts at the university. The reasons why the researcher adapted the reading passage from TOEFL with some modifications were that the tests are standardized. Pierce (1994) notes that TOEFL tests are standardized, highly secure and internationally administered for assessing the language proficiency level of foreign language speakers. Furthermore, the same author points out that TOEFL tests have high reliability and validity. They have substantial reliability estimates between 0.87 and 0.90 (Pierce, 1994).

The test also included questions that measured students' critical thinking abilities on the basis of the texts. At this point, the critical thinking questions were prepared in an open-ended question format. In sum, 25 critical thinking questions were prepared for the test. The tests were prepared and assessed based on Facione's (1995) critical thinking model. Facione (1995) states that a person engaged in critical thinking uses a core of self-regulated cognitive skills like interpretation, analysis, inference, evaluation, and explanation to form logical reasoning and judgments (p. 3). Herein, the first test was a pre-test that was given to both control and experimental groups to make sure that the two groups had equivalent capacities before the treatment. The second test was a post-test that had been given to both groups at the end of the intervention to check whether the intervention had an effect on the experimental group. All the above-listed assessment models were employed to assess students' critical thinking skills except 'evaluation'. Assessing students' 'evaluation' skill was not used in this specific study since it is highly exposed to subjectivity, as proved in the pilot study. Hence, it was not found convenient in this particular study. Finally, the results of the pre-test and post-test scores were analyzed using an independent T-test.

The Reliability of the Tests

In order to ensure the reliability of the pre-test of the critical thinking questions, two university TEFL professors rated each item to check whether the tests are relevant, appropriate, and consisting of adequate items for the intended purpose and then Cohen's kappa statistical measurement was employed to compute the results of the raters' agreement. So, kappa coefficient measure of agreement between the two TEFL professional raters regarding the pre-test of critical thinking lies in the category of 'very good' was 88%, and the value of the kappa measure of agreement is 0.638. Therefore, according to Koch's benchmark (1977) percent agreement rate classification, the result obtained by the kappa coefficient measure of agreement between the two raters can be considered as perfect agreement, and then the test was used for the intended purpose. Therefore, the pre-test questions which are designed to measure students' critical thinking were reliable.

Similarly, 25 critical thinking questions were rated by two TEFL professionals: the Cohen's kappa coefficient measure of agreement between the two TEFL professional raters regarding the post-test critical thinking questions which lie in the category of 'very good' was 92%, and the value of the kappa measure of agreement was −0.569. Therefore, according to Koch's benchmark (1977) the percent agreement rate classification, the result obtained from the kappa coefficient measure of agreement between the two raters can be considered as a perfect agreement. As a result, the questions were found reliable for the intended purpose.

Internal Consistency of the Tests

In the same way, the internal consistency of both the pre- and post-tests of critical thinking skills' questions were computed using split-half reliability tests. Based on this, the value of the half-split reliability of the pre-tests of critical thinking questions' correlation between forms is 0.309, and the value of the Spearman-Brown coefficient is 0.871. Similarly, the value of the post-test critical thinking correlation between forms is 0.39, and the Spearman-Brown coefficient is 0.85. This indicates that the Spearman-Brown coefficient or the aggregate result exceeds the value of the correlation between forms in both tests. Therefore, based on the above analysis of reliability interpretation, the value of the pre- and post-tests' results of this data is reliable. In general, the values of all types of tests' half-test reliability results revealed that the tests' internal consistency was reliable for the intended purpose.

The Validity of the Tests

As stated above, the tests were adapted from the TOEFL exams. In order to determine whether the reading passages matched the students' level, Gunning Fog Index Formula was used to measure the difficulty level of reading texts. The difficulty level of the pre- and post-tests were 11 and 12 respectively. Gunning (1968) recommends that difficulty levels of reading passages between 8 and 13 are considered as appropriate for undergraduate students. Furthermore, to assess the validity of the tests, the designed test was given to two TEFL university professors to check whether the test is relevant, appropriate, and consisting of adequate items for the intended purpose. Moreover, the tests have been assessed whether they have a mix of easier and harder questions. In all the above criteria, the tests were rated by two professionals in the categories of very good, good, or poor. Finally, the rating results were computed using the kappa measure of agreement. Apart from the two professors, the test was also given to two of research advisors to be commented on and checked whether the tests consist of the required face and content validity for the intended purpose. Eventually, the advisors' and two of the professors' comments were followed in the final preparation of the tests. Lastly, with all amendments, the tests were administered to collect the required data.

Procedures of the Study

As discussed above in the research design, students from two EFL groups were assigned randomly as control and experimental groups. The trained instructor was assigned to teach Communicative language skills I in these groups. In this study, the main focus area is reading and writing. Other language learning activities were shared by both groups. Therefore, in the conventional approach, students often read the text and do a set of activities focusing on comprehension, while the instructor's feedback would focus on correcting students' errors to the given answers. Similarly, writing would be done or practiced on prescribed topics focusing on end products of the hand-written text, and then the instructor's feedback would concern the grammatical accuracy.

However, the newly designed integrated reading and writing instructional procedure was implemented for the experimental group with the following three steps. Firstly, apart from brainstorming discussions, there were writing before reading activities like predicting contents in the form of paragraphs

or writing a short paragraph on different sayings related to the texts before going through the texts (Write-before-you-read activities). Secondly, the students were supposed to write their response to the literal and open-ended questions like comparing their prediction, identifying the main idea of paragraphs, paraphrasing, summarizing, interpreting texts, analyzing, explaining, inferring information, writing the gist of the text (Read and respond through writing activities). Finally, the students were involved in connecting activities like text-to-self, text-to-text and text-to-world activities (Write-text connecting activities). Indeed, when the students failed to implement the above cognitive strategies by themselves, the instructor was supposed to guide and teach explicitly each selected strategy until the students mastered it and worked independently, and the instructor's feedback was given through reflective observation.

The implementation of the instructional strategy for both the control and experimental group lasted for 12 weeks with three sessions per week and was held from October 1, 2019, to December 30, 2019. Finally, a critical thinking test was given to the two groups to see if there is a change in students' critical thinking results. Eventually, the tests were corrected by two different TEFL instructors to avoid unintended biases in correcting the subjective items, and the results the students obtained in the 25 critical thinking questions were used for analysis. Finally, the two instructors' average score was taken for analysis, and then it was analyzed and computed through an independent t-test to see if a significant difference existed between the groups in terms of critical thinking skills development.

Data Analysis, Findings of the Study, and Discussions

In this part, data analysis, findings, and their discussion of the study are presented in accordance with the students' critical thinking skills' test results.

RQ; Is there any significant difference between the two randomly selected intact groups of students in terms of critical thinking skills before the treatment?

Table 1

Descriptive and inferential statistics of the experimental and control groups—pre-test results of students' critical thinking skills

Group		N	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Cohen's d effect size
Inference	control group	48	11.770	2.746	–2.063	94	.072	0.421
	experimental group	48	12.791	2.052				
Interpretation	control group	48	12.250	2.693	–.671	94	.504	0.136
	experimental group	48	12.583	2.142				
Explanation	control group	48	12.583	3.167	–.868	94	.388	0.177
	experimental group	48	13.125	2.943				
Analysis	control group	48	14.791	3.395	–.530	94	.598	0.108
	experimental group	48	15.125	2.733				

As shown in Table 1, there were 48 participants both in the control and experimental groups. The mean scores of the control and experimental groups in terms of each critical sub-skill is presented sequentially. Among the critical thinking elements, inference is one of the sub-skills. The mean score of the experimental (group 2) (M = 12.791; SD = 2.052) was slightly higher than in group 1 (M = 11.771; SD = 2.746). Here, the two groups' mean difference is –1.0208 which is an insignificant means difference. Thus, the results show that students do not differ significantly in inferring before the treatment. Likewise, an independent-samples t-test revealed no statistically significant difference between group 1 (the control group) and group 2 (the experimental) mean scores since t (94) = –2.063; the Sig. (2-tailed) value = .072, P > 0.05. In the same way, the effect size is Cohen's d = 0.421 signifying a medium effect size. Therefore, the results show that there is no statistically significant difference between the two groups in their inference skills before the treatment.

Correspondingly, in terms of students' interpretation, both groups *had* almost equal mean scores on the pre-test. The mean difference between the groups is –.333. Therefore, between the two groups, a non-significant mean difference was obtained. This indicated that the two groups were equal before the treatment. Moreover, the result of an independent samples t-test signified that there is no significant difference between the experimental and control groups in their interpretation. The t (94) = –.671; the Sig. (2-tailed) value = .388, P > 0.05. In the same way, the result of the effect size is Cohen's d = 0.136, specifying a small effect size. Therefore, there was no significant difference

between the two groups of students in their interpretation achievement before the treatment.

Similarly, students' explanation achievement mean score of group 2 (M = 13.125; SD = 2.943) was slightly higher than that of group 1 (M = 12.58; SD = 3.167). The difference between the two groups in their mean scores is –.541. So, the difference margin is very narrow, and this portrays that the two groups did not differ in their explaining achievement before the treatment. The independent samples t-test result shows that there is no significant difference between the experimental and the control group in their explanation achievement since t (94) = –.868; the Sig. (2-tailed) value = .388, P > 0.05. Likewise, the result of the effect size is Cohen's d = 0.177 signifying a small effect size. With this in mind, the statistical evidence showed the absence of statistically significant difference between the two groups in terms of their explanation achievement.

Likewise, students' achievement in analyzing texts was also compared. Based on this, the mean scores of group 2 was (M = 15.125; SD = 2.733) and group 1 (M =l4.791, SD = 3.395) respectively; their mean difference is 1.333 which is insignificant. The results of the independent samples t-test result showed the absence of a significant difference between the control and experimental groups scores in terms of their analysis achievement since t (94) = –.530; the Sig. (2-tailed) value = –.333, P > 0.05. In the same way, the result of the effect size is Cohen's d = 0.108, showing a small effect size. Therefore, from these results, it is possible to conclude that the two groups of students did not differ in their analyzing achievement before the treatment.

The independent sample test results also confirm that the groups' mean differences are not statistically significant in all the above listed critical skill components in the pre-test results or before the treatment. Therefore, there is no significant variance between the experimental and control groups in terms of inferences, interpretations, explanations, and analyses skills' achievement before the treatment. As a result, the null hypothesis is accepted.

Critical Thinking Post-test Result Comparison between the Control and Experimental Groups

The following research question and corresponding null hypothesis were addressed in this section:

RQ: Is there a statistically significant difference between the achievement of the control and experimental groups of students in their critical thinking skills after the treatment?

Null hypothesis (H01): There is no statistical difference between the achievement of the control and experimental groups in their critical thinking skills achievement after the treatment.

Table 2

Descriptive and inferential statistics of the experimental and control groups—post-test result of students' critical thinking skills

Group Statistics								
Group		N	Mean	Std. Deviation	t	df	Sig. (2-tailed)	Cohen's d effect size
Inference	Control group	48	12.25	3.131	–3.634	94	.000	0.74
	Experimental group	48	14.5	2.931				
Interpretation	Control group	48	12.08	4.556	–5.187	94	.000	1.05
	Experimental group	48	16.33	3.385				
Explanation	Control group	48	12.25	4.397	–5.762	94	.000	1.05
	Experimental group	48	17.	3.643				
Analysis	Control group	48	14.25	3.491	–5.184	94	.000	1.176
	Experimental group	48	18.33	4.193				

As it is clearly displayed in Table 2, the number of students in group 1 (control group) and group 2 (experimental group) is equal to 48 participants. The table shows the results of students' critical thinking sub-skills like inference, interpretation, explanation, and analysis after the treatment. Inference is one of the sub-skills. As displayed in the table, there is a significant difference between the experimental (M = 14.5; SD = 2.931) and the control group (M = 12.25; SD = 3.131) in inferring information from a given reading text. As a result, the mean difference between the two groups is –2.25. The independent sample t-test result depicts that there is a significant difference between the experimental and control groups of students in text inference since the t (94) = –3.634; Sig (2-tailed) value = 000, $p < 0.05$. Furthermore, the result of the effect size shows that Cohen's d = 0.74, specifying a strong effect. Therefore, the result of the independent samples t-test indicates that there is a statistically significant difference between the experimental and the control groups of students on the inference post-test. In other words, the experimental group of students scored higher than the control group in terms of inferring information after the intervention. In general, all the above descriptive and inferential statistical data confirm that the intervention that had been implemented in the

experimental group in the EFL classroom had a positive effect on students' inference skills.

By the same token, as the data demonstrated in Table 2 shows, there is a significant mean difference between the experimental (M = 16.331; SD = 3.385) and the control group (M = 12.083; SD = 4.556) in text interpretation in the post-test results. Thus, the mean difference between the two groups is –4.25. So, this descriptive statistics result justifies the claim that the experimental group surpassed the control group in interpreting on the critical thinking skills post-test due to the intervention. This is supported by the independent sample t-test result which shows that there is a significant difference between the experimental and control group in text interpretation; the Sig. (2-tailed) value = 000, $p < 0.05$. Moreover, the result of the effect size shows that Cohen's d = 1.05, specifying a strong effect. Thus, the results of the independent samples t-test verify that there is a statistically significant difference between the experimental and control groups in interpretation due to the intervention. In brief, the experimental group scored higher than the control group on critical thinking post-test. This clearly demonstrates that integrating reading and writing intervention that had been practiced by the experimental group in the EFL classroom had a positive effect on their interpretation skills.

The data exhibited in Table 2 shows that there is a significant mean difference between the experimental (M = 17; SD = 3.643) and the control group (M = 12.25; SD = 4.397) with respect to text explanation. Hence, the mean difference between the two groups is –4.75. So, this result attests that the experimental group exceeded the control group on the text explaining achievement on the post-test after the intervention. This result is verified by the independent sample t-test result that shows a significant difference between the experimental and control groups on text explanation since the t (94) = –5.762; the Sig. (2-tailed) value = 000, $p < 0.05$. Furthermore, the effect size shows that Cohen's d = 1.176, showing a strong effect. Therefore, the results of the independent sample t-test reveal that there is a significant difference between the experimental and the control group in the explanation achievement in the critical thinking skills post-test. Generally, the evidence obtained from these statistical data shows that there is statistically significant difference between the two groups of students on explaining texts in the critical thinking skills post-test result. In a nutshell, it is possible to conclude that the experimental group's test result surpassed the control group text explanation achievement as a result of reading and writing integrated instructional intervention.

Table 2 also shows that there is a mean difference between the experimental (M = 18.333; SD = 4.193) and the control group (M = 14.250; SD = 3.491) in text analyses in the post-test results. Consequently, the mean difference between the two groups is –4.083. Hence, these results assert that the experimental group outscored the control group on text analyzing achievement after the intervention.

Also, the independent sample t-test analysis shows that there is a significant difference between the experimental and control group on text analysis for the t (94) = –5.184; the Sig. (2-tailed) p value = 000. Moreover, the effect size shows that Cohen's d = 1.058, indicating a strong effect. Thus, the statistical evidence obtained from the independent sample t-test pointed towards the idea that there is a statistically significant difference between the experimental and the control group in the text analysis achievement on the critical thinking skills post-test. Altogether, the descriptive and inferential statistical data endorse that there is a statistically significant difference between the experimental and control group in text analysis skills. In other words, the experimental group scored higher than the control group in terms of analyzing texts on critical thinking skills post-test as a result of the intervention. This clearly indicates that the treatment that had been implemented in the experimental group in the EFL classroom had brought a positive effect on students' analytical skills. To sum up, all these data would seem reasonable to point to the likelihood that the null hypothesis is rejected.

Discussion

In this part, to investigate the existence of the difference between the two groups, a 12-week integrated reading and writing instructional intervention was implemented on the experimental group and the traditional reading and writing segregated approach on the control group. Finally, the two groups' critical thinking pre-test and post-test results were compared. Based on this, the students' pre-test critical thinking result was homogeneous before the treatment. However, the post-test results demonstrate that the students in experimental group performed well ($p < .05$) in terms of various aspects of critical thinking like inference, interpretation, explanation, and analysis skills compared to the control group whose scores on each of these scales showed a lower result.

Furthermore, the experimental groups of students' post-test question items were deeply analyzed to see which aspects of the critical thinking skills the experimental group did better than the control group. The analysis of the experimental group students' results revealed that they manifested better results in inferring information like constructing reasonable meaning from a text, identifying the author's purpose or the intended message he/she wants to convey, identifying information that was not clearly stated in the text, and framing syntheses of related ideas into a coherent view and so on. Similarly, the experimental group was good at interpreting texts, identifying key ideas

from the text, paraphrasing a text, guessing the meanings of unfamiliar terms, identifying the purpose of the text, and so on. Also, the experimental group's results at explaining texts like clarifying, summarizing, describing, justifying were much better than the control groups. Likewise, the experimental group was good at analyzing texts by identifying main ideas from a text, finding a relationship between texts, identifying similarities and differences between claims in the specified text, finding the relationship of sentences or paragraphs, and so on. Therefore, the results imply that the treatment that had been implemented in the experimental group in the EFL classroom had brought a positive effect on students' critical thinking skills.

The results of this experimental study substantiate the results obtained by Li and Yang (2014) who state that reading to write practice helps students to shift from passive reception of knowledge to active seeking for knowledge and to move from rote learning to the practical use of knowledge in solving problems. The results obtained in the research also confirm that students could infer, interpret, explain, and analyze the required information on their own more effectively than the control group. Similarly, the result is also consistent with Hirvela's (2004) proposition that using both writing and reading skills in harmonious integration in EFL instruction enhances students' understanding, composing skills, and the ability to look at things critically. Likewise, the findings also agree with the claims of Graham and Hebert (2010) who argue that understanding a text involves writing about it, and writing about a text should enhance comprehension, for it provides students with a tool for visibly and permanently recording, connecting, analyzing, personalizing and manipulating key ideas in a text. Therefore, the results of the current experimental group's critical thinking improvement concurred with the above findings.

The findings of the current study imply that when reading and writing are integrated into instruction, it helps students to combine input and output together at the same time in the EFL classroom. Reading is input, and writing is output, therefore, when students read a text, there are contents and language elements that they can explore in the text. Thus, when they construct written responses for an open-ended comprehension, they apply both comprehension and composition concurrently. In doing so, they show their understanding of what they have read in the text, for writing is important for thinking. When they write in response to the text, the knowledge they get from the text can be transformed into their writing. The results obtained in this research also confirm this reality. In contrast, when reading and writing are taught separately in classroom instruction, students do not connect input and output together. Rather, their focus will be on discrete skills, and there is no way to use language in a holistic way and, as a result, it hampers students' thinking development. Therefore, integrating reading and writing in instruction has immense benefits for students' critical thinking advancement. Moreover, the results of the cur-

rent study also affirm that integrating reading and writing skills instruction to enhance learners' critical thinking can be also practiced in the Ethiopian EFL setting.

Conclusion

In this study, the data collected quantitatively revealed that learners who received an integrated reading and writing instruction demonstrate higher critical thinking skills than the learners who learned in the conventional way. The intervention boosted students' ways of thinking in various dimensions particularly, on inference, interpretation, explanation, and analysis respects. In a nutshell, these results show that integrating the two skills gives students more opportunities to construct their own learning than the conventional approach does, and it also encourages them to be a self-reliant learner and help them develop their self-confidence to take responsibility for their own learning, especially with respect to foreign language learning. Moreover, the findings show that integrated reading and writing instruction into EFL is a good predictor of critical thinking skills improvement for learners as it transforms them from passive to active readers and writers, and also maximizes students' language use instead of simple memorization of facts and grammar rules. As the results of the students' improvement in the critical thinking post-test indicate, integrating the two skills empowered the students to practice the critical thinking components in meaningful ways. The results are in line with DuBrowa's (2011) findings that the integration of reading and writing enables students to develop both critical thinking and critical literacy that promote students' ability to transform information for their own purposes.

Furthermore, EFL instructors should be aware that implementing reading and writing integrated instruction helps learners to be immersed in reading and writing which demand the use of the two language skills at a time, and then it may make learners be motivated to use authentic language and enables them to interact naturally with the target language and develop implicit knowledge as well. Moreover, it develops students' learning in all disciplines because it requires them to become more actively engaged in what they are learning and with that engagement, greater academic success will come and that in turn increases students' motivation. In the same way, it is more relevant to teach reading and writing in an integrated way while teaching communicative language skills in higher institutions to enhance students' critical thinking in the EFL classroom.

Finally, the researchers of this study recommend other researchers to continue to explore the potential effects of this pedagogical approach at different grade levels, using multiple data collection instruments.

Limitations of the Study

There were some limitations that affected this study. Some were associated with the area of sample and sampling; the sample in this study was handpicked on purpose from one university which may not represent most students from all universities in Ethiopia; however, this is done by considering that the students in this university can possibly come from every corner of the country. Furthermore, though every effort has been made to investigate this theme as thoroughly as possible, it is impossible to claim confidently that the researchers were able to scrutinize every single bit of areas in comprehensively finding all possible researches on this respect.

References

Alemu, A. (2016). The effects of explicit instruction in critical thinking on student achievement in writing. (Unpublished PhD Dissertation). Addis Ababa University.

Al-Dosari, H. (2016). Effects of an integrated reading and writing approaches on improving writing skills of EFL students. *Dirasat Educational Sciences, 43* (1), 761–771. https://doi.org/10.12816/0030028

Alghonaim, A. S. (2018). Explicit ESL/EFL reading-writing connection: An issue to explore in ESL/ EFL Settings. *Theory and Practice in Language Studies, 8*(4), 385–392. https://doi.org/10.17507/tpls.0804.04

Carson, J., & Leki, I. (1993). Reading in the composition classroom: Second language perspective. Heinle & Heinle. https://doi.org/10.1017/S027226310001336X

Cavdar, G., & Doe, S. (2012). Learning through writing: Teaching critical thinking skills in writing assignments. *PS: Political Science and Politics, 45* (2), 298–306. https://www.jstor.org/stable/i40070009

Cobine, Gary R. (1995). Writing as a response to reading. ERIC Digest, ED386734. https://catalogue.nla.gov.au/Record/4117942

DuBrowa, M. (2011). Integrating reading and writing: One professor's story. *Journal of Research and Teaching in Developmental Education, 28*(1), 30–33. https://eric.ed.gov/?id=EJ960411

Elder, L., & Paul, R. (2009). Critical thinking: Strategies for improving student learning. *Journal of Developmental Education, 32*(3), 40. https://eric.ed.gov/?id=EJ868672

Eun-Hee, N., & Myeong-Hee, S. (2020). Teaching activities and students' preferences in integrated English reading and writing classes. *English Teaching,* 75 (2), 69–91. https://doi.org/10.15858/engtea.75.2.202006.69. http://journal.kate.or.kr

Facione, P. A. (1992). The California critical thinking skills test. Academic Press. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.131.8583&rep=rep1&type=pdf

Fisher, A. (2001). *Critical thinking: An introduction*. Cambridge University Press. https://www.worldcat.org/title/critical-thinking-an-introduction/oclc/47983908

Geremew, L. (2009). A study on the academic writing requirements: Four departments in Focus. (Unpublished PhD dissertation). Addis Ababa University.

Getu, A. (2015). Assessment of integration of content-based instruction in teaching English. (Unpublished PhD dissertation). Addis Ababa University,

Graham, S., & Hebert, M. A. (2010). Writing to read: Evidence for how writing can improve reading. *A Carnegie corporation time to act report*. Washington, DC. Alliance for Excellent Education. https://www.carnegie.org/publications/writing-to-read-evidence-for-how-writing-can-improve-reading

Haregewoin A. (2008). The effect of communicative grammar on the accuracy of academic writing. (Unpublished PhD dissertation). Addis Ababa University.

Hao, X., & Sivell, J. (2002). Integrating reading and writing in EFL composition in China. Paper presented at the annual meeting of the Canadian association of applied Linguistics: Humanities and social sciences. Congress, Toronto. https://files.eric.ed.gov/fulltext/ED468599.pdf

Hailah, A. (2020). Evaluating the effectiveness of integrating reading and writing pedagogy in EFL setting and teachers' perceptions. *English Language Teaching, 13*(5), 177–190. https://doi.org/10.5539/elt.v13n5p177

Hirvela, A. (2004). *Connecting reading and writing in second language writing instruction*. The University of Michigan Press.

Kintsch, W., & Vipond, D. (1979). Reading comprehension and readability in educational practice and psychological theory. In L. G. Nilsson (Ed.), *Perspectives on memory research* (pp. 329–365). https://philpapers.org/rec/KINRCA-3

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* (March 1977), 33(1), 159–174. PMID: 843571.

Li, Z., & Yang, C. (2014). Reading-to-write: A practice of critical thinking. *Journal of Arts and Humanities 3*(5), 67–71.

Nigusse, D. (2019). The effects of integrated reading-and-writing practice on EFL learners' performance and self-efficacy of reading comprehension and summary writing on grade eight students. (Unpublished PhD dissertation). Bahir Dar University.

Tajzad, M., & Namaghi, S. (2014). Exploring EFL learners' perceptions of integrated skills approach: A grounded theory. *English Language Teaching,* 7 (11), 92–98. https://doi.org/10.5539/elt.v7n11p92

Teshome, T., & Prasana, S. (2016). Students' writing self-efficacy and writing apprehension relating to their writing performance: Reflection on Ethiopian first year university students. *International Journal of Humanities, Language and Literature*, 3(1), 22–26. https://1library.net/document/z12m6dey-students-efficacy-apprehension-relating-performance-reflection-ethiopian-university.html

Vanniarajan, Swathi. (1990). *Language learning strategy: What every teacher should know* by Rebeca L. Oxford. *Issues in Applied Linguistics,* 1(1), 115–120. https://doi.org/10.5070/L411004984

Spack, R. (1985). Literature, reading, writing, and ESL: Bridging the gaps. *TESOL Quarterly,* 19(4), 703–725. https://doi.org/10.2307/3586672

Richards, J. C., & Rodgers, T. S. (1986). *Approaches and methods in language teaching.* Cambridge University Press. https://www.novaconcursos.com.br/blog/pdf/richards-jack-c.-&-rodgers.pdf

Rosenblatt, L. M. (2004). The transactional theory of reading and writing: Theoretical models and processes of reading. *International Reading Association, 48*(1), 1363–1398 https://northtrap.wordpress.com/2012/07/15/the-transactional-theory-of-reading-and-writing-louise-m-rosenblatt

Tsai, J. (2006). Connecting reading and writing in college EFL courses. *The Internet TESL Journal, 12* (12), 465–485. http://iteslj.org/Articles/Tsai-ReadingWritingConnection.html

Wondifraw, D. (2018). Infusing intellectual standards of critical thinking on students' critical reading performance. (Unpublished PhD dissertation). Addis Ababa University.

Ayalew Tilahun, Berhanu Simegn, Mulugeta Teka

Die Untersuchung der Praxis vom integrierten Lese- und Schreibunterricht zur Förderung des kritischen Denkens bei EFL-Studenten

Zusammenfassung

Das grundlegende Ziel der vorliegenden Arbeit war es, die Praxis des integrierten Lese- und Schreibunterrichts im Hinblick auf die Entwicklung des kritischen Denkens bei den EFL-Studenten der Universität Bahir Dar zu untersuchen. Dabei wurde ein quasi-experimentelles Vortest-Nachtest-Design eingesetzt. An der Untersuchung nahmen 96 Studenten im Hauptfach Englisch teil, die nach dem Zufallsprinzip einer Versuchs- (n=48) bzw. einer Kontrollgruppe (n=48) zugeordnet wurden. Die Kontrollgruppe wurde mittels einer separaten Lese- und Schreibmethode unterrichtet, während die Versuchsgruppe 12 Wochen lang drei Stunden pro Woche mittels einer neu entwickelten, integrierten Lese- und Schreibmethode lernte. Anschließend wurden 25 Vor- und Nachtests mit Fragen zum kritischen Denken bearbeitet, um die Entwicklung des kritischen Denkens bei den Studierenden zu bewerten. Zur Überprüfung der Zuverlässigkeit und der internen Konsistenz von den beiden Tests wurden Kappa-Statistiken (Interrater- und Split-Half-Reliabilitätstest) verwendet. Schließlich bediente man sich bei der Datenauswertung eines unabhängigen t-Tests, dessen Ergebnisse zeigten, dass sowohl die Kontroll- als auch die Versuchsgruppe in Bezug auf deren Niveau in den Vortests zum kritischen Denken gleich abschnitten. Allerdings stellte sich am Ende der Untersuchung heraus, dass der integrierte Lese- und Schreibunterricht gegenüber den herkömmlichen Methoden zur Förderung des kritischen Denkens bei Studierenden überlegen ist.

Schlüsselwörter: integrierter Lese- und Schreibunterricht, Sprachpädagogik, Lesekompetenz, Schreibkompetenz, kritisches Denken

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 129–155
https://doi.org/10.31261/TAPSLA.10353

Seray Tanyer
Anadolu University, Turkey
https://orcid.org/0000-0001-6190-8651

Samet Deniz
Anadolu University, Turkey
https://orcid.org/0000-0002-4493-816X

Turkish EFL Learners' Acquisition of Psych Verbs and Unaccusative Verbs: A Replication Study on Underpassivization and Overpassivization

Abstract

The processability account anticipates that learners will make more underpassivization errors than overpassivization errors since passivization entails more processing. Although one study on psych verbs and a few on unaccusatives examined Turkish L2 learners' acquisition, no research compared a single set of learners' acquisitions of these verbs together from a processing point of view. In this regard, the current study aims to investigate whether the processing complexity of passivization influences acquisition of psych and unaccusative verbs. It also questions whether general accuracy levels in Grammaticality Judgement Task (GJT) and degree of familiarity with target verbs are related to their level of accuracy with individual psych and unaccusative verbs. 33 undergraduate-level university students performed on the GJT and a Word Familiarity Rating Task (WFRT). The GJT included 38 items with 12 sentences for psych verbs, 12 sentences for unaccusative verbs, 12 sentences for distracters and two sentences for examples. The WFRT was a survey questioning familiarity with six psych and six unaccusative verbs. To analyse the data, a set of nonparametric tests and descriptive statistics were used. The results revealed that learners performed more accurately on unaccusatives than on psych verbs. They did more underpassivization errors by accepting ungrammatical active constructions of psych verbs. Their performances on psych and unaccusative verbs went parallel with their general accuracy levels in GJT while their degree of familiarity with and accuracy level for two verbs do not correlate with each other. The results suggest that such factors as processability and L1 transfer seem to impact the acquisition.

Keywords: second language acquisition, psych verbs, unaccusative verbs, underpassivization, overpassivization

Research on L1 acquisition indicates that acquisition of passive voice by native English children is postponed to the later stages during their language development (de Villiers & de Villiers, 1973). This later development of child language in terms of passivization has been attributed to the complexity of syntactic operations and problems in semantic-syntactic mapping (Pienemann, 2005). In a scenario where processability plays a role in the acquisition of English as an L2, passive constructions requiring more processing will be more difficult to acquire as compared to active constructions. As a result, L2 learners of English can potentially make more errors with passive constructions, favor active constructions over passive constructions, and even wrongly accept or use active constructions in the context where passives are necessary.

The underuse and overuse of passive constructions with two particular verb classes (i.e., psych verbs and unaccusative verbs) have already been observed in the studies conducted with Turkish learners of L2 English (i.e., Kurtoğlu, 2006; Montrual, 2001). What is deficient is a comparison between the acquisition of psych and unaccusative verbs from a processing point of view. As this view assumes that linguistic components necessitating less processing can be acquired earlier than those necessitating more processing, L2 learners of English can be expected to make more underpassivization errors with psych verbs than overpassivization errors with unaccusative verbs. Hence, the current research mainly intends to investigate Turkish EFL learners' acquisition of psych verbs and unaccusative verbs with a specific focus on underpassivization and overpassivization.

Literature Review

The literature suggests different hierarchies for thematic roles (e.g., Bresnan, 2001; Foley & Van Valin, 1984; Givon, 1984; Jackendof, 1972) and grammatical functions (Keenan & Comrie, 1977) as seen in (1) and (2) below.

(1) Thematic Hierarchy (Bresnan, 2001, p. 307)
Agent > Beneficiary > Experiencer/Goal > Instrument > Patient/Theme > Locative

(2) Relational Hierarchy (Keenan & Comrie, 1977)
(core) (non-core)
SUBJ > OBJ > OBJq > OBJq > COMPL > ADJUNCT

Depending on the thematic and relational hierarchies, Pieneman et al. (2005) suggested the Lexical Mapping Hypothesis which has recently been resolved as follows:

> In second language acquisition learners will initially map the highest available role in the semantic hierarchy onto a minimally specified SUB/TOP. We call this: canonical mapping. Then they learn to map further arguments onto grammatical functions (GF) other than SUBJ or OBJ. Finally, they learn to attribute prominence to a particular thematic role lower in the semantic hierarchy by promoting it to SUBJ. At this stage they also learn to defocus the highest role by suppressing it or mapping it onto a GF other than SUBJ. We call these non-canonical mapping. (Kawaguchi, 2013, p. 93)

As detailed above, canonical mapping in English necessitates at least two arguments in its transitive actualization. *Agent* as the more prominent role is placed on *Subject* while *Patient* as the less prominent role is placed on *Object* position. There also exist some intransitive verbs (i.e., unergative verbs, e.g., cry) which shows a canonical mapping operation. The only argument of these verbs would be an *Agent* or *Experiencer* which is a high role in the thematic hierarchy and is mapped on *Subject* position.

(3) Canonical – transitive The dog ate the meat.
Agent *Patient* (thematic roles)
Subject *Object* (grammatical functions)

(4) Canonical – intransitive The child cried.
Agent (thematic role)
Subject (grammatical function)

Apart from canonical mapping, two types of non-canonical mapping can be observed in English that are either *structurally* or *lexically* created. The non-canonical mapping created structurally "is usually called 'structural' because the alternative lexical entry creates a structural frame which is regular and predictable" (Kawaguchi, 2013, p. 96). This kind of non-canonical mapping can be exemplified with passive (e.g., active eat versus passive be eaten, (5)) and causative constructions (make X wash Y).

(5) Passive mapping The meat was eaten by the dog.
Patient *Agent*
Object *Subject*

The non-canonical mapping produced lexically, on the other hand, "are intrinsically required by the lexical verb" (Kawaguchi, 2013, p. 96), which makes them irregular and unpredictable as compared to the first one. So, the L2 learners need to discover and learn them as how the occasion requires. Two different verb classes which build non-canonical mappings in English and are within the scope of the current research are unaccusative verbs and psych verbs.

Unaccusative verbs are a sub-category of intransitive verbs which take Themes as their Subjects. This category of verbs includes verbs of occurrence (e.g., happen, occur), and dis/appearance (e.g., die, disappear; Levin, 1993,

pp. 30, 258) which are defined as non-alternating unaccusatives since they do not have any transitive or passive alternatives (6).

(6) a. Unaccusative — An accident happened.
Theme

b. Transitive — *The woman happened an accident.
Subject *Theme*

c. Passive — *The accident was happened.
Theme

The other verb type forming non-canonical mappings in English are psych verbs which are a subcategory of transitive verbs (e.g., frighten, surprise, disappoint; although intransitive use of them is possible for some verbs and forms, it is not a concern of the current paper), and they basically "describe the bringing about of a change in psychological or emotional state" (Levin, 1993, p. 189). In English, these verbs take at least two arguments in their transitive and passive use that are *Theme* and *Experiencer*. However, locating *Experiencer* in *Subject* position by using some psych verbs intransitively results in an ungrammatical use in English (see 7c).

(7) a. Transitive — The news surprised the people.
Theme *Experiencer*

b. Passive — The people were surprised at the news.
Experiencer *Theme*
The people were surprised when they heard the news.
Experiencer

c. Intransitive — *The people surprised when they heard the news.
Experiencer

Hahn (2011) argues that if processability plays a role in the acquisition of English as an L2, passive constructions requiring more processing will be more difficult to acquire as compared to active constructions. Therefore, L2 learners of English can potentially make more errors with passive constructions, favor active constructions over passive constructions, and even wrongly accept or use active constructions in the context where passives are necessary. Since psych verbs map the *Experiencer* in the subject position (i.e., The people were surprised…), they can be regarded as a subcategory of passive constructions. This implies that the processing load of passivization is also valid for psych verb constructions, too. For this reason, L2 learners of English are supposed to deal with processing complexity of passivization in acquiring psych verb constructions.

The incorrect underuse of passives with psych verbs was documented in a study of Montrul (2001) conducted with Turkish learners of L2 English. Although intransitive psych verbs are accepted ungrammatical in English (i.e., *The hunter frightened), Turkish L2 learners were found to label this

use of psych verbs as grammatical. Overpassivization can, on the other hand, be a probable area of errors for L2 learners of English. Turkish learners, for example, have been observed while overpassivizing in the contexts where they are supposed to prefer active construction with unaccusative verbs (Kurtoğlu, 2006). If Turkish learners are observed while making such an underpassivization error in (7c) more than the overpassivization error in (6c), the reason of learners' underpassivization in English psych verb constructions can stem from the processing complexity of passivization. This claim has also been supported by Kawaguchi (2013, p. 99) which indicated that "within lexically non-canonical mapping, … unaccusatives are acquired before psych verbs." Therefore, the findings of the translation task conducted with Japanese L2 learners of English revealed the following acquisition order: *canonical* > *unaccusative verb* > *psych verb.*

The above few studies have indicated that Turkish learners of L2 English can underpassivize psych verbs and overpassivize unaccusative verbs. If processability account has a role in constraining L2 acquisition of English, linguistic dimensions necessitating less processing can be hypothesized to be acquired earlier than those necessitating more processing. Therefore, Turkish learners can commit more underpassivization errors with psych verbs as compared to overpassivization errors with unaccusative verbs. As no particular research has been conducted related to this issue with a specific reference to Turkish learners of L2 English, the current study has attempted to shed a humble light on it by investigating the acquisition of unaccusative and psych verbs simultaneously and comparatively. While quite a few studies (i.e., Kurtoğlu, 2006; Montrul, 2001) have investigated the nature of Turkish learners' acquisition of psych verbs and unaccusative verbs, no research have been conducted to examine these verbs together from a processing account of language acquisition and for a single set of subjects. The current study, therefore, aims to investigate Turkish L2 learners' acquisition of psych verbs and unaccusative verbs, and find out whether the development of these learners' interlanguage is constrained by the processability account of language acquisition or not. It also aims to investigate whether Turkish L2 learners' perceived word familiarity is related to their acquisition of psych verb and unaccusative verb constructions.

Research Questions

RQ1. Do ELT students perform more accurately on any verb type: psych verbs vs. unaccusative verbs?

RQ2. Do ELT students' performances on psych verbs and unaccusative verbs differ across their accuracy levels in GJT (i.e., low, medium, high)?

RQ3. Do ELT students with low, medium and high accuracy of GJT perform better on any verb type: psych verbs vs. unaccusative verbs?
RQ4. Are ELT students more familiar with any verb type: psych verbs or unaccusative verbs?
RQ5. Is there any relationship between ELT students' degree of familiarity and level of accuracy with psych and unaccusative verbs?
RQ6. Do ELT students perform more accurately on any construction of psych and unaccusative verbs: passive vs. active constructions?

In parallel with these research questions, three main hypotheses developed by Hahn (2011) and reported below have been tested. Among them, RQ1 and RQ6 are relevant to Hypothesis 1, RQ2 and RQ3 are related to Hypothesis 2 while RQ4 and RQ5 are concerned with Hypothesis 3.

Hypothesis 1: Passives charge processing cost, and thus learners will make more errors due to underpassivization than overpassivization. As a result, psych verb constructions, which require passivization, will be more difficult to acquire than unaccusatives, which do not require it.
Hypothesis 2: Once learners reach a stage where passivization does not charge extra-cost, learners' underpassivization errors will reduce, and accordingly, their errors with psych verb constructions will decrease rapidly.
Hypothesis 3: Learners' difficulty with psych verb constructions and unaccusative constructions will not depend on their familiarity with individual verbs. (Hahn, 2011, p. 63)

Method

Participants

A total of 33 undergraduate-level university students enrolling in the program of English Language Teaching (ELT) at a state university of Turkey participated in the study. The participants consisted of first-year (N = 24, 73%), second-year (N = 7, 21%) and fourth-year (N = 2, 6%) students. Of the 33 participants, 18 were females (55%) and 15 were males (45%). This group of students were chosen as the research sample because the current study was designed as a replication of the research conducted by Hahn (2011) whose participants were comprised of the university students studying English as their major and minor. The participants of the current study were members of the course called Critical Reading which is developed for the second half of first-year ELT program during the 2017–2018 academic year. The first-year

program is also full of courses aiming to improve students' L2 (i.e., English) such as Contextual Grammar, Academic Reading, Written Communication, Listening Comprehension, Oral Communication Skills, and Academic Writing and Report Writing.

Instruments

Two different tasks developed by Hahn (2011) were applied in the current study: (1) Grammaticality Judgement Task (GJT) and (2) Word Familiarity Rating Task (WFRT). The first one, GJT, were a total of 38 items with 12 sentences for psych verbs, 12 sentences for unaccusative verbs, 12 sentences for distracters and two sentences for examples (see Appendix A). The GJT included six psych verbs (i.e., *disappoint*, *embarrass*, *frighten*, *frustrate*, *shock*, *surprise*) and six unaccusative verbs (i.e., *arrive*, *die*, *disappear*, *fall, happen*, *occur*) with which six pairs of grammatical and ungrammatical sentences were formed as exemplified in (8) and (9). It also includes 12 distracting sentences theorized to be problematic for L2 learners and serving the purpose of preventing the students from concentrating excessively on psych and unaccusative verb constructions. For more details related to task design process, you may take a look at the "Instrument and Procedure" section of Hahn (2011, pp. 63–66).

(8) Psych-verb constructions
Grammatical: The girl *was disappointed* when she found her grade. (Item 19)
Ungrammatical: The boy *disappointed* because he got a bad grade. (Item 31)

(9) Unaccusative-verb constructions
Grammatical: A letter from her mom *arrived* two days ago. (Item 34)
Ungrammatical: Your package *was arrived* a few hours ago. (Item 20)

The second task applied was the WFRT and it was a survey questioning familiarity with a total of 12 psych and unaccusative verbs (see Appendix B). In this task, the participants were supposed to indicate their degrees of familiarity with individual psych and unaccusative verbs on a 4-point Likert scale ranging from "1 = very unfamiliar" to "4 = very familiar."

Data Collection and Analysis

The data were collected in the second semester of 2017–2018 academic year at a Turkish state university in Turkey. The participants were from the two sections of Critical Reading class and the study was conducted during class time. The two tasks were implemented at the same time with GJT being the

first and WFRT being the second one asked to be filled out. Different from the previous study (Hahn, 2011), an untimed application of the GTJ was preferred because of practical reasons, and additionally, participants were requested not to turn back and revise their responses for the task. "Volunteer participation" was an adopted criterion during the data collection process therefore only the students having this volunteerism performed the two target tasks by signing the certificate of consent (see Appendix A).

As for data analysis, firstly, the individual performances of the participants were calculated. In GJT, each accurate response corresponded to "1" point while each inaccurate response corresponded to "0." The sum of the points participants collected from GJT was equivalent to their accuracy scores. In addition to the total score of GJT, the scores for psych verb and unaccusative verb constructions as well as distracters were also calculated. As for WFRT, degree of familiarity with psych and unaccusative verbs were determined according to the values chosen by the participants on the 4-point Likert scale. Therefore, the familiarity degree of each participant for individual and two groups of verbs ranged across 1, 2, 3, and 4. Based on these scores, descriptive statistics were tabulated and a set of nonparametric tests were conducted with the aim of answering research questions.

In order to answer RQ1 (i.e., Do ELT students perform more accurately on psych verbs or unaccusative verbs?), a Wilcoxon Signed-Ranks Test was conducted to compare the participants' performances on psych and unaccusative verbs. To investigate (RQ2) whether ELT students' performances on psych verbs and unaccusative verbs differ across their accuracy levels in GJT (i.e., low, medium, high), two different Kruskal-Wallis Tests were conducted, which was followed by a sets of Mann-Whitney U tests as follow-ups. For RQ3 (i.e., Do ELT students with low, medium and high accuracy of GJT perform better on any verb type including psych verbs and unaccusative verbs?), three different Wilcoxon Signed-Ranks Tests were performed to compare the performances of low, moderate and high accurate participants on psych and unaccusative verbs. In order to find out whether ELT students are more familiar with psych verbs or unaccusative verbs (RQ4), a Wilcoxon Signed-Ranks Test was conducted to compare the participants' degree of familiarity with psych and unaccusative verbs. The Pearson correlation coefficient was used to investigate the possible relationship between degree of familiarity and level of accuracy with psych and unaccusative verbs (RQ5). Lastly, so as to investigate whether ELT students perform more accurately on any of the two constructions of psych and unaccusative verbs (i.e., passive vs. active constructions), two sets of Wilcoxon Signed-Ranks Tests were conducted (RQ6). In addition to these nonparametric tests, descriptive statistics of the related variables were also reported in the Results section in a detailed manner.

Results

As a first step to data analysis, the normality of the data has been tested. A Shapiro-Wilk's Test (p = .031) and the visuals of histograms, normal Q-Q plots and box plots showed that the data from GJT were not normally distributed, with a skewness of –0.539 (*SE* = 0.409) and a kurtosis of –0.769 (*SE* = 0.798). Also, the data from WFRT were found not to be normally distributed with a skewness of –1.572 (*SE* = 0.409) and a kurtosis of 2.786 (*SE* = 0.798). As the assumption of normality has not been met for this sample, the researchers used nonparametric tests for the analyses.

Table 1.

Descriptive statistics: Psych-v, unaccusative-v, other structures and GJT

	N	Min.	Max.	Mean	Std. Deviation	Std. Error Mean
GJT_Psych-V	33	1	12	9.424	2.598	0.452
GJT_Unaccusative-V	33	3	12	9.787	1.916	0.333
GJT_Other Structures	33	5	11	8.545	1.715	0.298
GJT (Total)	33	18	35	27.757	4.416	0.768

As indicated in the Table 1, the participants of the current study performed more accurately on unaccusative verbs (*M* = 9.787, *SD* = 1.916) than on psych verbs (*M* = 9.424, *SD* = 2.598).

Table 2.

Wilcoxon Signed-Rank Test: Psych-v vs. unaccusative-v in GJT

GJT_Unaccusative-V GJT_Psych-V	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	12	13.75	165	–.583*	.560
Positive Ranks	15	14.20	213		
Equal	6				
TOTAL	33				

*Based on negative ranks.

To answer **RQ1** (i.e., *Do first-year ELT students perform more accurately on any verb type: psych verbs vs. unaccusative verbs?)* the following step was to conduct a Wilcoxon Signed-Ranks Test in order to compare the performances of the participants for two types of constructions: psych verbs (*Mdn* = 10) and unaccusative verbs (*Mdn* = 10). The results (see Table 2) indicated that despite the values found by the descriptive statistics (Table 2),

unaccusative verb ranks were not statistically significantly higher than psych verb ranks ($Z = –0.583$, $p > .05$).

Table 3.

The formula applied to calculate three accuracy levels in GJT

Low	M – SD = Y (Y and above scores)	Y = 23.341 X = 32.173
Medium	M + SD = X and M – SD = Y (The scores between X and Y)	
High	M + SD = X (X and above scores)	

To answer **RQ2** (i.e., *Do first-year ELT students' performances on psych verbs and unaccusative verbs differ across their accuracy levels in GJT (i.e., low, medium, high)?*), firstly, the accuracy profile of the participants for GJT have been decided. For this purpose, the accuracy score of each participant for GJT has been calculated. While scoring the participants' responses for the task, each accurate response corresponded to "1" point. Each inaccurate response, on the other hand, corresponded to "0" point. Therefore, the sum of the points participants collected for 36 items composed the accuracy score for GJT. As stated in Table 3, the mean value and the standard deviation of the mean for GJT were found as follows: $M = 27.757$, $SD = 4.416$. Based on these values, the obtained accuracy scores in GJT were divided into three levels as *Low-Accurate*, *Medium-Accurate* and *High-Accurate* using the formula detailed in Table 3.

Table 4.

The accuracy profiles of the participants for GJT

Level	Accuracy in GJT	Min.	Max.	*N*	[%]
Low-accurate	23 and below	18	23	7	21
Medium-accurate	between 23 and 32	24	31	20	61
High-accurate	32 and above	32	35	6	18
TOTAL				33	100

As a result, the participants with the accuracy scores of "23" and below were determined as *low-accurate* ($N = 7$, 21%) while the ones with the accuracy scores of "32" and above were decided as *high-accurate* ($N = 20$, 61%). The rest of the participants having the accuracy scores between "23" and "32" were, on the other hand, defined as *medium-accurate* ($N = 6$, 18%) as depicted in Table 4.

Table 5.

Descriptive statistics: Accuracy for psych-v and unaccusative-v across 3-levels in GJT

Accuracy Level GJT	GJT_Psych-V			GJT_Unaccusative-V		
	Mean	Std. Deviation	Std. Error Mean	Mean	Std. Deviation	Std. Error Mean
Low (*N* = 7)	6.571	2.820	1.065	8.142	2.609	0.986
Medium (*N* = 20)	9.850	1.926	0.430	9.900	1.447	0.323
High (*N* = 6)	11.833	0.408	0.166	11.333	0.816	0.333

In Table 5, the related values of low-accurate, medium-accurate and high-accurate participants for psych verbs and unaccusative verbs have been tabulated. As seen in the table, the mean scores of low and medium accurate participants for unaccusative verbs (M = 8.14, SD = 2.60 and M = 9.90, SD = 1.44 respectively) are higher than their mean scores for psychverbs (M = 6.57, SD = 2.82 and M = 9.85, SD = 1.92). But, the situation is totally opposite for the high-accurate participants. Their scores for psych-verbs (M = 11.83, SD = 0.40) are higher than their scores for unaccusitive-verbs (M = 11.33, SD = 0.81). In addition to the descriptive statistics, a set of non-parametric tests (i.e., Kruskal-Wallis Test) have also been conducted in order to discover the performances on psych verbs and unaccusative verbs across three GJT accuracy levels.

Table 6.

Kruskal-Wallis Test: Psych-v and unaccusative-v across 3 accuracy levels of GJT

Accuracy Score	*Groups*	*N*	*Mean Rank*	*Chi-square*	*df*	*p*
GJT_Psych-V	Low	7	6.21	15.733	2	0.000
	Intermediate	20	17.78			
	High	6	27.00			
	TOTAL	33				
GJT_Unaccusative-V	Low	7	9.64	10.071	2	0.007
	Intermediate	20	16.78			
	High	6	26.33			
	TOTAL	33				

The results of the Kruskal-Wallis Test (see Table 6) indicated that there is a statistically significant difference among three group of learners' performances on psych verb constructions ($H(2) = 15.733$, $p < .001$), with a mean rank of 6.21 for the *low-accurate* students, 17.78 for the *medium-accurate* students and 27 for the *high-accurate* students. The same statistically significant difference

also exists among learners' performances on unaccusative-verb constructions ($H(2) = 10.071$, $p < .001$), with a mean rank of 9.64 for the *low-accurate* students, 16.78 for the *medium-accurate* students and 26.33 for the *high-accurate* students. These findings show us that as the total level of accuracy in GJT increases, the level of accuracy with psych verbs and unaccusative verbs increase individually, too.

A series of Mann-Whitney U Tests were conducted as follow-ups in order to detect the significant differences in verb accuracy occurring across three different general accuracy levels (i.e., low, intermediate, and high). The results indicated significant differences between the unaccusative verb scores of *low* (*Mdn* = 9) and *high* (*Mdn* = 11.5) accurate ($U = 2.5$, $p = .007$), and *medium* (*Mdn* = 10) and *high* (*Mdn* = 11.5) accurate participants ($U = 22.5$, $p = .019$) while the difference is not significant between the scores of *low* (*Mdn* = 9) and *medium* (*Mdn* = 10) accurate participants ($U = 37$, $p = .062$). They also revealed significant differences between the psych verb scores of *low* (*Mdn* = 7) and *medium* (*Mdn* = 10) accurate ($U = 19$, $p = .004$), *low* (*Mdn* = 7) and *high* (*Mdn* = 12) accurate ($U = 0$, $p = .002$), and *medium* (*Mdn* = 10) and *high* (*Mdn* = 12) accurate participants ($U = 16.5$, $p = .006$).

Table 7.

Wilcoxon Signed-Rank Test: Psych-v vs. unaccusative-v with low-accuracy in GJT

GJT_Unaccusative-V_Low GJT_Psych V_Low	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	2	3.25	6.50	–0.841*	.400
Positive Ranks	4	3.63	14.50		
Equal	1				
TOTAL	7				

* Based on negative ranks.

As indicated in the Table 5, the low-accurate participants performed more accurately on unaccusative verbs ($M = 8.142$, $SD = 2.609$) than on psych verbs ($M = 6.571$, $SD = 2.820$). In order to answer **RQ3** (i.e., *Do first-year ELT students with low, medium and high accuracy of GJT perform better on any verb type including psych verbs and unaccusative verbs?)*, the following step was to conduct a set of Wilcoxon Signed-Ranks Test in order to compare the performances of the low-accurate participants for two types of constructions: psych verbs (*Mdn* = 7) and unaccusative verbs (*Mdn* = 8). The results (see Table 7) indicated that unaccusative verb ranks were not statistically significantly higher than psych verb ranks ($Z = –0.841$, $p > .05$).

Table 8.

Wilcoxon Signed-Rank Test: Psych-v vs. unaccusative-v with medium-accuracy in GJT

GJT_Unaccusative-V_Medium GJT_Psych-V_Medium	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	9	10	90	–0.204*	.839
Positive Ranks	10	10	00		
Equal	1				
TOTAL	20				

* Based on negative ranks.

The same result is also valid for the medium-accurate and high-accurate participants. The medium-accurate participants performed slightly more accurately on unaccusative verbs (M = 9.90, SD = 1.447) than on psych verbs (M = 9.85, SD = 1.926). In order to compare the performances of the medium-accurate participants for psych verb (Mdn = 10) and unaccusative verb (Mdn = 10) constructions, the following step was to conduct a Wilcoxon Signed-Ranks Test. The results (Table 8) indicated that unaccusative verb ranks were not statistically significantly higher than psych verb ranks (Z = –0.203, p > .05).

Table 9.

Wilcoxon Signed-Rank Test: Psych-v vs. unaccusative-v with high-accuracy in GJT

GJT_Unaccusative-V_High GJT_Psych-V_High	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	2	1.50	3.00	–1.342*	.180
Positive Ranks	0	.00	.00		
Equal	4				
TOTAL	6				

*Based on positive ranks.

The high-accurate participants, on the other hand, performed slightly more accurately on psych verbs (M = 11.833; SD = 0.408) than on unaccusative verbs (M = 11.333; SD = 0.816). In order to answer **RQ3**, one more Wilcoxon Signed-Ranks Test was conducted in order to compare the performances of the high-accurate participants for psych verb (Mdn = 12) and unaccusative verb (Mdn = 11.5) constructions. The results (Table 9) indicated that psych verb ranks were not statistically significantly higher than unaccusative-verb ranks (Z = –1.342, p > .05).

Table 10.

Descriptive statistics: Psych-v and unaccusative-v in WFRT

	N	Min.	Max.	Mean	Std. Deviation	Std. Error Mean
WFRT_Psych-V	33	1	4	3.300	0.825	0.143
WFRT_Unaccusative-V	33	1	4	3.442	0.718	0.125

The values in descriptive statistics (Table 10) indicate that the ELT students are more familiar with unaccusative verbs (M = 3.442, SD = 0.718) than psych verbs (M = 3.300, SD = 0.825). In addition to descriptive statistics, a nonparametric test (i.e., a Wilcoxon Signed-Ranks Test) was also conducted in order to answer **RQ4** (i.e., *Are first-year ELT students more familiar with any verb type including psych verbs or unaccusative verbs?*). The aim was to compare the participants' degree of familiarity with psych verbs (Mdn = 3.666) and unaccusative verbs (Mdn = 3.666).

Table 11.

Wilcoxon Signed-Rank Test: Familiarity with psych-v vs. unaccusative-v

WFRT_Unaccusative-V WFRT_Psych-V	N	Mean Rank	Sum of Ranks	Z	p
Negative Ranks	7	12	84	–1.650*	.099
Positive Ranks	16	12	192		
Equal	10				
TOTAL	33				

* Based on negative ranks.

The results (Table 11) indicated that unaccusative verb ranks were not statistically significantly higher than psych verb ranks (Z = –1.650, p = .099). Although a slight difference is observed between familiarity means of two types of verbs (Table 10), this difference is not at any significant level. Meanwhile, Table 12 below also indicates that the familiarity ranks of two types of verbs (i.e., psych (six verbs) and unaccusative (six verbs)) do not also match up with the accuracy ranks of the same verbs in GJT, which has been detailed below.

Table 12.

Familiarity (WFRT) and accuracy (GJT) ranks & means of psych-v and unaccusative-v

Verb	Familiarity Rank	Mean_WFRT	Accuracy Rank	Mean_GJT
Happen_Unaccusative-V	1	3.67	2	1.787
Surprise_Psych-V	2	3.59	9	1.484
Shock_Psych-V	3	3.55	3	1.727
Die_Unaccusative-V	4	3.52	1	1.818
Disappoint_Psych-V	4	3.52	11	1.393
Fall_Unaccusative-V	5	3.48	5	1.666
Disappear_Unaccusative-V	6	3.42	7	1.575
Arrive_Unaccusative-V	7	3.38	10	1.424
Embarrass_Psych-V	8	3.19	4	1.697
Occur_Unaccusative-V	9	3.18	8	1.515
Frighten_Psych-V	10	3.12	6	1.606
Frustrate_Psych-V	11	2.91	8	1.697

In Table 12, the familiarity and accuracy ranks/means for individual psych and unaccusative verbs have been reported. The familiarity ranks of the verbs have been tabulated in a descending order, which means that the participants are found to be most familiar with *happen* and least familiar with *frustrate*. However, the verb that the participants performed on most accurately is *die*, while *disappoint* is observed to be the most challenging one. All in all, the values in Table 13 indicate very little correspondence between the familiarity and accuracy ranks/means of the verbs.

The answer to the **RQ5** (i.e., *Is there any relationship between first-year ELT students' degree of familiarity and level of accuracy with psych and unaccusative verbs?)* also supports the finding above (Table 13) because the participants' degree of familiarity with and accuracy level for psych verbs (r = .174; p = .203 > .05) do not correlate with each other at any significant level. The same result is also valid for unaccusative verbs since the participants' degree of familiarity with and accuracy level for unaccusative verbs (r = –.257, p = .262 > .05) do not also correlate with each other at any significant level. So much so that, unaccusative verbs show a negative correlation between its degree of familiarity and accuracy level.

In addition to grouping the two verb categories, degree of familiarity and accuracy level of each verb was compared individually too. The results indicated no significant relationship between familiarity degrees and accuracy levels of six different psych verbs (disappoint (r = –.271, p = .127); embarrass (r = .096, p = .596); frighten (r = .325, p = .065); frustrate (r = .272, p = .126); shock

(r = .120, p = .508), surprise (r = .068, p = .705)). The same result was also observed for six different unaccusative verbs (arrive (r = –.028, p = .877); die (r = –.247, p = .166); disappear (r = –.212, p = .236); fall (r = .085, p = .636); happen (r = .181, p = .313); occur (r = –.111, p = .538).

Table 13.

Accuracy for psych-v constructions: Number & rate of the participants

			Items	*N* (33)	[%]
Item 8	(G)	P	*The little girl **was frightened**, so she called the police.*	31	93.9
Item 13	(G)	P	*He **was embarrassed** because he could not speak English.*	30	90.9
Item 35	(G)	P	*Everyone **was shocked** to hear the news.*	30	90.9
Item 23	(G)	P	*The people **were frustrated** because they lost all the money.*	28	84.8
Item 16	(UNG)	A	*All the people **shocked** to hear the news.*	27	81.8
Item 21	(G)	P	*The hikers **were surprised** to see a big bear.*	27	81.8
Item 26	(UNG)	A	*The teacher **embarrassed** because she made a mistake.*	26	78.8
Item 19	(G)	P	*The girl **was disappointed** when she found her grade.*	23	69.7
Item 31	(UNG)	A	*The boy **disappointed** because he got a bad grade.*	23	69.7
Item 3	(UNG)	A	*The woman **surprised** to hear the loud noise.*	22	66.7
Item 29	(UNG)	A	*The cat **frightened** when it saw a big dog.*	22	66.7
Item 11	(UNG)	A	*The woman **frustrated** because she could not find help.*	22	66.7

G = Grammatical, UNG = Ungrammatical, P = Passive, A = Active

Table 13 shows the number and rate of participants accurately responding to the twelve items of psych verb constructions in GJT. As the symbols *N* and [%] indicate, the number and rates of participants making accurate judgements for the listed items are higher for the passive construction of psych verbs than active construction of it. The first four lines followed by the sixth and eighth lines of the table include passive constructions with higher number and rates of participants responding accurately to the related items. This finding has also been verified by the nonparametric test results reported below.

Table 14.

Wilcoxon Signed-Rank Test: Active vs. passive constructions of psych-v

GJT_Psych-V_Active GJT_Psych-V_Passive	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	15	13.6	204	–2.582*	.010
Positive Ranks	7	7	49		
Equal	11				
TOTAL	33				

* Based on positive ranks.

In order to answer **RQ6** (i.e., *Do first-year ELT students perform more accurately on any construction of psych and unaccusative verbs (i.e., passive vs. active constructions)?)*, a Wilcoxon Signed-Ranks Test was conducted in order to compare the participants' degree of accuracy for the two types of psych verb constructions: passive (M = 5.121; SD = 1.192; Mdn = 5) and active (M = 4.303; SD = 1.811; Mdn = 5). The results (Table 14) indicated that psych verb passive construction ranks were statistically significantly higher than psych verb active construction ranks (Z = –2.582, $p \leq .01$).

Table 15.

Accuracy for unaccusative-v constructions: Number & rate of the participants

			Items	*N* (33)	[%]
Item 30	(G)	A	*Their parents **died of** cancer when they were young.*	32	97
Item 25	(G)	A	*A terrible accident **happened** last night.*	31	93.9
Item 18	(UNG)	P	*Several storms **were occurred** in Seattle last year.*	30	90.9
Item 7	(G)	A	*The boy **fell into** the swimming pool.*	28	84.8
Item 9	(UNG)	P	*Her mother **was died of** cancer 5 years ago.*	28	84.8
Item 37	(UNG)	P	*So many things **were happened** to her last week.*	28	84.8
Item 27	(UNG)	P	*The girl **was disappeared into** the forest.*	27	81.8
Item 33	(UNG)	P	*The cup **was fallen from** the table.*	27	81.8
Item 20	(UNG)	P	*Your package **was arrived** a few hours ago.*	26	78.8
Item 15	(G)	A	*The thief **disappeared into** the dark.*	25	75.8
Item 34	(G)	A	*A letter from her mom **arrived** two days ago.*	21	63.6
Item 4	(G)	A	*Several fires **occurred** in Seoul last year.*	20	60.6

G = Grammatical, UNG = Ungrammatical, P = Passive, A = Active

Table 15 shows the number and rate of participants accurately responded to the twelve items of unaccusative verb constructions in GJT. As the symbol *N* and [%] indicate, the number and rate of participants making accurate judge-

ments for the listed items do not show a certain accumulation neither for passive construction nor for active construction of unaccusative verbs. The first two lines followed by the fourth of the table include active constructions while the third and the following five lines include passive constructions. This discrete distribution has also been verified by the nonparametric test results reported below.

Table 16.

Wilcoxon Signed-Rank Test: Active vs. passive constructions of unaccusative-v

GJT_Unaccusative-V_Active GJT_Unaccusative-V_Passive	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	14	10.21	143	–1.481*	.139
Positive Ranks	6	11.17	67		
Equal	13				
TOTAL	33				

* Based on positive ranks.

For **RQ6,** one more Wilcoxon Signed-Ranks Test was also conducted in order to compare the participants' level of accuracy for the two types of unaccusative verb constructions: passive (M = 5.030; SD = 1.237; Mdn = 5) and active (M = 4.757; SD = 0.969; Mdn = 5). The results (Table 16) indicated that unaccusative verb passive construction ranks were not statistically significantly higher than unaccusative verb active construction ranks (Z = −1.481, p = .139).

Table 17.

Wilcoxon Signed-Rank Test: unaccusative-v passive constructions vs. psych-v active constructions

GJT_Unaccusative-V_Passive GJT_Psych-V_Active	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	8	9.75	78	–1.844*	.065
Positive Ranks	15	13.20	198		
Equal	10				
TOTAL	33				

For **RQ6,** one more Wilcoxon Signed-Ranks Test was also conducted in order to compare the participants' level of accuracy for the passive constructions of unaccusative-verb (M = 5.030; SD = 1.237; Mdn = 5) and active constructions of psych verbs (M = 4.303; SD = 1.811; Mdn = 5). The results (Table 17) indicated that unaccusative verb passive construction ranks were not statistically significantly higher than psych verb active construction ranks (Z = –1.844, p = .065).

Table 18.

Wilcoxon Signed-Rank Test: unaccusative-v active constructions vs. psych-v passive constructions

GJT_Unaccusative-V_Active GJT_Psych-V_Passive	*N*	Mean Rank	Sum of Ranks	*Z*	*p*
Negative Ranks	15	11.77	176	–1.677*	.094
Positive Ranks	7	10.93	76.5		
Equal	11				
TOTAL	33				

* Based on positive ranks.

For **RQ6,** the last Wilcoxon Signed-Ranks Test was conducted in order to compare the participants' level of accuracy for the active constructions of unaccusative verb (M = 5.121; SD = 1.192; Mdn = 5) and passive constructions of psych verbs (M = 4.757; SD = 0.969; Mdn = 5). The results (Table 18) indicated that unaccusative verb active construction ranks were not statistically significantly higher than psych verb passive construction ranks (Z = –1.677, p = .094).

Discussion

Hypothesis 1 claimed that it would be difficult to acquire psych verb constructions than unaccusative verb constructions since passives charge processing cost. For this reason, the participants had been hypothesized to be making more underpassivization errors than overpassivization errors. The above results confirmed the first hypothesis of the study to some extent. The descriptive statistics indicated that the participants performed more accurately on unaccusative verbs than on psych verbs; however, this accuracy difference was not statistically significant (RQ1). Therefore, there is still a possibility of this result to emerge by chance. Also, the result related to the participants' accuracy on active and passive constructions of psych verbs indicated that psych verb passive construction ranks were statistically significantly higher than psych verb active construction ranks. This result shows us that the participants did more underpassivization errors by accepting the ungrammatical active constructions of psych verbs (e.g., The cat *frightened* when it saw a big dog) as correct and by performing more accurately on passive constructions of psych verbs. (e.g., The little girl *was frightened*, so she called the police). In contrast, the difference between participants' accuracy levels for passive (e.g., The girl *was dis-*

appeared into the forest) and active (e.g., The thief *disappeared into* the dark) unaccusative verb constructions were not statistically significant.

The above analyses also demonstrated that the performances on psych verbs and unaccusative verbs go parallel with their general accuracy levels in GJT. Once the participants' accuracy levels in GJT increased, their underpassivization errors, in other words, their errors with psych verb constructions decreased synchronously. The same result is valid for unaccusative verb constructions with an exception that is the low and medium level participants' performances are close to each other for this verb class. This exception can suggest that unaccusative verbs are acquired at the later levels of proficiency because low and moderate-level student performances do not differentiate from each other.

When the performances of low, medium and high-level participants on psych verbs and unaccusative verbs were compared, it was found that the low, medium and high-level participants did not perform differently on psych and unaccusative verbs. In other words, the low-levels performed less accurately both on psych and unaccusative verbs while the high-levels performed more accurately on both verb types. This suggests that the difficulty level of psych and unaccusative verb constructions do not vary for the same level of learners. Therefore, once learners reach a stage, both underpassivization and overpassivization errors reduced, the gap in the performances between psych and unaccusative constructions did not enlarge and none of them got any advantage over the other. Therefore, the Hypothesis 2 was not confirmed by the current results.

As for the Hypothesis 3, it is not possible to regard degree of familiarity as a source of accuracy for psych and unaccusative verbs. The results firstly indicated that the participants' degree of familiarity with two verb types did not vary from each other. Moreover, their degree of familiarity with and accuracy level for unaccusative and psych verbs do not correlate with each other at any significant level. Even, these two variables demonstrated a reverse relationship for some unaccusative (i.e., *arrive*, *die*, *disappear*, *occur*) and psych (i.e., *disappoint*) verbs. Therefore, it can be concluded that the participants' performance on psych and unaccusative verbs in the GJT was not influenced by their perceived degree of familiarity with these verb types, which confirms the third hypothesis of the study.

The performance difference raised between active (ungrammatical) and passive (grammatical) constructions of psych verbs implies that the participants do not have as much difficulty in approving the grammatical (passive) constructions as they do in rejecting the ungrammatical (Active) constructions. They were, instead, inclined to regard active constructions (e.g., The cat *frightened* when it saw a big dog) as correct. As the participants' accuracy for psych verbs do not correlate significantly with their familiarity, and as their performance on active and passive constructions of psych verbs differentiate at a significant level, the participants' responses to two different constructions

rather than individual verbs can be interpreted as their actual performances. As seen in Table 13, except "disappoint" (since the number/ratio of participants responding accurately for two constructions were equal for "disappoint"), all the passive constructions of five psych verbs were outperformed in contrast with their active constructions.

As for unaccusative verbs, the participants' level of accuracy for passive and active constructions were not statistically significantly different from each other. Instead, it is possible to observe some differences in verb level. For instance, while the grammatical constructions (i.e., active) of the verbs *die*, *fall*, and *happen* are accepted more accurately than their ungrammatical constructions (i.e., passive), it is totally opposite for the other three verbs (i.e., *arrive*, *disappear*, *occur*). For these verbs, it was more difficult for the participants to approve grammatical items than rejecting ungrammatical ones. Therefore, it is possible to state that the participants' performances vary across different verbs. One possible explanation for this variety can be the different levels of exposure to these verbs throughout their language acquisition process.

The potential support for this rationale can come from the lists of word frequencies in English. While *die*, *fall*, and *happen*, for example, are among the frequent words according to the two common wordlists (i.e., New General Service List (NGSL) and BNC/COCA Headword List), the other three verbs, *arrive*, *disappear* and *occur*, are in higher rankings of these lists (see Table 19). Also, English Vocabulary Profile (EVP) reports that *die*, *fall*, and *happen* are used by L2 English learners at the levels of A1, A2, and A2 respectively while *arrive*, *disappear* and *occur* are found to be used by the learners at A2, B1, and B2 levels.

Table 19.

Accuracy (GJT), familiarity (WFRT) and frequency (NGSL, BNC/COCA Headword List, EVP) ranks of psych-v and unaccusative-v

Verb	Accuracy Rank	Familiarity Rank	NGSL	BNC/COCA Headword List	EVP
Die_Unaccusative-V	1	4	493	1st 1000	A1
Happen_Unaccusative-V	2	1	245	1st 1000	A2
Shock_Psych-V	3	3	4118	2nd 1000	B2
Embarrass_Psych-V	4	8	2576	2nd 1000	C2
Fall_Unaccusative-V	5	5	353	1st 1000	A2
Frighten_Psych-V	6	10	2249	1st 1000	B2
Disappear_Unaccusative-V	7	6	1762	2nd 1000	B1
Frustrate_Psych-V	8	11	4913	2nd 1000	-
Occur_Unaccusative-V	8	9	728	2nd 1000	B2
Surprise_Psych-V	9	2	684	1st 1000	B1
Arrive_Unaccusative-V	10	7	574	1st 1000	A2
Disaapoint_Psych-V	11	4	1973	2nd 1000	B1

The significant difference between the participants' performances on active and passive constructions of psych verbs, the participants', though non-significant, outperformance on unaccusative verbs as well as the accuracy on unaccusative verbs varying based on individual verbs and constructions imply that processing passivization cannot be the only effective factor determining the acquisition of psych and unaccusative verb constructions. Specifically, cross-linguistic factors and L1 interference can be some possible sources of the challenges faced by this sample group as indicated by Montrul (2001) who found that the crosslinguistic difference in transitivity between English and Turkish was responsible for Turkish learners' ungrammatical use of psych verbs because these learners used active psych-verb form where the Experiencer is the subject (e.g., *The hunter frightened).

Conclusion

The processability account anticipates that learners will make more under-passivization errors than overpassivization errors since passivization entails more processing. In a scenario where processability plays a role in the acquisition of English as an L2, passive constructions requiring more processing will be more difficult to acquire as compared to active constructions. As a result, L2 learners of English can potentially make more errors with passive constructions, favor active constructions over passive constructions, and even wrongly accept or use active constructions in the context where passives are necessary. In order to investigate these points, a set of research questions has been asked and answered.

Firstly, the results indicated that ELT students did not significantly performed more accurately on psych verbs or unaccusative verbs. The second point handled was whether the performances on psych and unaccusative verbs differ across the accuracy levels in GJT. The results showed that as the total level of accuracy in GJT increased, the level of accuracy with psych verbs and unaccusative verbs increased, too. They also indicated significant differences between the unaccusative verb scores of *low* and *high* accurate, and *medium* and *high* accurate participants while the difference is not significant between the scores of *low* and *medium* accurate participants. As for psych verbs, the results revealed significant differences between the psych verb scores of *low* and *medium* accurate, *low* and *high* accurate, and *medium* and *high* accurate participants. Thirdly, the comparison between the performances of the low-accurate participants for psych verbs and unaccusative verbs indicated that unaccusative verb ranks were not statistically significantly higher than psych

verb ranks. The same result is also valid for the medium-accurate and high-accurate participants. Fourthly, when the familiarity with psych and unaccusative verbs was investigated, the results indicated that ELT students were not more familiar with any verb type. Also, the familiarity ranks of two verbs did not also match up with the accuracy ranks of the same verbs in GJT. The participants' degree of familiarity with and accuracy level for two verbs do not correlate with each other at any significant level. Lastly, the study compared the participants' degree of accuracy for two types of verb constructions (i.e., passive and active), the results indicated that active construction ranks were statistically significantly different than passive construction ranks for each verb. All in all, the results suggested that beyond processability, such factors as L1 interference and word frequencies in English (Table 19) seem to impact the acquisition process.

References

Browne, C., Culligan, B., & Phillips, J. (2013). The New General Service List. Retrieved from: http://www.newgeneralservicelist.org

Hahn, Hye-ryeong, (2011). Korean EFL learners' underpassivization and overpassivization: Psych verbs and unaccusatives. *Foreign Language Education Research, 14*, 54–76.

Kawaguchi, S. (2013). The relationship between lexical and syntactic development in English as a second language (pp. 92–106). In A. F. Mattsson & C. Norrby (Eds.), *Language acquisition and use in multilingual contexts: Theory and practice*. Media-Tryck, Lund University.

Kurtoğlu, Ö. (2010). (Over)passivization errors of Turkish learners of English as a foreign language. *Journal of Linguistics and Literature, 7*(1), 29–53.

Kurtoğlu, Ö. (2006). *A study on the passivisation errors of Turkish learners of English as a foreign language* (Master's Thesis). Retrieved from Council of Higher Education Thesis Center (205789) https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Levin, B. (1993). *English verb classes: A preliminary investigation*. University of Chicago Press.

Montrul, S. (2001). First-language-constrained variability in the second-language acquisition of argument-structure-changing morphology with causative verbs. *Second Language Research, 17*(2), 144–194.

Nation, I. S. P. (2012). The BNC/COCA word family lists. Retrieved on September, 2018 from: https://www.victoria.ac.nz/lals/about/staff/paul-nation#vocab-lists

Pienemann, M. (2005). An introduction to processiblity theory. In M. Pienemann (Ed.,), *Cross-linguistic aspects of processibilty theory* (pp. 1–60). John Benjamins.

Pienemann, M., D Biase, B., & Kawaguchi, S. (2005). Extending processiblity theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of processiblity theory* (pp. 199–251). John Benjamins.

The English Vocabulary Profile (n.d.). Retrieved from: http://www.englishprofile.org/

Appendices

Appendix A

Grammaticality Judgement Task (GJT) – Hahn (2011)

Participant Name: …….. **Year of Study :** (1st) (2nd) (3rd) (4th)
Gender : ( ) Female ( ) Male

INSTRUCTION: This study is conducted as a preliminary study for a PhD dissertation to investigate if the acquisition of specific constructions is affected by processing complexity of passivization. Please put a mark next to the sentence if you think that the test sentence is **Grammatical (Correct)** or **Ungrammatical (Incorrect)**. Make your decisions only once, and do not turn back and correct your original answers after a second thought. Thank you very much for your cooperation and contribution.

Certificate of Consent

I have been invited to participate in this research about acquisition of specific constructions in English. I have read the related details above and consent voluntarily to be a participant in this study.

Date: …. / …. / ……..
Signature:

	CORRECT	INCORRECT
Example: 1) The dog eating the meat.		X
Example: 2) The dog was eating the meat.	**X**	
3) The woman surprised to hear the loud noise.		
4) Several fires occurred in Seoul last year.		
5) This is the company that he likes to work.		
6) Jane told us that she was going to church.		
7) The boy fell into the swimming pool.		
8) The little girl was frightened, so she called the police.		
9) Her mother was died of cancer 5 years ago.		
10) The boy who finished all the assignments were so tired.		
11) The woman frustrated because she could not find help.		
12) The doctor asked him if he feels okay.		
13) He was embarrassed because he could not speak English.		
14) This is the topic that we want to talk about.		
15) The thief disappeared into the dark.		
16) All the people shocked to hear the news.		
17) Mr. Jones asked me what the problem was.		
18) Several storms were occurred in Seattle last year.		
19) The girl was disappointed when she found her grade.		
20) Your package was arrived a few hours ago.		
21) The hikers were surprised to see a big bear.		
22) Mary told her father that she has a fever.		
23) The people were frustrated because they lost all the money.		

24) I asked him what kind of movies does he like.		
25) A terrible accident happened last night.		
26) The teacher embarrassed because she made a mistake.		
27) The girl was disappeared into the forest.		
28) The little boy who passed all the subjects was very proud.		
29) The cat frightened when it saw a big dog.		
30) Their parents died of cancer when they were young.		
31) The boy disappointed because he got a bad grade.		
32) People who have a lot of friends do not have time to watch TV.		
33) The cup was fallen from the table		
34) A letter from her mom arrived two days ago.		
35) Everyone was shocked to hear the news.		
36) The people in the room didn't know what was the problem.		
37) So many things were happened to her last week.		
38) People who have only one child wants to have more children.		

Appendix B

Word Familiarity Rating Task (WFRT) – Hahn (2011)

INSTRUCTION: Please mark the degree of your familiarity with the individual words on a 4-point Likert scale:

(1) = Very Familiar
(2) = A Bit Familiar
(3) = Quite Familiar
(4) = Very Familiar

	(1) Very Unfamiliar	**(2) A Bit Familiar**	**(3) Quite Familiar**	**(4) Very Familiar**
Example: eat			X	
surprise				
happen				
arrive				
disappear				
shock				
die				
disappoint				
occur				
embarrass				
fall				
frighten				
frustrate				

Appendix C

Accuracy for the structures used as distracters in GJT: Number & rate of the participants accurately answering

		Items	*M*	*N* (33)	[%]
Item 14	(G)	Geography is the subject that Joseph wants to take.	.969	32	97
Item 28	(G)	*The little boy who passed all the subjects was so proud.*	.969	32	97
Item 32	(G)	Students who do not pass the math exam must take math again.	.939	31	93.9
Item 17	(G)	Mrs. Jones asked me what the problem was.	.909	30	90.9
Item 6	(G)	Jane told us that she was going to church.	.727	24	72.7
Item 10	(UNG)	The young woman who finished all the assignments were so tired.	.727	24	72.7
Item 38	(UNG)	Students who does not want to take geography can take history.	.696	23	69.7
Item 24	(UNG)	I asked what kind of movies does he like.	.636	21	63.6
Item 36	(UNG)	The people in the room didn't know what was the problem.	.575	19	57.6
Item 22	(UNG)	Mary told her father that she has a fever.	.484	16	48.5
Item 5	(UNG)	A café is a small restaurant that people can get a light meal.	.454	15	45.5
Item 12	(UNG)	The doctor asked him if he feels okay.	.454	15	45.5

G = Grammatical, UNG = Ungrammatical (*Min.* = 0; *Max.* = 1)

Seray Tanyer, Samet Deniz

Der Erwerb von psychologischen und unakkusativen Verben durch türkische EFL-Studenten: eine Replikationsstudie zur Unterpassivierung und Überpassivierung

Zusammenfassung

Die Komplexität des Transformierbarkeit legt die Vermutung nahe, dass Lernende mehr Fehler hinsichtlich der Unterpassivierung als Überpassivierung begehen werden, weil die Passivierung einen höheren Aufwand an Transformation erfordert. Obwohl eine Studie zu psychologischen und einigen unakkusativen Verben deren Erwerb durch türkische Fremdsprachenlernende untersuchte, wurde bisher in keiner Arbeit der Erwerb der beiden genannten Verbgruppen durch Lernende unter dem Gesichtspunkt der Transformation einzeln analysiert. In diesem Zusammenhang ist die vorliegende Studie darauf abgezielt, zu untersuchen, ob die Komplexität der Passivtransformation den Erwerb von psychologischen und

unakkusativen Verben beeinflusst. Außerdem wird die Frage behandelt, ob das allgemeine Genauigkeitsniveau in Grammaticality Judgement Task (GJT) und der Grad der Vertrautheit mit Zielverben mit dem Genauigkeitsniveau in Bezug auf die einzelnen psychologischen und unakkusativen Verben zusammenhängen. 33 Universitätsstudenten im Grundstudium nahmen an einem GJT und Word Familiarity Rating Task (WFRT) teil. Der GJT umfasste insgesamt 38 Punkte, und zwar 12 Sätze mit psychologischen Verben, 12 Sätze mit unakkusativen Verben, 12 Sätze mit Distraktoren und 2 Beispielsätze. Der WFRT war ein Fragebogen zur Vertrautheit mit 6 psychologischen und 6 unakkusativen Verben. Zur Datenauswertung wurden eine Reihe von nichtparametrischen Tests sowie deskriptive Statistiken verwendet. Die Ergebnisse machten deutlich, dass die Lernenden mit unakkusativen Verben genauer als mit psychologischen umgehen konnten. Sie begingen mehrere Unterpassivierungsfehler, indem sie ungrammatische Aktivkonstruktionen mit psychologischen Verben zuließen. Ihre Leistungen in Bezug auf die psychologischen und unakkusativen Verben zeigten Parallelen zu ihrem allgemeinen Genauigkeitsniveau in GJT, während ihr Grad der Vertrautheit mit und Genauigkeitsniveau von zwei Verben nicht miteinander korrelierten. Die Ergebnisse deuten darauf hin, dass Faktoren wie Transformierbarkeit und muttersprachlicher Transfer den Erwerb zu beeinflussen vermögen.

Schlüsselwörter: Fremdsprachenerwerb; psychologische Verben; unakkusative Verben; Unterpassivierung; Überpassivierung

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 157–176
https://doi.org/10.31261/TAPSLA.9152

Tan Arda Gedik https://orcid.org/0000-0003-1429-9675
Friedrich-Alexander Universität Erlangen-Nürnberg, Germany
Yağmur Su Kolsal https://orcid.org/0000-0002-2659-4447
Middle East Technical University, Turkey

A Corpus-based Analysis of High School English Textbooks and English University Entrance Exams in Turkey

Abstract

This study explores the disconnect between the English textbooks studied in high schools (9th–12th grades) and the English tested on Turkish university entrance exams (2010–2019). Using corpus linguistics tools such as AntWordProfiler, TAALED, and the L2 Syntactic Complexity Analyzer (L2SCA), this paper analyzes the lexical diversity and syntactic complexity indices in the sample material. A comparison of official textbooks and complementary materials obtained from the Ministry of National Education against the official university entrance exams demonstrates that: (i) differences in lexical sophistication level can be observed between the two corpora, the lexical sophistication level of the exam corpus was higher than that of the textbook corpus, (ii) there is a statistically significant difference between the two corpora in terms of lexical diversity, the exam corpus has a significantly higher level of lexical diversity than the textbook corpus, (iii) statistically significant differences also existed between the two corpora regarding the syntactic complexity indices. The syntactic complexity level of the exam corpus was higher than that of the textbook corpus. These findings suggest that Turkish high school student taught English with official textbooks have to tackle low-frequency and more sophisticated words at a higher level of syntactic complexity when they take the nationwide exam. This, in turn, creates a negative backwash effect, distorting their approach to L2, and raising other concerns about the misalignment between the official language education materials and nationwide exams.

Keywords: corpus linguistics, lexical diversity, syntactic complexity

Textbooks and Exams

English language teaching in Turkey has been a topic for long hours of debate in many layers of the society. With this in mind, the English curriculum in Turkey has witnessed many changes over the years (Hatipoğlu, 2016). The most drastic change in the recent years has been the lowering of the grade in which students learn English, the first foreign language to be taught at schools, from 4th to 2nd. In addition, the change in educational model which experienced a shift from a eight years of elementary school and four years of high school type of division of grades to four years of primary, four years of middle and four years of high school. This has required many to adopt a different approach to language learning. The national curriculum claims that the new model accommodates these changes, and the textbooks used in Turkish English as a Foreign Language (EFL) setting have also been tweaked and enhanced over the years. The national curriculum for English language for the term of 2018–2019 by Ministry of Education also states that the new curricular model puts emphasis on the use of authentic language in an authentic context, a consideration, the importance of which Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) emphasizes.

The main goal of the new English curriculum for secondary schools is engaging learners of English in stimulating, motivating, and enjoyable learning environments to render them independent, fluent, and effective users of the language (Milli Eğitim Bakanlığı, 2018). Rather than adopting a singular teaching methodology, the curriculum sets recurring teaching and language principles which are based on the acknowledgment of the international status of English, the components of communicative competence and the integration of four main language skills. These claims of an enhanced educational model for the textbooks is very important in an EFL context, since textbooks are among the most widely used EFL teaching materials (Allen, 2008). The marked presence of textbooks in EFL classrooms signifies the need for analyzing the content and problems associated with the success of the EFL programs (Choi, 2008).

Textbooks can be considered a route map for any English language teaching (ELT) program: not only sources of information but also a factor influencing the program's structure and destination. A wrong selection can later be a source of regret. That holds true for government-imposed books, which give little opportunity for modification (Sheldon, 1998). In a wide variety of occasions in many countries, textbooks are designed with the aim of preparing the students for standardized tests, and while this widespread tendency in EFL can be a source of criticism, textbooks need to fulfill that aim. In Turkey, textbooks are mainly

used to prepare students who are to take high stakes exams. These exams are also referred to as nationwide university entrance exams. The textbooks are provided across Turkey at the beginning of each semester, free of charge, to establish equality (Gençoğlu, 2017). Some scholars have analyzed the discrepancies and a lack of correspondence between English textbooks and high stakes university entrance exams for English in various other contexts (Underwood, 2010; Tai & Chen, 2015; Nur & Islam, 2018). Although the English textbooks used at Turkish high schools are not directly aimed at addressing the English university entrance exam, the textbooks are handed out as an aid to improve students' overall proficiency. The exam, on the other hand, is a multiple-choice proficiency exam without subsections that test productive skills such as speaking and writing.

In the light of these, to achieve academic success in Turkey, students are obligated to succeed in the nationwide exam, but are the textbooks adequately preparing the students to cope with the exams? To the best of the researchers' knowledge, no study has scrutinized the lexical and syntactic complexity of high school English textbooks and the university entrance exams from a statistical standpoint so far within the Turkish context. Hence, this study aims to analyze English high school textbooks and the complementary materials that are currently in use throughout the country and English university entrance exams that were administered in the past ten years in terms of lexical sophistication, lexical diversity, and syntactic complexity using corpus linguistics analysis tools.

In sum, the current study aims to serve as (i) a non-biased source of findings while bridging the research gap, (ii) a gateway between the exam preparation committee and the textbook writers, (iii) the voice of students who struggle with vocabulary item and syntactic differences between the textbooks and exams.

Literature Review

English Language Teaching and Testing Situation in Turkey

The situation of EFL teaching in Turkey is a troubled area. Kırkgöz (2007) mentions that with Turkey's negotiations with the EU, English saw a rise of importance (e.g., to comply with the EU regulations like CEFR leveled textbooks). Attempts at accommodating for the rising importance of English competencies include international collaborations with schools in the EU in addition to modification of textbooks according to the new model. These factors have been the primary influences on the EFL teaching situ-

ation in Turkey. Kırkgöz (2007) also mentions two phases: 1863–1997, 1997 and onwards. 1863 marks the beginning of ELT in what was back then the Ottoman Empire. The year 1997, on the other hand, was of great importance as the compulsory grade in which English was taught was lowered from 6th to 4th grade. In other words, the content of many textbooks had to be re-evaluated, and this was another significant change to the EFL teaching situation in the near past. This could be associated with the never-ending change of ELT policies which attempt to make foreign language education better and increase the level of proficiency among school-age children, and as a result, the general demographic in Turkey.

As previously mentioned, Hatipoğlu (2016) mentions that Turkey has "one important high-stakes exam, which determines whether students gain entry to prestigious colleges or tertiary institutions" (p. 2). The study done by Hatipoğlu (2016) also reveals that a big number of pre-service teachers believe that high stakes exams play a dramatically life-changing role in one's future. Furthermore, it is revealed that due to the detrimental consequences of the negative backwash effect of unplanned high stakes exams and changes to the curriculum, many students regard English as a sum of the parts they separately learn. Hatipoğlu (2016) claims the following for the EFL teaching situation in Turkey:

> The short historical overview presented in the first part of the paper reveals an unsettled and frequently changing system where, in majority of the situations, changes were not based on empirical research, educational theories, or assessment models but rather on political and practical reasons. This reveals an inadequate understanding and skewed interpretation of testing and assessment. (p. 142)

Comparing English Language Testing and Teaching Materials in Other Contexts

English language testing is a topic that cannot be overlooked. Using multiple-choice based exams has been widely accepted as a way of testing many subjects, and English is not an exception. Many countries conduct various university entrance exams that utilize multiple-choice questions. Moreover, the lack of correspondence between textbooks and university entrance exams seems to be a recurring theme among other countries. In a study done by Nur and Islam (2015) in Bangladesh, the findings highlighted a clear disconnect between the intended English assessment policy directions and the practiced pattern. The analysis of data also indicated that a backwash of such "disconnect between policy and practice substantially intercedes the overall quality of secondary English education" (p. 100).

Underwood (2010) conducted a similar study to the present one in Japan, comparing English textbooks and the Japanese university entrance exam for English. Underwood (2010) states that over the years, there has been a greater alignment between the textbooks and exams in terms of readability and lexical sophistication. Nevertheless, Underwood (2010) notes that there is still more improvement required in terms of lexical overlap between the analyzed materials.

Another different approach to the same topic was carried out by Tai and Chen (2015) in Taiwan. Their study compared English textbooks in high schools to the national university entrance exam, and the frequency of marked structures, namely relative, adverbial, and passive clauses, was attained by utilizing AntConc and Readability Test Tool. In other words, their study scrutinized the two corpora from a syntactic analysis point-of-view. They reported statistically significant results between the corpora. Although there have been many studies analyzing the relationship between syntactic complexity and L2 writing (Lu & Ai, 2015, Kyle, 2016; Kyle & Crossley, 2018), studies that scrutinize syntactic complexity levels to compare exams to textbooks have been very few (Mirshojaee & Sahragard, 2015). Nevertheless, these findings, where textbooks and exams are compared, demonstrate a lack of correlation between the above-mentioned corpora and affirm the fact that "skewed interpretation of testing and assessment" (Hatipoğlu, 2016, p. 142) is a recurring theme in other parts of the world.

Lexical Sophistication, Diversity

Read (2000) determines four different ways of identifying lexical richness: lexical density, lexical diversity, lexical sophistication, and proportion of errors. Lexical sophistication and lexical diversity are two essential terms out of those four for the present investigation as lexical density and proportion of errors are more often researched in corpora that are produced by learners. To measure lexical sophistication, researchers have calculated the total number of advanced or sophisticated words in a text (Laufer & Nation, 1995). Nevertheless, there has not been a consensus on what a sophisticated/advanced word is. Yet, overall, many seem to agree that the use of word frequency as a tool to identify whether a word is advanced or not has been the widely accepted way of approaching this issue (Bardel et al., 2012). Namely, low-frequency words and how many times those appear in a text appear to stand out as the most reliable way of approaching sophisticated words (Hyltenstam, 1988; Laufer & Nation, 1995; Read, 2000; Vermeer, 2004).

Bardel et al. (2012) approach lexical sophistication as the percentage of sophisticated or advanced words in a text, including the first one thousand

(K1), the first two thousand (K2), the first three thousand (K3), and academic word list (AWL) words in the corpora. The researchers argue that the lexical sophistication level(s) of non-native speakers (NNS) of a language can prove to be a source of knowledge when it comes to testing L2 knowledge. In other words, lexical sophistication can be employed as a way of determining whether a NNS has reached native-like proficiency in terms of vocabulary size. Their argument also extends to the vocabulary size of the teaching material employed to teach L2 since the more low-frequency words the learners are exposed to, the higher native-like proficiency they are likely to have. To measure the lexical sophistication level of a text or corpus, a procedure called lexical frequency profiling first carried out by Laufer and Nation (1995), corpus linguistics tools such as AntwordProfiler (Anthony, 2012) are utilized. AntwordProfiler enables finding the coverage of aforementioned word lists in a corpus. In recently conducted studies of Kwary et al. (2018), Du (2019), Beauchamp and Constantinou (2020), AntwordProfiler was used to analyze lexical frequency profiles.

Lexical diversity, on the other hand, refers to "the range of different words used in a text, with a greater range indicating a higher diversity" (McCarthy & Jarvis, 2010, p. 381). The researchers also argue that lexical diversity can be used to determine the "writing quality of a text, vocabulary knowledge, speaker competence, Alzheimer's onset, hearing variation as well as socioeconomic status" (p. 381) of interlocutors in a conversation. Lexical diversity introduces two different sub-terms: type-token ratio (TTR; RootTTR and LogTTR), and the measure of textual lexical diversity (MTLD). While RootTTR and LogTTR are basically calculation of the TTR level of a text using a root and a log formula, in the case of MTLD, the text is divided into segments based on the TTR value of each segment. Each segment finishes when the TTR level reaches .72 (Toruella & Capsada, 2013) and the calculation of MTLD is done by dividing the length of the text in number of words by segments.

These two other terms are introduced because determining the lexical diversity level of a text has been problematic as lexical diversity indices may display sensitivity to the length of a text (McCarthy & Jarvis, 2010). Researchers like Biber (1989) have produced reliable analyses of corpora as they seem to have been aware of this sensitivity, however, researchers such as Ertmer et al. (2002) and Miller (1981) who have not demonstrated their awareness of this issue may have produced misleading analyses of corpora. McCarthy and Jarvis (2010), however, believe that MTLD, RootTTR, and LogTTR results are of a validating nature for analyzing a text and have corrective features and factors that help researchers yield a more reliable analysis. In this study, TAALED version 1.3.1. was used to this end. TAALED (The Tool for the Academic Analysis of Lexical Diversity) is used in calculating the lexical density of a corpus for types and tokens and eight indices of lexical diversity (Kyle, 2018). Studies of

Bulté and Roothooft (2020) and Skalicky et al. (2020) are recent examples of the use of TAALED for lexical diversity analysis.

With all of this mentioned, Crossley et al. (2011) draw on the importance of lexical proficiency explaining parts of lexical proficiency, as a cognitive construct, as exposure to lexically diverse corpora, lexical-semantic relations, and coherence of core lexical items. Thus, lexical proficiency is also a very salient indication of academic success in L2 (Daller, Van Hout, & Treffers-Daller, 2003) that is interconnected with the focus of this paper.

Given the context of the EFL teaching situation not only in Turkey but also in other countries, the following question arises: do English textbooks used in high schools and English university entrance exams correspond to each other in terms of lexical complexity? What is more important is that no matter what kind of approach the institutions follow, if the textbooks and exams do not match in terms of lexical richness (lexical sophistication and diversity in this paper's case), the students are left in a position of disadvantage where what they learn does not prepare them for the examinations. As mentioned, and demonstrated by many scholars (McCarthy & Jarvis, 2010; Crossley et al., 2011; Bardel et al., 2012), lexical richness goes hand in hand with the number of low-frequency words introduced in L2 textbooks and materials. It would be unimaginable to ignore this fact and create textbooks and exams disconnected from each other. This, in turn, would raise another important question in many readers' minds: do we test what we teach? When this is not the case, when what is not taught is being tested or vice versa, many students suffer from what is called a negative backwash effect. This, in turn, demotivates them and distorts their perception of and approach to L2, forcibly changing their notion of language from a tool of communication with which they can create and share to a distorted one on which they must (or are expected to) perform various assigned tasks to be considered proficient.

Syntactic Complexity

Syntactic complexity is one of the crucial elements in language testing and evaluation of L2 learners (Wang & Slater, 2016). To assess the syntactic complexity of a text, sentence level and word level measures have been proposed such as ratio of T-units to clauses and syntactic variety of tenses (Ellis & Yuan, 2004; Larsen-Freeman, 2006; Nelson & Van Meter, 2007; Norrby & Håkansson, 2007). This is because syntactic complexity seems to have become a vital indicator of a text's complexity and comprehensibility (Wang, 1970). Many scholars report that this complexity goes higher in more proficient L2 users (Lu, 2011; McNamara et al., 2010; Ortega, 2003). These L2 users, in correlation with their proficiency, produce syntactically lengthier pieces of texts

compared to less-proficient L2 users (Frase et al., 1999; Grant & Ginther, 2000; Ortega, 2003). A heightened use of subordination was also reported (Grant & Ginther, 2000). Therefore, it is fair to explain syntactic complexity in the lines of "measures such as length of production unit, amount of subordination or coordination, [and] range of syntactic structures" (Kim, 2014, p. 32). Park (2012) suggests that the mean length of clause and sentence as well as the number of complex nominals in clauses and T-units are of salient indicators for L2 proficiency. T-unit is one of the tiniest but most important indexes in evaluating syntactic complexity (Hunt, 1965). Wolfe-Quintero et al. (1998) in their study revealed that mean length of T-unit, dependent clauses, mean number of clauses per T-unit, and mean length of clause were the best indicators of syntactic complexity.

Mean length of clause (MLC) is the average number of words per clause. It can be referred to as a global measure of syntactic complexity. Many studies also point to a salient correspondence between MLC and proficiency levels (Cumming et al., 2005; Ortega, 2003; Wolfe-Quintero et al., 1998). In contrast to MLC, the mean length of T-unit (MLT) builds another layer of specific examination of the complexity. That is, dependent clauses might be indistinguishable in MLC, but MLT, due to its T-unit nature, specifies them. Ortega (2003) and Wolfe-Quintero et al. (1998) demonstrated that just like MLC, MLT also shows great correlation with high proficiency levels. T-units may not always be enough on their own, and another index may be required. A complex T-unit per T-unit (CT/T) is the proposed index by Casanave (1994) and Lu (2011). What makes this a complex T-unit is, this time the T-unit is expected to host an independent and a dependent clause at the same time. However, CT/T is not proven to be statistically significant in relation to language development; in other words, learners' proficiency is not reflected through this index. Nevertheless, the studies (Casanave, 1994; Lu, 2011) done on CT/T only compared the production of L2 learners and thus their proficiency. CT/T has not been examined from the point of language testing and evaluation.

This study attempts to see whether there is a contrast between the two corpora. Complex nominals per T-unit (CN/T) is a syntactic construction that has nominal clauses, nouns with adjectives, possessives, prepositional phrases, and/or infinitives/gerunds. Despite studies not reporting a significant relationship between proficiency and CN/T numbers (Wolfe-Quiero et al., 1998; Lu, 2010), Dean (2017) demonstrates a significant connection between L2 proficiency and CN/T. Table 1 illustrates the definitions of the syntactic indices used in this study based on Lu's (2010) article.

Table 1

Syntactic indices

	Explanation
MLC	Mean length of clauses
MLT	Mean length of T-units
CT/T	# of complex T-units per T-unit
CN/T	# of complex nominals per T-unit

Lu (2010) reported five categories of syntactic complexity measures. These were: length of production unit, amount of subordination, amount of coordination, level of phrasal complexity, and overall sentence complexity. The L2 Syntactic Complexity Analyzer (L2SCA) uses 14 indices based on Lu's (2010) categories. During this study, the following four indices were employed to examine the syntactic complexity levels: MLC and MLT identify the length of the production unit. CT/T identifies the amount of subordination and CN/T examines the degree of phrasal complexity. All these indices have been investigated to seek relations between proficiency and production. However, the current study assumes that textbooks should prepare students on all four indices and that exams should correspond to them. If the textbooks fall behind the exams in terms of syntactic complexity, this will ensure that proficiency levels of the students are not tested on the same level as the textbooks prepare them to be. Furthermore, the three categories addressed in the present study (lexical sophistication, lexical diversity, and syntactic complexity) would affect the comprehension of a text the most, especially in dealing with standardized tests. Quite clearly, comprehension and proficiency are cognitive heavy processes (Kalyuga, 2006). Thus, these indices, because they indicate complexity which affect comprehension and proficiency, may possibly indicate the relation between sentence complexity and syntactic processing of the sentences. Both corpora could be examined in relation to other ten indices as well, but to keep uniformity across the two corpora, the same set of indices were utilized, namely MLT, MLC, CT/T and CN/T.

Hence, the present study aims to examine the following research questions:

(i) Are there statistically significant differences in terms of lexical sophistication and lexical diversity between the textbook and exam corpus?

(ii) Are there statistically significant differences in terms of syntactic complexity between the textbook and exam corpus?

Methodology

To answer the questions above, all data were gathered online either from eba.gov.tr (for English textbooks) or from ösym.gov.tr (for English university entrance exams), ÖSYM being the Measurement, Selection and Placement Center, the sole body responsible for preparing and administering the nationwide entrance exams and the placement of students, while EBA is the online platform where students and teachers alike can access educational content, among which are textbooks. English textbooks and other complementary materials (i.e., corresponding workbooks and listening transcripts) that are currently in use from 9th through 12th grade were identified and downloaded in .pdf format. Meanwhile, English university entrance exams between the years 2010–2019 were identified and downloaded in .pdf format. In total, there were eight textbooks and ten exams. The textbooks covered each grade in high schools (9th–12th grade) and were published by the following publishing houses; (MEB) *Relearn*, *Teenwise*, *Progress* for 9th; *Count Me In*, *Gizem* for 10th; *Sunshine*, *Silverlining* for 11th; and *Count Me In* for 12th grades with their accompanying workbooks. Regardless of the publishing house of the books, the respective CEFR level for grades were as follows: A1–A2 for 9th grade, A2+–B1 for 10th grade, B1+–B2 for 11th grade and B2+ for 12th grade. The total number of tokens in the textbook corpus was 301.255. The ten exams were all prepared and released by ÖSYM between the years of 2010–2019 with a total token number of 66.913. While these books are produced by different publishing houses, they all have to follow the same regulations put forward by MEB, and their products (textbooks) have to go through a series of assessments and evaluation by a committee allocated by MEB itself.

Once the data collection was over, the followings were executed in a progressive order: (a) convert all the .pdf files into .docx files using an online document converter; (b) clean both corpora of any mistakes, typos and unnecessary signs or images which may have been caused by the conversion and may interfere with the results; (c) convert the clean .docx files into compatible .txt files for the analysis tools; (d) run both AntwordProfiler, TAALED and the L2 Syntactic Complexity Analyzer (L2SCA) on all the documents and save the results in .csv files; (e) run the .csv files' output through SPSS for statistical analysis, including descriptive analysis and a series of independent samples t-tests); (f) interpret the results.

While for lexical sophistication, AntWordProfiler (Anthony, 2012) was used to examine both corpora, for lexical diversity, Kristopher Kyle's TAALED version 1.3.1. was employed. TTR, RootTTR, LogTTR, and MTLD were selected as the indices to conduct the comparison between the two corpora. As mentioned in the literature review, because these indices have corrective

features that are required when working with longer texts, they were chosen reliable indices. As for syntactic complexity, the L2SCA (Lu, 2010) was employed to analyze MLT, MLC, CT/T and CN/T because of the following two reasons: (i) the researchers specifically wanted to focus on whether sentence and clause lengths were statistically different across corpora even though the token numbers are vastly different (thus MLT and MLC were selected), (ii) the amount of subordination, as mentioned in the literature review, would affect one's comprehension (hence, CT/T and CN/T were selected). Finding out the differences between the two would then show the researchers whether students are trained well enough for a timed examination regarding decoding syntactically heavily subordinated clauses. Another reason is that the scope of this study would need to be broader to examine all the syntactic indices at once.

Results

Lexical Sophistication and Lexical Diversity

The mean difference between the two corpora regarding the percentage of K1, K2, and AWL words were conducted with the SPSS software. For the following results, assumptions of equal variance and normality were met. Although the descriptive means results or K1 and K1 between the two corpora demonstrated means resembling each other, the means for AWL displayed a mismatch. As illustrated in Figure 1, the textbook corpus scored a higher mean in its use of K1 and K2 words (*MK1: 79.96%, SDK1: 1.93501; MK2: 6.64%, SDK2: .76213*) than the exam corpus (*MK1: 79.52%, SDK1: 1.65094; MK2: 6.15%, SDK2: .46871*). On the other hand, the exam corpus had a significantly higher coverage of academic words (*MAWL: 5.65%, SDAWL: 1.16101*) than the textbook corpus (*MAWL: 2.71%, SDAWL: 1.12163*). This finding was further proven with the following results. Independent t-tests results indicated that the corpora did have a drastically salient significance level for AWL. While K1 and K2 displayed insignificant statistical results (*K1: .556; K2; $p = .87$, $p > 0.5$*), AWL displayed a statistically significant result (*AWL: $p = .000 < 0.5$*). Descriptive statistics suggest that, on average, the exam corpus contained more low-frequency words than the textbook corpus as the textbook corpus demonstrated a higher usage of higher frequency words in mean (K1 and K2) and that the use of academic words was significantly low in the textbook corpus than in the exam corpus.

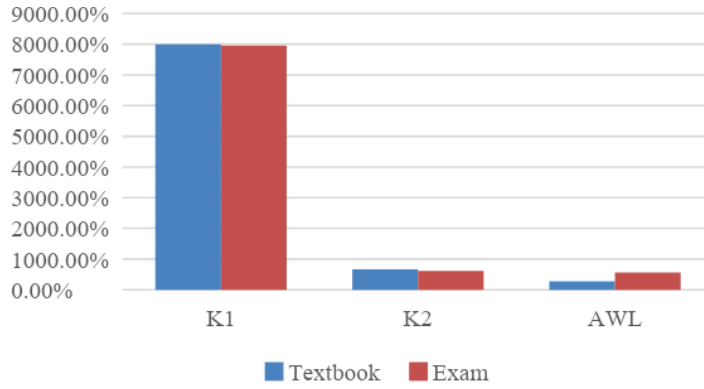


Figure 1. Lexical sophistication overlap

Unlike lexical sophistication findings, lexical diversity findings displayed greater differences in the mean between the two corpora in TTR, LogTTR, and MTLD. The assumptions of equal variance and normality were met. It is evident that, regardless of TTR type, the exam corpus always scored a higher mean value (*MTTR: .2335, SDTTR: .016959; MRootTTR: 18.096, SDRootTTR: 1.50964; MLogTTR: .8372, SDLogTTR: .010753; MMTLD: 59.8613, SDMTLD: 4.90247*) than the textbook corpus (*MTTR: .1212, SDTTR: .006937; MRootTTR: 17.1479, SDRootTTR: .793944; MLogTTR: .7864, SDLogTTR: .002871; MMTLD: 55.2500, SDMTLD: 3.97819*). These numbers indicate that the exam corpus was lexically more diverse than the textbook corpus on average. The mismatch of lexical diversity was proven by independent t-tests results ($p < .05$). These results were statistically significant except for Root TTR (*TTR: .000; RootTTR: .105; LogTTR: .000; MTLD: .042,* $p < .05$) and supported the claim that the exam corpus was lexically more diverse than the textbook corpus. Except Root TTR ($p = .105 > .05$), all other variables prove a notable variation for the corpora. Using Cohen's d (Cohen, 2013), the effect size of the differences between the two corpora regarding lexical diversity can be further explained. The effect sizes for the lexical diversity indices that were found are as follows; TTR: 8.6%, RootTTR: 0.78%, LogTTR: 6.45%, and MTLD: 1.03%. In other words, the previously mentioned percentage indicates the amplitude of the gap of lexical diversity between the two corpora. Figure 2 shows the lexical diversity overlap.

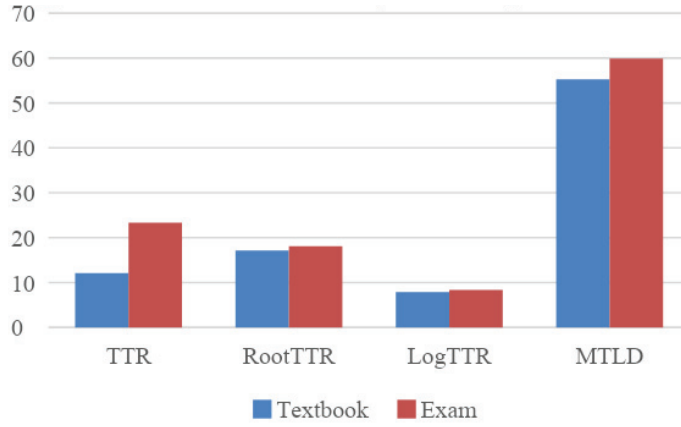


Figure 2. Lexical diversity overlap

Syntactic Complexity

Corresponding to the previous findings in the lexical section, syntactic complexity indices indicate significant differences regarding MLT, MLC, CT/T, and CN/T. The means of exams were higher (*MMLT: 15.47, SDMLT: 3.39884; MMLC: 9.80, SDMLC: 1.79345; MCT/T: .4631, SDCT/T: .12678; MCN/T: 1.84, SDCN/T: .61127*) than the textbooks means (*MMLT: 10.40, SDMLT: 2.67762; MMLC: 7.97, SDMLC: 1.26800; MCT/T: .2609, SDCT/T: .13291; MCN/T 1.01, SDCN/T: .43015*). (See Figure 3 for the differences). On the surface, it seems as if the exams were syntactically more complex than the textbook corpus. The results of the independent T-test further proved this point by displaying a significance level of ($p = .000 < 0.5$). Departing from our lexical findings, results for all four indices examined in this study performed a significance level ($p = .000 < 0.5$). These numbers suggest that the exam corpus was notably more complex than the textbook corpus regarding syntactic complexity. The implications of this finding are discussed in the next section.

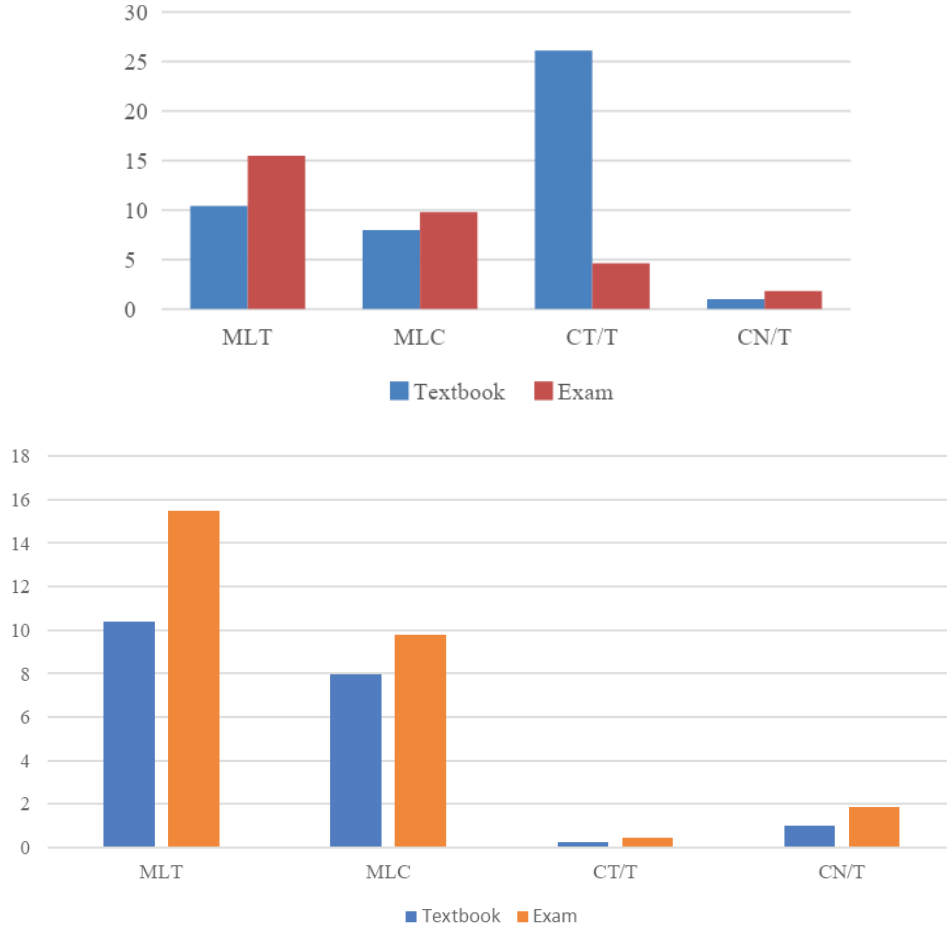


Figure 3. Syntactic complexity overlap

Discussion and Conclusion

The present research paper explored the lexical sophistication, lexical diversity, and syntactic complexity differences between the English high school textbooks and the English university entrance exams in Turkey.

Descriptive statistics suggest that lexical sophistication levels (for AWL) between the corpora demonstrate a considerable variation. Although the coverage of K1 and K2 were not significantly different between the two corpora, the coverage of the AWL was found to be significantly different. This indicates that the exam corpus contains more academic words than the textbook

corpus. Furthermore, because lexical sophistication level in AWL is lower for the textbook corpus, the learners who conduct English lessons with these textbooks are less likely to encounter low-frequency words AWL words than the AWL lexical items available in the exam corpus. This would indicate that these students would be less likely to encounter words that render them near-native-like. The exam corpus, on the other hand, proves to be lexically more sophisticated regarding AWL and contain less high-frequency AWL words in its inventory. Although K1 and K2 levels showed similar results, one should still note the slight variation between the corpora, especially when there needs to be a one-to-one correspondence between the exam and textbook materials. Frequency words also indicate that the decrease in the overlap correlates with the increase in the gap between the two corpora in terms of lexical alignment.

Results for the lexical diversity levels of the corpora tell a similar story. The differences in TTR, RootTTR, LogTTR, and MTLD among the corpora suggest that a statistically significant mismatch is present between the two corpora. More practical interpretation is averagely speaking, in every 100 words, the textbook corpus introduces ten new (different) words. This increases the lexical diversity gap between the two corpora, leading to poor input in the textbook corpus compared to the exam corpus. The statistical findings for lexical sophistication and diversity levels give the stakeholders (e.g., students, test and textbook-writers, English language teachers) a better insight and reinforce the recurring claim that the textbooks do not prepare students for the upcoming high stakes exams in terms of lexis.

The findings in lexical sophistication and diversity match with the findings of Yu's study (2018). Yu suggests that Turkish learners of English, in their academic writings, have the highest "coverage of the high-frequency words, namely the first and second 1,000 words" (Yu, 2018, p. 167). Furthermore, Yu's study, comparing Turkish speakers' written output to five other NNS groups, proves that Turkish learners of English demonstrate very poor lexical sophistication and diversity performances. These findings correspond to the current findings in this study, suggesting a cause-and-effect relationship of the materials used and tested. That is, if the materials used in classroom are more compelling regarding lexical sophistication and diversity, when they are tested in nationwide English exams, they are more likely to be acquired (see positive backwash effect, Heaton, 1989). Therefore, to improve the performance of Turkish learners of English, "vocabulary lists of academic, substitutional, and discipline-based words should be provided" (Yu, 2018, p. 168) in textbook materials.

Syntactic complexity findings are, perhaps, the most dramatic results in this study. Descriptive statistics results for syntactic complexity indices (MLC, MLT, CT/T and CN/T) always demonstrate a higher mean in the exam corpus. This means that on average, exam takers are likely to spend more time reading the

sentences (MLC). Due to higher means of MLT (and T-unit's nature which is "one main clause with all subordinate clauses attached to it" (Hunt 1965, p. 20) in the exam corpus, exam takers are more likely to be under a cognitive load to process the syntactic packaging compared to the textbook corpus. As with MLT, CT/T also significantly affects the exam takers processing times significantly as CT/Ts pack more complex T-units. Complementarily, higher means of CN/T indicates a heavier syntactic load for the exam takers, to decode the complex nominals. The difference between the two corpora was statistically significant for all indices. Namely, if students are to prepare for the high stakes exams using the government imposed books, then the chances of students' success (unless they have access to external educational materials and teachers who are aware of this mismatch, or this mismatch has been addressed by the exam and textbook preparation teams) is very low because of the mismatch between MLC, MLT, CT/T, and CN/T levels.

The pedagogical implications of this study are as follows: because there is a remarkable differentiation of lexical sophistication, lexical diversity and syntactic complexity levels, the students who have used these textbooks and taken these exams may have been forced to develop a more distorted idea of L2 (in this case, English). This distorted idea (also known as negative backwash effect) reinforces that languages can be split into smaller units and that no matter how hard they study for the English university entrance exam using government-based textbooks, they run the risk of not being able to succeed in the high-stakes English university exams. Another important point to explain is that students who use these textbooks are likely to struggle with exam fatigue due to heavy syntactic processing even from the very beginning of the exam. Moreover, this study can be beneficial for the major stakeholders of English language teaching in Turkey, namely, the textbook and exam-writers, the English language teachers, and the students. These stakeholders, with the findings at hand, can communicate and reconcile this apparent gap of lexical knowledge expected from students in the high stakes exams. The textbook and exam writers also need to work collaboratively to account for these to provide a more reliable exam experience for everyone, on equal grounds. The discussion of equal grounds can also be expanded to include the inequalities across socio-economically advantaged and disadvantaged students. Most students who come from a disadvantaged background may not have access to lexically and syntactically more compelling textbooks and may be more likely to fail in the university entrance exam while the advantaged students are ever so subtly favored and made to succeed as they already have access to more compelling language learning materials. This may not be the case for everyone in Turkey, but it might disclose an important—mostly overlooked—inequality that affects the lives of many young students who just wish to be successful but cannot figure out why they keep failing.

Although this study attempts to bridge the gap in the literature of Turkish corpus linguistics, it has several limitations. First, the study has relatively small corpora and only discovers the current situation of the corpora that are in use; Second, the study includes only four syntactic complexity indices out of fourteen. Future studies should consider these limitations and conduct a study that can utilize larger corpora and evaluate the overlap and mismatch of lexical sophistication/diversity and syntactic complexity alignment levels of the corpora.

References

Allen, H. W. (2008). Textbook materials and foreign language teaching: Perspectives from the classroom. *The NECTFL Review, 62*, 5–28.

Bardel, C., Gudmundson, A., & Lindqvist, C. (2012). Aspects of lexical sophistication in advanced learners' oral production: Vocabulary acquisition and use in L2 French and Italian. *Studies in Second Language Acquisition, 34*(2), 269–290. https://doi.org/10.1017/S0272263112000058

Beauchamp, D., & Constantinou, F. (2020). Using corpus linguistic tools to identify instances of low linguistic accessibility in tests. *Research Matters: A Cambridge Assessment publication, 29*, 10–16.

Biber, D. (1989). A typology of English texts. *Linguistics, 27*, 3–43.

Bulté, B., & Roothooft, H. (2020). Investigating the interrelationship between rated L2 proficiency and linguistic complexity in L2 speech, *System, 91*, 1–16.

Casanave, C. P. (1994). Language development in students' journals. *Journal of second language writing, 3*(3), 179–201.

Choi, I. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing,* 25 (1), 39–62.

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Academic press.

Crossley, S. A., & Salsbury, T. (2010). Using lexical indices to predict produced and not produced words in second language learners. *The Mental Lexicon, 5*(1), 115–147.

Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). What is lexical proficiency? Some answers from computational models of speech data. *TESOL Quarterly, 45*(1), 182–193.

Cumming, A., Kantor, R., Baba, K., Erdosy, U., Eouanzoui, K., & James, M. (2005). Differences in written discourse in independent and integrated prototype tasks for next generation TOEFL. *Assessing Writing, 10*(1), 5–43.

Daller, H., Van Hout, R., & Treffers-Daller, J. (2003). Lexical richness in the spontaneous speech of bilinguals. *Applied linguistics, 24*(2), 197–222.

Dean, A. C. (2017). Complex Dynamic Systems and Interlanguage Variability: Investigating Topic, Syntactic Complexity, and Accuracy in NS-NNS Written Interaction. *Working Papers in TESOL & Applied Linguistics, 17*(1), 56–97.

Du, W. (2019). Analysis on the development of lexical complexity in Chinese science students' English writing. *Noble International Journal of Social Sciences Research, 4*(7), 116–120.

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition, 26*(1), 59–84.

Ertmer, P. A., Bai, H., Dong, C., Khalil, M., Hee Park, S., & Wang, L. (2002). Online professional development: Building administrators' capacity for technology leadership. *Journal of Computing in Teacher Education, 19*(1), 5–11.

Fletcher, P. (1985). *A child's learning of English*. Blackwell.

Frase, L. T., Faletti, J., Ginther, A., & Grant, L. (1999). *Computer analysis of the TOEFL test of written English*. Educational Testing Service.

Gençoğlu, C. (2017, October). *Republic of Turkey Ministry of National Education*. COMCEC, Ankara, Turkey.

Grant, L., & Ginther, A. (2000). Using computer-tagged linguistic features to describe L2 writing differences. *Journal of Second Language Writing, 9*(2), 123–145.

Hatipoğlu, Ç. (2016). The impact of the university entrance exam on EFL education in Turkey: Pre-service English language teachers' perspective. *Procedia-Social and Behavioral Sciences,* 232, 136–144.

Hunt, K. W. (1965). Grammatical structures written at three grade levels. *NCTE Research Report No. 3,* 2–176.

Hyltenstam, K. (1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual & Multicultural Development, 9*(1–2), 67–84.

Kalyuga, S. (2006). Rapid assessment of learners' proficiency: A cognitive load approach. *Educational Psychology, 26*(6), 735–749. https://doi.org/10.1080/01443410500342674

Kim, J. Y. (2014). Predicting L2 Writing Proficiency Using Linguistic Complexity Measures: A Corpus-Based Study. *English Teaching*, *69*(4), 27–51.

Kirkgoz, Y. (2007). English language teaching in Turkey: Policy changes and their implementations. *RELC Journal, 38*(2), 216–228.

Kyle, K. (2016). *Measuring syntactic development in L2 writing: Fine grained indices of syntactic complexity and usage-based indices of syntactic sophistication* [Georgia State University]. http://scholarworks.gsu.edu/alesl_diss/35/

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

Kyle, K. (2019). Measuring Lexical Richness. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 454–475). Routledge.

Kwary, D., Artha, A., & Amalia, Y. (2018). Lexical word-class distributions in research articles of four subject areas. *Studies about Languages, 33*, 108–118.

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

Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied linguistics, 16*(3), 307–322.

Lu, X. (2011). A corpus-based evaluation of syntactic complexity measures as indices of college-level ESL writers' language development. *TESOL Quarterly, 45*(1), 36–62.

Lu, X., & Ai, H. (2015). Syntactic complexity in college-level English writing: Differences among writers with diverse L1 backgrounds. *Journal of Second Language Writing, 29*, 16–27.

Milli Eğitim Bakanlığı. (2018). Ortaöğretim İngilizce Dersi Öğretim Programı. Retrieved from: http://mufredat.meb.gov.tr/ProgramDetay.aspx?PID=342

Mirshojaee, S. B., & Sahragard, R. (2015). Reading comprehension passages of Iranian general English books and MA reading comprehension tests: A corpus analysis. *Journal of Modern Research in English Language Studies, 2*(2), 77–98.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior research methods, 42*(2), 381–392.

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*(1), 57–86.

Miller, D. P. (1981). The depth/breadth trade-off in hierarchical computer menus. In *Proceedings of the Human Factors Society 25th Annual Meeting* (pp. 296–300). HFES.

Nelson, N. W., & Van Meter, A. M. (2007). Measuring written language ability in narrative samples. *Reading & Writing Quarterly, 23*(3), 287–309.

Norrby, C., & Håkansson, G. (2007). The interaction of complexity and grammatical processability: The case of Swedish as a foreign language. *International Review of Applied Linguistics in Language Teaching, 45*(1), 45–68.

Nur, S., & Islam, M. (2018). The (Dis)Connection between Secondary English Education Assessment Policy and Practice: Insights from Bangladesh. *International Journal of English Language Education, 6*(1), 100–132.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied linguistics, 24*(4), 492–518.

Park, S.-Y. (2012). A corpus-based study of syntactic complexity measures as development indices of college-level L2 learners' proficiency in writing. *Korean Journal of Applied Linguistics, 28*(3), 139–160.

Read, J. (2000). *Assessing vocabulary*. Cambridge University Press.

Sheldon, L. E. (1998). Evaluating ELT textbooks and materials. *ELT Journal, 42*(4), 237–246.

Skalicky, S., Duran, N., & Crossley, S. A. (2020). Please, please, just tell me: The linguistic features of humorous deception. Retrieved from: osf.io/qdjmn

Tai, S., & Chen, H.-J. (2015). Are teachers test-oriented? A comparative corpus-based analysis of the English entrance exam and junior high school English textbooks. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds.), *Critical CALL – Proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 518–522). Research-publishing.net. http://dx.doi.org/10.14705/rpnet.2015.000386

Thomas, D. (2005). Type-Token Ratios in one teacher's classroom talk: An investigation of lexical complexity. University of Birmingham.

Torruella, J., & Capsada, R. (2013). Lexical statistics and typological structures: A measure of lexical richness. *Procedia-Social and Behavioral Sciences*, *95*, 447–454.

Underwood, P. (2010). A comparative analysis of MEXT English reading textbooks and Japan's National Center Test. *RELC Journal, 41*(2), 165–182.

Vermeer, A. (2004). Vocabulary size in Dutch L1 and L2 children. In P. Bogaards & B. Laufer (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 173–189). John Benjamins Publishing Company.

Wang, M. D. (1970). The role of syntactic complexity as a determiner of comprehensibility. *Journal of Verbal Learning and Verbal Behavior, 9*(4), 398–404.

Wang, S., & Slater, T. (2016). Syntactic complexity of EFL Chinese students' writing. *English Language and Literature Studies, 6*(1), 81–86.

Wolfe-Quintero, K., Inagaki, S., & Kim. H. Y. (1998). *Second language development in writing: Measures of fluency. Accuracy and complexity*. University of Hawaii Press.

Yu, X. (2018). *Analyses and comparisons of three lexical features in native and nonnative academic English writing* [University of Central Florida]. https://stars.library.ucf.edu/etd/6061

Tan Arda Gedik, Yağmur Su Kolsal

Eine korpusbasierte Analyse englischer Lehrbücher für die Oberschule und englischer Hochschulaufnahmeprüfungen in der Türkei

Zusammenfassung

Die vorliegende Studie untersucht die Diskrepanz zwischen dem Inhalt von englischen Lehrbüchern, die man in den Oberschulen (9. bis 12. Klasse) verwendet, und Englischkenntnissen, die während Aufnahmeprüfungen an türkischen Universitäten (2010–2019) geprüft werden. Unter Verwendung von korpuslinguistischen Werkzeugen wie AntWordProfiler, TAALED bzw. L2 Syntactic Complexity Analyzer (L2SCA) werden anhand des Untersuchungsmaterials die lexikalische Vielfalt und syntaktische Komplexität analysiert. Aus dem Vergleich der offiziellen Lehrbücher und zusätzlichen Materialien des Ministeriums für Nationale Bildung mit den offiziellen Hochschulaufnahmeprüfungen lässt sich schließen, dass: (i) es treten Unterschiede im lexikalischen Niveau zwischen den beiden Korpora auf – das lexikalische Niveau des Prüfungskorpus war höher als das des Lehrbuchkorpus, (ii) zwischen den beiden Korpora besteht ein statistisch signifikanter Unterschied in Bezug auf die lexikalische Vielfalt – das Prüfungskorpus hat ein wesentlich höheres Niveau der lexikalischen Vielfalt als das Lehrbuchkorpus, (iii) es gibt statistisch signifikante Unterschiede zwischen den beiden Korpora hinsichtlich der syntaktischen Komplexität – das Niveau der syntaktischen Komplexität im Prüfungskorpus war höher als das im Lehrbuchkorpus. Die angeführten Schlussfolgerungen deuten darauf hin, dass türkische Oberschüler, die aus offiziellen Lehrbüchern Englisch lernen, bei landesweiten Prüfungen mit dem seltener gebrauchten und anspruchsvolleren Wortschatz auf höherem Niveau der syntaktischen Komplexität umgehen müssen. Dies wiederum führt zu einem negativen Backwash-Effekt, der ihre Einstellung zur Fremdsprache verzerrt und weitere Bedenken hinsichtlich Abweichungen zwischen den offiziellen Sprachlehrmaterialien und landesweiten Prüfungen aufkommen lässt.

Schlüsselwörter: Korpuslinguistik, lexikalische Vielfalt, syntaktische Komplexität

Reviews

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 179–182
https://doi.org/10.31261/TAPSLA.10425

David Singleton and Larissa Aronin (Eds.), *Twelve Lectures on Multilingualism*. Bristol: Multilingual Matters, 2019, ISBN 978-1-78892-205-0

This timely and most welcome handbook includes a state-of-the-art overview of multilingualism research from a variety of angles and perspectives. It is organized around four main parts dealing with (i) multilingualism in society and education, (ii) aspects of individual multilingualism, (iii) the psycholinguistics and neurolinguistics of multilingualism, and (iv) forms of multilingualism in the past and present.

Part 1 of the book on educational and societal perspectives includes four lectures. The first two chapters provide a comprehensive description of what multilingualism entails. Aronin presents defining features of a very complex phenomenon. She refers to historical and current multilingualism. As Aronin puts it, language use has changed dramatically as some languages are more dominant than others. This idea leads to the notion of dominant language constellations (DLC) also presented in Lecture 2. Aronin and Gabryś-Barker both refer to DLC as a new stable pattern of organization that shows the languages employed in a given speech community and how these are used. This dynamic description accounts for reality of language use both at an individual and societal level. As argued by Gabryś-Barker, considering DLC as an approach may change actual perspectives in research. The way participants and communities are described in bilingualism and multilingualism studies would be more realistic and could thus help explain patterns of multilingual language acquisition and use. Interestingly, Gabryś-Barker raises another important contribution of multilingualism research to applied linguistics in general, namely, that of qualitative methodology. Currently, many scholars argue for the need to use mixed-methods and it often involves data triangulation and the inclu-

sion of qualitative data collection procedures. Introspection techniques and narratives may show patterns of acquisition that would be invisible in plain statistical data. Research by Gabryś-Barker (2013, 2017), which also examines the use of metaphors, is an excellent example of this fact. In addition to multilingualism methodology, this second chapter deals with key concepts that help the reader understand the scope of multilingualism. One of these concepts is Herdina and Jessner's Dynamic Model of Multilingualism (2002) which is also mentioned in Lecture 3. Hufeisen and Jessner tackle the psycholinguistics of multiple language learning by presenting research to date and challenges for future studies. These authors argue for the need to create new guidelines in teacher training and language instruction. The final chapter of this first part of the volume also deals with language teaching and it considers norms for language education. Cenoz and Gorter refer to minority languages in educational policies as one type of multilingual education (see Gorter, Zenotz, & Cenoz, 2014). Content and Language Integrated Learning (CLIL) is described when tackling the educational framework of majority language speakers. As raised in their work (Cenoz, 2015), CLIL may be seen as an effective way to learn a third language but there is still a need to include all learners' languages.

Part 2 of the volume tackles individual multilingualism throughout four lectures. Edwards in Lecture 5 raises two main points. On the one hand, the author states that languages do not evolve identically in multilinguals. In fact, factors like motivation, attitudes or levels of fluency may influence their development. On the other hand, Edwards challenges the direct connection between multilingualism and intelligence. The repertoires that multilinguals possess and their interaction are the focus of Chapter 6. De Angelis examines the factors that boost cross-linguistic influence (CLI) in language learning. In so doing, the author points to recency of use, second language status and typological proximity. One of the most powerful factors in language acquisition is motivation, and research on this topic has largely focused on L2 as argued by Ushida in Lecture 7. This author refers to notions like identity that may best embrace the complex and individual experiences of multilingual learners. Another factor that has raised much interest is that of age. Lecture 8 is devoted to age and multilingualism. According to Muñoz and Singleton, who have largely investigated this topic (Muñoz & Singleton, 2011), significant exposure may be responsible for better acquisition rates in younger learners. They also refer to the great influence of attitudes and motivation, as well as the key role of quality in multilingual schools. Interestingly, these authors conclude that learners of all ages may become successful multilingual individuals.

Parts 3 and 4 of the volume include two lectures each. Festman deals with the psycholinguistics of multilingualism in Lecture 9 of the present volume.

She describes a number of concepts used in psycholinguistics together with some assumptions and empirical evidence. The author highlights some open research areas that refer to the representation of L1 or L3 and the extent to which these may differ. In line with Ushida's recommendation (see Lecture 7), Festman argues for the use of mixed-methods in multilingualism research. Kadyamusuma, Higby, and Obler tackle the neurolinguistics of multilingualism and they present evidence against the idea that languages are stored in different compartments in the multilingual brain. These authors also refer to interesting findings on the cortical organization of language representation and its relationship with dementia.

Forms of past and present multilingualism are included in the last part of the volume. Lecture 11 deals with historical multilingualism. Braunmüller comments on the traces of covert multilingualism and he raises the possibility of comparing those who had Latin as L2 during the Roman Empire and current L2 learners. The presence and spread of English as a Lingua Franca and the fact that its use is often restricted to academia or the media in some sociolinguistics contexts may be one of the causes of current receptive multilingualism. This is the last topic of the book and it is examined by Ten Thije in Lecture 12.

As the editors state in the introduction of the book, we live in a multilingual world where citizens from a variety of cultures and identities interact, learn languages and are virtually connected. Research no longer compares native and non-native speakers of a given language but it now focuses on intercultural speakers with dominant linguistic repertoires that vary across time. The complexity of multilingualism that was raised by Hufeisen, Cenoz, and Jessner (2001) almost two decades ago is now also the norm in L3 research. We now know more about the peculiarities of multilingual language learners and speakers from psycholinguistic, educational or sociolinguistic perspectives and this volume is a wonderful source to introduce graduate students, scholars or interested parties to that knowledge.

The strength of the volume also lies in the outstanding scholars who have contributed to this book. While all authors are widely recognized scholars, I may specifically refer to the three founders of the International Association of Multilingualism, Jasone Cenoz, Ulrike Jessner, and Britta Hufeisen as well as co-founding members like Larisa Aronin, David Singleton, Danuta Gabryś-Barker whose work has inspired that of other scholars with an interest in the teaching, acquisition and use of languages beyond the second or third one. *Twelve Lectures on Multilingualism* is a must-read in any graduate or postgraduate course on languages in contact, multilingual education or third language acquisition. This book is also a great excuse to review what has already been done and the challenges that remain in this exciting, innovative,

and interdisciplinary area of study that has captivated many of us, that of multilingualism.

https://orcid.org/0000-0002-2130-2409

Pilar Safont
Universitat Jaume I (Castelló – Spain)

References

Gabryś-Barker, D. (2013). Face to face with one's thoughts: On thinking multilingually. In M. Pawlak and L. Aronin (Eds.), *Essential topics in applied lingusitics and multilingualism: Studies in honour of David Singleton* (pp. 185–204). Springer.

Gabryś-Barker, D. (2017). New approaches to multilingualism research. Focus on metaphors and similes. In D. Gabryś-Barker, D. Gałajda, A. Wojtaszek, & P. Zakrajewski (Eds.), *Multicultiralism, multilingualism and the self* (pp. 77–95). Springer.

Cenoz, J. Hufeisen, B., & Jessner, U. (2001). *Looking beyond second language acquisition: Studies in tri- and multilingualism*. Stauffenburg.

Cenoz, J. (2015). Content-based instruction and content and language integrated learning. The same or different? *Language, Culture and Curriculum*, 28, 1–24.

Gorter, D. Zenotz, & Cenoz, J. (2014). *Minority languages and multilingual education*. Springer.

Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, 44, 1–35.

Theory and Practice of Second Language Acquisition
vol. 8 (1), 2022, pp. 183–187
https://doi.org/10.31261/TAPSLA.12089

Mariusz Kruk and Mark Peterson (Eds.), *New Technological Applications for Foreign and Second Language Learning and Teaching* Hershey, PA: IGI Global, 2020, ISBN 9781799825937, 388 pages

New Technological Applications for Foreign and Second Language Learning and Teaching published in 2020 is part of the IGI Global book series Advances in Educational Technologies and Instructional Design (AETID), which assembles state-of-the-art research in (language) teaching and learning with technologies. The present peer-reviewed volume, edited by Mariusz Kruk and Mark Peterson, is dedicated to exploring diverse current practices, research, and implications in technology-enhanced language learning and teaching. In their preface, the editors touch on the role of technology in today's society and introduce the readers to the aims and contents of the volume.

The book is divided into two sections: a theoretical and a practical one. The former, entitled Theoretical Considerations comprises the first five chapters that look into computer assisted language learning (CALL) and the research of selected subfields within it. The remaining ten chapters are subsumed under the heading Practical Applications, covering empirical research into digital tool and technology-enhanced pedagogy integration into the language classroom. Each chapter is rounded off by a list of recommended literature and a list of key terminologies (and brief definitions) used in the chapter.

John Blake commences the theoretical considerations with a contribution entitled "Intelligent CALL: Using Pattern Matching to Learn English" (Chapter One) which discusses the application of natural language processing for language learning. The author first provides a history of CALL and intelligent CALL, and then discusses patterns and their use for language learning through

a discovery learning, data-driven learning (DDL) approach. Throughout, contextual factors surrounding EFL learners whose L1 is Japanese are addressed and two tools developed specifically for Japanese university EFL learners and pilot studies of their effectiveness are presented.

In "Pedagogical Considerations for Successful Implementation of Virtual Reality in the Language Classroom" (Chapter Two), Ryan Lege, Euan Bonner, Erin Frazier, and Luann Pascucci introduce the VR Application Analysis Framework. This framework was designed to help language teachers analyze commercial virtual reality applications in terms of their feasibility for language learning purposes. In addition to the components of the framework, its application in analyzing a selection of VR applications is included in the chapter.

In Chapter Three—"The Case for Qualitative Research Into Language Learning in Virtual Worlds"—Luisa Panichi presents an extensive account of research into virtual worlds in CALL with the aim of "opening up our thinking as a community to wider scrutiny and debate in the interest of the validity of our research on the one hand, and future research decisions and directions on the other" (p. 48). She argues for the continued validity of qualitative approaches within the field which is characterized by the researcher-practitioner configuration and the fluidity of the medium studied, which constantly produces unchartered territory for exploration.

Mark Peterson, Jeremy White, Maryam Sadat Mirzaei, and Qiao Wang devote Chapter Four to "A Review of Research on the Application of Digital Games in Foreign Language Education." The authors first consider features of digital games which are attributed educational merit form a general educational, a cognitive and social linguistic perspective. These features are then supplied with evidence through the synthesis of the 26 research studies reviewed. In addition to a valuable overview, the chapter uncovers numerous avenues for future investigation.

In "Gamification for Technology-Enhanced Language Teaching and Learning" (Chapter Five), Joan-Tomàs Pujolà and Christine Appel draw the readers' attention to different terminology which is often used interchangeably despite significant conceptual differences (e.g., game-based learning vs. gamification) and explain how these approaches differ. An overview of the dearth of research into gamification of language learning experiences is followed by useful design recommendations, tools and resources for gamifying language learning through technologies. The chapter closes by suggesting that future research in this domain should draw on replication studies and design-based research to study innovative design and attempt generalizations of findings.

Section Two, which assembles practical applications of technologies in language learning and teaching, begins with Farhana Ahmed's contribution on "Investigating Learner Autonomy and 21st Century Skills in Blended Tech-Enhanced Language Learning" (Chapter Six). In the context of a blended

learning EAP course which employs products of the Google Suite for communication and collaboration, learner autonomy is explored using Benson's (2011) learner autonomy framework. A discussion of the implications of the findings for practitioners including "the need for educating students about the specific technologies that are being used by the teacher" (p. 129) conclude the chapter.

Chapter Seven, authored by Alberto Andujar and Fidel Çakmak, uncovers the potentials of "Foreign Language Learning Through Instagram: A Flipped Learning Approach." This mixed-methods case study in secondary education examines how leaners perceived the use of teacher recorded instructional videos uploaded to Instagram for instruction outside the classroom prior to in-class practice. Although the platform was chosen due to its responsive design (e.g., layout adapts to screen size) and its popularity among adolescents, the log data revealed that fewer than half of the students accessed the videos.

Chapter Eight, "Using Digital Storytelling to Handle Second Language Writing Anxiety and Attitudes: A Longitudinal Experiment," by Seyit Ahmet Çapan, describes a quasi-experimental design to measure and compare language mistakes, writing anxiety and attitudes towards writing of two groups of university students with elementary L2 skills. While the one group produced stories using pen and paper, the other one created digital stories. Findings suggest that the latter perceived writing in a more positive light.

Juan Francisco Coll-García reports "Students' Perceptions Toward an International Telecollaboration Project Through an Engineering-Themed Online Simulation in a Language-Learning Setting" in Chapter Nine. The findings highlight the pedagogical need for learners to understand the rationale behind such projects, the different aims pursued by all involved parties and how they connect to the curriculum.

Chapter Ten by Hiroshi Hasegawa, Julian Chen, and Teagan Collopy also investigates learner perceptions: "First-Year Japanese Learners' Perceptions of Computerised vs. Face-to-Face Oral Testing: Challenges and Implications." The computerized oral testing mode is described as learners voice-recording their answers to questions posed in the form of pre-recorded teacher videos achieved through a computer program. While this chapter indicates that it focuses on a first analysis of a larger data set and on student experience, the reader is left curious as to how the computer-based oral tests are assessed and how assessment affects the time reduction reported as one of the main benefits of this examination mode for teachers.

Ferit Kiliçkaya focuses on "Using a Chatbot, Replika, to Practice Writing Through Conversations in L2 English: A Case Study" in Chapter Eleven. The qualitative case study explores learner perceptions of practicing writing in the L2 through instant messaging with a chatbot which through this interaction increasingly replicates the human participant. The discussion concludes by

stressing the need to "consider course objectives and pedagogical uses" (p. 231) of technology integration into the language classroom.

In Chapter Twelve, "Considerations for Future Technology Development Based on EFL Teachers' Integration of Technology," Tim Kochem, Ananda Astrini Muhammad, Yasin Karatay, Haeyun Jin, and Volker Hegelheimer present the Global Online Course which trains English language teachers in employing tools and digital resources for teaching the four skills, grammar and vocabulary. On the basis of naturalistic inquiry, the team of researchers explore language teachers' interaction with the course content and their set of ideas of technology integration. This chapter includes preliminary findings and discusses how these and the course can contribute to spreading knowledge of technology-enhanced teaching, providing an overview of useful tools, and in contributing to what they term "usable technology" development (p. 254).

Lina Lee addresses intercultural and interpersonal learning through virtual language exchanges in Chapter Thirteen ("Promoting Interpersonal and Intercultural Communication With Flipgrid: Design, Implementation, and Outcomes"). Using the video response platform Flipgrid, an L2 Spanish class consisting of L1 English-speakers interacted with an L2 English class whose L1 is Spanish through short videos they recorded of themselves in the respective L2. The discussion includes insights into tasks, pedagogical considerations and shortcomings of (employing) the video platform.

In Chapter Fourteen, Hsien-Chin Liou and Tzu-Wei Yang focus on "Data-Driven Learning at the English Drafting Stage." This mixed methods, action research study examines Asian EFL university students' use of monolingual (COCA) as well as bilingual (TANGO and Totalrecall) corpora in data-driven learning (DDL) while composing the draft version of an essay; screencasts of students' activity on the computer served as one source of data. The findings suggest that the students primarily employed corpora to define patterns but can nonetheless be described as "emergent pattern hunters" (p. 293).

DDL also constitutes the research interest of Chapter Fifteen, the final chapter, by Nina Vyatkina. In "Corpus-Informed Pedagogy in a Language Course: Design, Implementation and Evaluation" the researcher proposes a pedagogical framework for incorporating DDL into L2 teaching practice. Following a design-based research approach, the development, implementation and evaluation of a DDL university course for German as an L2 are described using the ADDIE model as adapted by Colpaert (2006) as a descriptive framework.

This volume provides detailed insights in the scope and breadth of practice in and research into the intersection between language learning and technology. The definitions, keywords, and suggestions of further reading make the volume accessible to novices and non-specialists, yet the scope of topics and applications make it an interesting resource for researchers and practitioners

as well. The latter will appreciate the large number of specific tools discussed, investigated and evaluated.

Throughout the volume, the important role of language learning pedagogy and instructional design are either stressed from the outset or emergent through the findings of practical applications. Areas which seem to have been largely neglected are data protection, privacy and further context specific ethical considerations of practical applications in general and specific technologies discussed in particular (e.g., Google Suite, Replika, Flipgrid, etc.). Especially when commercial applications or applications which collect (sensitive/personal) data of learners are involved, implications and possible consequences should be discussed and considered critically.

https://orcid.org/0000-0001-9140-7110

Caterina Hauser
University of Graz, Austria

References

Benson, P. (2011). *Teaching and researching: Autonomy in language learning*. Pearson Education Limited.

Colpaert, J. (2006). Pedagogy-driven design for online language teaching and learning. *CALICO Journal*, *23*(3), 477–497. https://doi.org/10.1558/cj.v23i3.477-497

STYLE GUIDE FOR THE AUTHORS

Please note that we are changing from APA 6th edition to newer 7th edition. Authors are requested to submit manuscripts formatted in APA style (*American Psychological Association*, 7th ed.).

All manuscripts must include an abstract in English (maximum of 250 words). After the abstract please provide keywords.

Main text: 12 Times New Roman
Long citations (more than 40 words): 10 Times New Roman, indent by 1 tab either side, one empty line above and below, no quotation marks.

1.5 spacing

APA headings

Level	Format
1	**Centered, Boldface, Uppercase and Lowercase Heading**
2	**Left-aligned, Boldface, Uppercase and Lowercase Heading**
3	**Indented, boldface, lowercase heading with a period.** Begin body text after the period.
4	***Indented, boldface, italicized, lowercase heading with a period.*** Begin body text after the period.
5	*Indented, italicized, lowercase heading with a period.* Begin body text after the period.

In-text citations (examples):

Author's name and date in brackets:
The experience of critical incidents and effective reflection upon them allows teachers to control their classroom actions more consciously and create critical events (CE's), which were described earlier as intended, planned and controlled (Woods, 1993).
Woods (1993) believes that critical events are structured and occur in well-defined staged of conceptualization . . .

Two authors:
(Ballantyne & Packer, 1995)
As Ballantyne and Packer (1995) demonstrate …

Three authors:
(Barker, Callahan, & Ferreira, 2009)

Subsequent use:
(Barker et al., 2009)

Six authors or more:
Lorenz et al. (1998) argued…
(Lorenz et al., 1998)

Authors whose last names are the same:
(D. Francis, 1985; H. Francis, 2004)

Online sources (unpaginated), provide paragraph or section title instead:
(Peterson & Clark, 1978, para. 4)
(Moss, Springer, & Dehr, 2008, Discussion section, para. 1)

No author, provide shortened title:
("Primary Teachers Talking," 2007)
(*Reflective Practice*, 2005, pp. 12–25)

Secondary citations:
Smith (as cited in Maxx & Meyer, 2000) noted that "there is . . . ."

Citation within citation:
As it has been noted that "there is no relevance . . . (Smith, 2005)" (Maxx & Meyer, 2000, p. 129).

& vs. and:
As Smithson and Stones (1999) demonstrated. . .
. . . as has been shown (Smithson & Stones, 1999) . . .

References

Selected examples (for more consult APA manual 7th ed.):

Book, one author:
Goldberg, A. (2006). *Constructions at work*. Oxford University Press.

Book, two authors and more:
Jarvis, S., & Pavlenko, A. (2008). *Crosslinguistic influence in language cognition.* Routledge.

Translated book:
Freud, S. (1960). *Jokes and their relation to the unconscious*. (J. Strachey, Trans.). Routledge & K. Paul. (Original work published 1905).

Edited book:
Flowerdew, J., Brock, M., & Hsia, S. (Eds.). (1992). *Second language teacher education.* City Polytechnic of Hong Kong.

Chapter in an edited book:
Goldberg, A., & Casenhiser, D. (2008). Construction learning and second language acquisition. In P. Robinson & N. C. Ellis (Eds.), *Handbook of cognitive linguistics and second language* acquisition (pp. 197–215). Routledge.

Article in a journal:
Hammarberg, B. (2010). The languages of the multilingual. Some conceptual and terminological issues. *International Review of Applied Linguistics in Language Teaching, 48*, 91–104.

Article online:
Tully, K., & Bolshakov, V. Y. (2010). Emotional enhancement of memory: How norepinephrine enables synaptic plasticity. *Molecular Brain,* 13 May. Retrieved from: http://www.molecularbrain.com/content/

Bakker, A. B., Hakanen, J. J., Demerouti, E., & Xanthopoulou, D. (2007). Job resources boost work engagement, particularly when job demands are high. *Journal of Educational Psychology*, *99*(2), 274–284. https://doi.org/10.1037/0022-0663.99.2.274

Magazines online:

Miller, G. (2014, September 4). Cinematic cuts exploit how your brain edits what you see. *Wired.* Retrieved from: http://wired.com/

Smith, A. (2007, June 12). Dying languages. *The Western Star*. Retrieved from: http://www.thewesternstar.com/

Blog:

Palmer, P. (2001). Now I become myself. *Yes Magazine,* blog post, 31 May. Retrieved from: http://www.yesmagazine.org/issues/working-for-life/now-i-become-myself

E-books:

Bolande, V. U. (1981). *On the psychology of humor.* Retrieved from: http://www.uflib.ufl.edu /ufdc/UFDC.aspx?n=palmm&c=psa1&m=hd2J&i=45367

Conference proceedings:

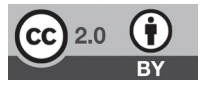
Souleles, N., & Pillar, C. (Eds.). (2014). Proceedings from the *First International Conference on the Use of iPads in Higher Education*. Paphos: Cyprus University of Technology.

Doctoral dissertation:

Churchwell, J. (2005). Becoming an academic: Factors that influence a graduate student's identity commitment (Doctoral dissertation). University of Michigan, Ann Arbor, MI.

Reachel, L. H. (2001). *Native languages and toponyms: Origins, meaning, and use* (Doctoral dissertation). Available from ProQuest dissertation and theses database. (Document ID 1964749161).

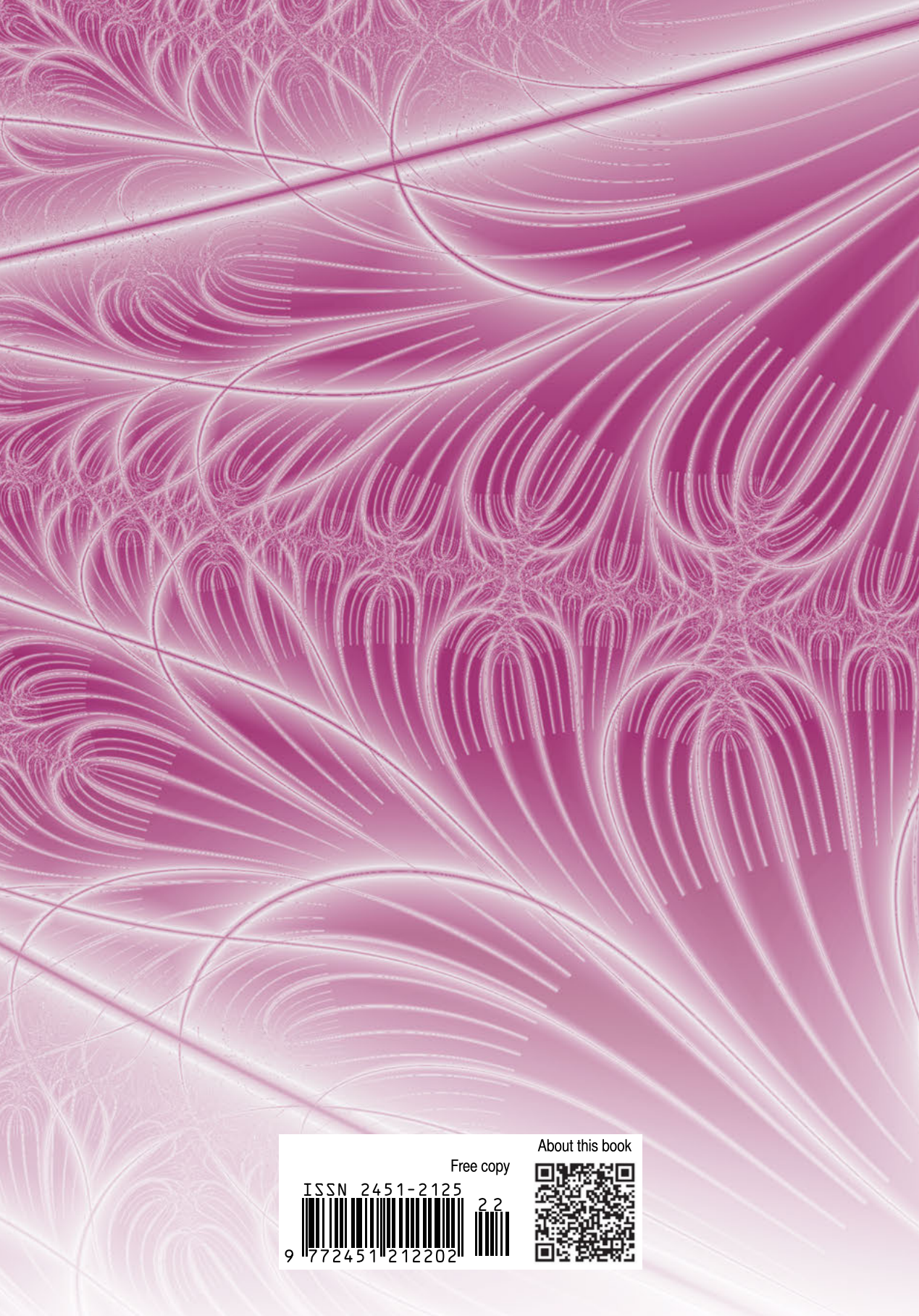
Cover photo: “big_blue” by Max Iter (Retrieved from www.flickr.com)



Copy-editing and proofreading: Gabriela Marszołek
Typesetting: Marek Zagniński
Cover preparation for printing: Paulina Dubiel

Electronic version is the original one.
The journal was previously published in printed form with the
ISSN 2450-5455
The journal is distributed free of charge
ISSN 2451-2125

Published by
Wydawnictwo Uniwersytetu Śląskiego
ul. Bankowa 12B, 40-007 Katowice
www.wydawnictwo.us.edu.pl
e-mail: wydawnictwo@us.edu.pl
First impression. Printed sheets: 12.0. Publishing sheets: 14.0.


Free copy
About this book
ISSN 2451-2125
9 772451 212202 22