

# Theory and Practice of Second Language Acquisition

Vol. 10 (1), 2024



UNIWERSYTET ŚLĄSKI  
WYDAWNICTWO

**Theory and Practice  
of  
Second Language Acquisition**

Vol. 10 (1), 2024

Editors-in-Chief

**Jolanta Latkowska**

University of Silesia in Katowice

**Adam Wojtaszek**

University of Silesia in Katowice

Language Editor

**David Schauffler**

University of Silesia in Katowice

Editorial Board

Janusz Arabski (University of Silesia in Katowice/Vistula University, Warsaw)  
Larissa Aronin (Oranim College of Higher Education/Trinity College, Dublin)  
Jasone Cenoz Iraqui (University of the Basque Country, Donostia – San Sebastian)  
Halina Chodkiewicz (Maria Curie-Skłodowska University, Lublin)  
Gessica de Angelis (Trinity College, Dublin)  
Anna Ewert (Adam Mickiewicz University, Poznań)  
Tammy Gregersen (University of Northern Iowa, Cedar Falls)  
Ulrike Jessner Schmid (University of Innsbruck, Innsbruck)  
Hanna Komorowska (University of Social Sciences and Humanities/University of Warsaw)  
Peter MacIntyre (Cape Breton University, Sydney)  
Anna Niżegorodcew (Jagiellonian University, Cracow)  
Aneta Pavlenko (Temple University, Philadelphia)  
Miroslaw Pawlak (Adam Mickiewicz University, Kalisz/State School of Higher Professional Education, Konin)  
Ewa Piechurska-Kuciel (University of Opole, Opole)  
Andrzej Porzuczek (University of Silesia in Katowice)  
David Singleton (Trinity College, Dublin/University of Pannonia, Veszprem)  
Eva Vetter (University of Vienna, Vienna)  
Ewa Waniek-Klimczak (University of Łódź, Łódź)  
Maria Wysocka (University of Silesia in Katowice)

This publication is indexed in the following databases:

**Arianta, BASE (Bielefeld Academic Search Engine), BAZHUM, CEEOL, CEON, DOAJ, 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.**



Attribution-ShareAlike 4.0 International (CC BY-SA 4.0)

## Contents

Preface (*Jolanta Latkowska, Adam Wojtaszek*). . . . .

### Articles

**Soheil Behdarvandirad**

A Systematic Review of Second Language Acquisition from the Perspective of Complex Dynamic System Theory . . . . .

**Robert Oliwa**

Contextualising the Hyflex Model of Instruction for Language Classes . . . . .

**Sabina Nowak, Aleksandra Rążewska**

Challenging but Rewarding Aspects of Telecollaboration: The Case of Virtual Israeli-Polish (VIP) Project . . . . .

**Emerson Case, Agnieszka Kaczmarek, Sebastian Zatylny**

Reading, Discussing, and Sharing: Creation of a Vibrant Transnational Online Learning Community through the International Book Club . . . . .

**Katarzyna Budzińska**

Foreign Language Classroom Anxiety in Emergency Remote Instruction . . . . .

**Nettie N. Ndou, Segun Emmanuel Adewoye**

Challenges of Learning in Second Language among South African School Learners with Developmental Language Disorder . . . . .

**Zofia Chlopek, Jacek Pradela**

Parents' Perceptions of Children's Bi-/Multilingualism . . . . .

**Fatih Ünal Bozdağ**

Probabilistic Analysis of English Dative Constructions in Academic Writings of English EFL Learners . . . . .

**Marta Sánchez-Saus Laserna**

Lexical Availability and Foreign Language Teaching: Main Contributions of a Growing Field . . . . .

**Zuzanna Wnuk**

Graphemic Awareness Development of Polish Learners of Chinese as a Foreign Language . . . . .

**Hiromu Okamura, Tomohiko Shirahata**

Factors Causing Overpassivisation of Unaccusative Verbs by Japanese Learners of English . . . . .

**Thiri Soe**

Investigating Critical Thinking in ELT Textbooks: A Systematic Literature Review of Textbook Evaluation Studies . . . . .

Style Guide for the Authors. . . . .



## A Good-bye Note

Resigning from my current position as editor-in-chief of our journal due to my retirement, I would like to thank Prof. Adam Wojtaszek, my co-editor-in-chief, and all the editorial team for the smooth and generous cooperation they have provided in creating and running this journal from its very beginnings in 2015 until now. It has been a wonderful and inspired time, the whole journey from scratch to a fairly well-established academic journal, enjoying a good reputation among scholars in the field. I would also like to extend my thanks to all of our referees whose thorough and rigorous assessment of submitted texts allowed us to develop and maintain high standards in the papers we published. And, last but not least, I would like to thank all the contributors to issues of TAPSLA in the period 2015–2023. Without you, the journal would not have become what it is now.

A lot of hard work has gone into producing the journal over the last eight years and hopefully will continue thanks to the publisher, the University of Silesia Press, and its well-qualified and dedicated team of administrators, coordinators, and copyeditors, from amongst whom I would like to give special mention to Dr Gabriela Marszolek-Kalaga, whose work with us has always been of the very highest standard.

This is the last issue of TAPSLA with me as one of its editors-in-chief. I do hope that TAPSLA will go on developing and continuing to raise its standards as a lively forum for the exchange of ideas in the field of second language acquisition, broadly understood.

 <https://orcid.org/0000-0003-0626-0703>

Thank you all very much.  
*Danuta Gabryś-Barker*  
(outgoing Editor-in-Chief)





## Preface

The Preface to this new issue of *Theory and Practice of Second Language Acquisition* offers—as it usually does—a short introduction for new readers to familiarize them with the origins and development of the journal. Additionally, it provides an introductory comment on the contents of the present issue (10/1).

*Theory and Practice of Second Language Acquisition* was founded as a journal offering a space for discussion among both Polish and foreign scholars working in the areas of generally understood second/foreign/multiple language acquisition and learning. It has clearly fulfilled its mission as a journal on the rise. The present volume marks the tenth year of its presence in the scholarly world, signalling the first significant anniversary of a journal which has already gained considerable popularity and esteem. We receive more and more qualified submissions from Polish and foreign researchers representing recent research in language acquisition studies. Since its founding, every successive issue of the journal has welcomed contributions from renowned researchers, including Peter MacIntyre, David Singleton, Larissa Aronin, Sarah Mercer, Tammy Gregersen, and Jean-Marc Dewaele, among others. Also, the fast-growing number of Open Access downloads testifies to the journal's increasing popularity, as does the backlog of articles already accepted and awaiting their turn to be assigned to individual issues for publication in print. This is reflected in the number of research papers published in each issue: for the first time in the journal's history, there are twelve, double the number to be found in the first volumes from 2015 and 2016. It has become our usual practice for each article to be available as Online First publication before it is assigned to a specific issue, thanks to which many authors can share their research with the academic community as soon as the copyediting phase is complete. It is the journal's ambition to showcase new trends and hitherto unknown venues for research in SLA, focusing both on theoretical discussions and on the practical solutions to problems that are based upon them. Our purpose is not only to publish and share with our read-

ers contributions from well-known and respected scholars, but also to promote young researchers from all over the world, who often present fresh and innovative ideas or open up new perspectives on issues already under discussion. In other words, the journal serves as a venue for the exchange of ideas between well-established academics and those who have been 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 both of topic choice and of the methodology of research. 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 maintain 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. It is published by the prestigious Polish academic publisher Wydawnictwo Uniwersytetu Śląskiego (University of Silesia Press), which provides an experienced team of editors to oversee the copyediting and technical side of the production. Updated information and all the issues published so far are available on the journal webpage at [www.tapsla.us.edu.pl](http://www.tapsla.us.edu.pl).

At the close of the first decade of its presence in the world of academia, *Theory and Practice of Second Language Acquisition* is undergoing the first important change in its leadership: in connection with her retirement Professor Danuta Gabryś-Barker, *TAPSLA*'s co-founder and co-editor-in-chief, has handed over her duties and responsibilities to Professor Jolanta Latkowska, who, as of 1st October 2023, is now the journal's co-editor-in-chief, together with Professor Adam Wojtaszek, who continues in this role which he assumed in 2015. This is the right moment to express thanks to Professor Gabryś-Barker, whose unsurpassed diligence, engagement, and devotion have transformed a platform first meant for disseminating papers that were presented at the International Conference on Foreign/Second Language Acquisition, organized by the Institute of Linguistics (previously the Institute of English) of the University of Silesia in Katowice, into an internationally recognised journal, indexed by the most important databases and steadily advancing towards the first quartile of the world's periodicals in the field of applied linguistics.

The present volume (10/1) offers a selection of texts, many of which report on qualitative research. Articles grouped in the first part of the volume share the common theme of online language learning and enforced electronic communication, which had to replace the usual face-to-face interaction between teachers and learners during the COVID-19 pandemic. All of them point

in varying ways to the positive aspects and potential benefits that may be drawn from this generally disheartening experience. There are a couple of papers that focus on the acquisition of selected aspects of grammar and attempt to find explanations for observed regularities, and several contributions that address some more general issues and various competencies developed in the process of language acquisition.

The first text, “A Systematic Review of Second Language Acquisition from the Perspective of Complex Dynamic System Theory” by Soheil Behdarvandirad, offers an interesting overview of recently published literature on the possible application of Complex Dynamic Systems Theory (CDST) to the explanation of the phenomena found in studies of the process of Second Language Acquisition. The author sets off from the assumption that there are so many different factors at play on the road to second language competence that it is virtually impossible to account for the observations satisfactorily if the perspective is narrowed down to a selection of variables or based on findings gathered during a short period of time. For this reason, the author advocates the perspective of CDST, which acknowledges a multitude of data points and the uniqueness of each individual development path. The literature review is not only a summary of recent findings, but also a collection of valuable methodological hints for future research based on the reviewed papers. In the second contribution, “Contextualising the Hyflex Model of Instruction for Language Classes,” author Robert Oliwa uses the exposure to digital media and electronically mediated communication which most of us were compelled to experience during the pandemic to propose a novel form of content presentation in language teaching. Focusing on the tertiary level, Oliwa develops and describes a new model incorporating skills that both teachers and learners have acquired in the potentials offered by online communication. This paper ends with some practical suggestions related to the applicability of the Hyflex Model to various educational conditions. The possible advantages of online communication for the development of language skills are investigated in the third paper of the volume, “Challenging but Rewarding Aspects of Telecollaboration: The Case of Virtual Israeli-Polish (VIP) Project” by Sabina Nowak and Aleksandra Rążewska. Unlike the previous contribution, this paper is based on observationally collected data and attempts to translate these findings into some educational conclusions which might increase the effectiveness of similar telecollaborative projects in the future. The authors investigated the influence of group dynamics and institutional/technological difficulties on the project’s flow and its results. Still on the theme of online communication, the fourth article, by Emerson Case, Agnieszka Kaczmarek, and Sebastian Zatylny, which is titled “Reading, Discussing, and Sharing: Creation of a Vibrant Transnational Online Learning Community through the International Book Club,” points to the advantages

of extending the usual experience of literature following from common reading programmes to the form of online discussion clubs. This study, which combines qualitative and quantitative longitudinal data, attempts to pinpoint the most significant benefits that can accrue from this particular form of contact with English-language literature. The topic of distance learning is continued in the fifth contribution, by Katarzyna Budzińska, titled “Foreign Language Classroom Anxiety in Emergency Remote Instruction.” The author takes the opportunity to investigate the widely studied construct of Language Anxiety in the new contextual framework created by the emergency remote teaching (ERT) conditions that we experienced in 2020 and 2021. Potential differences between two different educational environments, in the context of language anxiety experienced by learners, are explored by the author, as are effective ways to reduce anxiety.

The sixth contribution, “Challenges of Learning in Second Language among South African School Learners with Developmental Language Disorder,” by Nettie N. Ndou and Segun Emanuel Adewoye, focuses on a group of language learners with special educational needs, conditioned by their developmental language disorder. The authors apply a qualitative design to collect data from learners, language instructors, a speech therapist, and an educational psychologist, to receive a wide spectrum of perspectives. One of the conclusions emerging from the study is that using the target language as a medium of instruction significantly hampers the learning process and adds another difficulty to the already unfavourable situation of the group of learners under investigation. The next article, by Zofia Chłopek and Jacek Pradela, is titled “Parents’ Perceptions of Children’s Bi-/Multilingualism.” Instead of looking at the learners themselves, the authors chose to focus their attention on parents/caregivers and to find out how the bi-/multilingual competence of their children develops from their perspective. The results indicate that the parents’/caretakers’ observations echo the most significant facets described by researchers in earlier research. The author of the eighth paper, Fatih Ünal Bozdağ, concentrates on a quite specific area of developing grammatical competence of EFL, as evidenced by the title of the article, “Probabilistic Analysis of English Dative Constructions in Academic Writings of English EFL Learners.” Unlike most papers looking at this aspect of learners’ grammar, the author of this article does not compare learner language to native speaker data. Instead, only learner corpora are investigated in order to find emerging patterns and find new possible generalisations. One of the observations is that the learner’s native language is not a significant contributor to the observed variability. In the next paper, titled “Lexical Availability and Foreign Language Teaching: Main Contributions of a Growing Field,” Marta Sánchez-Saus Laserna presents cognitive and didactic aspects of using a French-born methodology called “lexical availability” in foreign language vocabulary instruction. Her focus is on teaching Spanish as a foreign

language and her major aim is to provide some useful suggestions related to optimal vocabulary selection in the teaching materials.

The tenth text of the present volume, “Graphemic Awareness Development of Polish Learners of Chinese as a Foreign Language,” by Zuzanna Wnuk, focuses on the teaching of the Chinese writing system to Polish learners of this language. Recognised as one of the most difficult aspects of the language to acquire for learners used to the Latin alphabet, Chinese characters turn out to be differently perceived and analysed at various stages of language proficiency development. The author, using a qualitative type investigation, managed to make a number of very interesting observations, which have a promising educational potential. The next text comes back to the issue of acquiring a specific grammatical construction in the English language, this time a certain form of passivisation. Hiromu Okamura and Tomohiko Shirahata, in the text “Factors Causing Overpassivisation of Unaccusative Verbs by Japanese Learners of English,” present interesting findings on Japanese learners of English. The earlier finding, that learners of English have the tendency to overpassivise unaccusative verbs, was postulated to have a number of different explanations. The authors wanted to find out which of the factors reported in earlier research was the strongest predictor for this observation. Their results show that for a particular group of Japanese learners of English one factor was indeed more often correlated with increased numbers of passives than others. The final text of the present volume shifts the focus to language teaching materials. Thiri Soe, in the text “Investigating Critical Thinking in ELT Textbooks: A Systematic Literature Review of Textbook Evaluation Studies,” presents a summary account of recently published evaluative textbook studies in which the issue of the inclusion of “critical thinking” was in some way addressed. The author reviewed 41 publications and identified four different aspects, which are further presented in the form of extensive concluding remarks. The contribution should be of great value to all interested in this topic who have not yet managed to find relevant publications on it. Thus, although the present volume does not include the customary book review section, the last paper could be treated as a certain kind of review of the subject-related literature.

We would like to emphasize that although the articles contained in this issue offer a selection of reports on empirical research, they are also strongly grounded in solid theoretical bases and overviews of the literature in given areas. The methodology used in these studies is mostly qualitative, but in some cases this is accompanied by quantitative analysis. The issue covers a wide array of different learning environments and languages, with a clearly dominating theme of distance learning mediated by electronic form of communication. We believe that the research presented here and the implications that can be derived from it have interesting potential not only for language practitioners

but also for teacher trainers and, equally importantly, for the content of teacher training programmes.

As our journal is fully published online in open access, our contributors are all able to reach a wide readership around the world to present their research, and thus to get feedback on their ideas. We hope that researchers, teachers, and students can all benefit from the present issue of *TAPSLA* and will find the articles published here not only useful but inspirational. We would therefore like to thank all the authors for their current contributions and, traditionally, extend our invitation to all Polish and foreign researchers and academics to share their work with us by submitting it to the journal in the future.

*Jolanta Latkowska*

 <https://orcid.org/0000-0002-9478-7689>

*Adam Wojtaszek*

 <https://orcid.org/0000-0003-0308-4337>



**Soheil Behdarvandirad**  
Texas State University, USA

 <https://orcid.org/0000-0003-4642-7993>

# **A Systematic Review of Second Language Acquisition from the Perspective of Complex Dynamic System Theory**

## **Abstract**

From the Complex Dynamic Systems Theory (CDST) point of view, second language development has unpredictable and non-linear patterns that vary from learner to learner. Keeping track of such dynamic development requires longitudinal studies with sufficient data points. The present systematic literature review attempts to present an overview on the previously conducted longitudinal studies that have investigated the development of second language subsystems from the CDST perspective. Starting from 1884 publications, the systematic searching strategy led to 45 articles which were examined in order to highlight the state of the art. The observations of the reviewed studies are conclusively supportive of the CDST principles in second language development. The synthesis of the findings of the papers will be presented and, finally, a multitude of suggestions for further research will be provided which can help future studies clarify the existing gaps that exist in the literature.

*Keywords:* second language development, second language acquisition, CDST, systematic review

## **Introduction**

During the past three decades a growing number of studies have started to focus on individual patterns of second language development instead of general developmental trends that can be observed among large groups (Larsen-Freeman, 1997; Larsen-Freeman, 2017; Han, 2020; Hiver et al., 2022; Rokoszewska, 2022). The introduction of Complex Dynamic Systems Theory (CDST) into the realm of applied linguistics has attracted attention to the un-

predictability and uniqueness of second language development among learners (Larsen-Freeman & Cameron, 2008; Verspoor et al., 2011). According to this theory, language consists of numerous interconnected systems where the whole system can be drastically impacted by small changes in one subsystem or in the initial state of the system (de Bot et al., 2007; van Geert & Dijk, 2002).

Unlike group studies where the focus is on general developmental patterns of larger numbers of language learners (Verpoor et al., 2021), examining second language development from the CDST perspective calls for longitudinal studies where usually fewer participants are examined (Lowie & Verspoor, 2019). Such studies have examined the development of different subsystems (e.g., syntactic and lexical complexity, accuracy, and fluency) with respect to different language skills (especially writing and speaking). This systematic literature review aims to present a general summary of the literature of CDST studies focused on second language development. After applying the inclusion criteria to the 1884 initially identified papers, 45 of them were selected for the final review. In addition to their findings, different methodological aspects of these studies are examined and compared, such as their duration, data points, number of participants, participant type, the investigated subsystems and skills, data collection methods, measurement, and, finally, as an important aim of this systematic review, the gaps and areas which require further investigation are mentioned and suggestions for future research are provided.

## Literature Review

The steps that we take during the journey of learning a new language have been of great interest for many researchers. Finding a specific path that all learners go through to acquire a second (or foreign) language can dissect this seemingly complicated developmental process. Many previous studies have tried to find general patterns in the development of various second language subskills in rather large sample sizes (Lowie et al., 2011; Verspoor et al., 2012) but such studies have inevitably overlooked the uniqueness of every individual's improvement path during the acquisition of a new language (Verspoor et al., 2021). The objective of capturing a generalizable pattern is too simplistic in reality (Larsen-Freeman, 2006, 1997). From the perspective of Information Processing model (IP model), the development of a second language follows a linear and predictable trend (de Bot et al., 2007; Shanker & King, 2002). In contrast, other theories such as Cognitive Linguistics, Functional Linguistics, Emergentism, and Competition Model take into consideration the numerous independent variables (psychological, social, environmental, etc.) that can play

significant roles at different levels of the development of a second language. Being non-linear and unpredictable, second language development has been characterized as dynamic and complex (Verspoor et al., 2017).

According to Dörnyei (2014), a system with at least two interlinked elements that can independently experience changes over time is considered dynamic or complex. Conducting empirical research in the field of dynamic systems, especially in social sciences, is indeed more difficult because of the almost innumerable interconnected elements that function independently and with each other at the same time, making the system unpredictable. The interconnectedness of all variables in a dynamic system leads to the fact that changes in one variable in the system influence the other coexisting variables, a characteristic called as “Complete Interconnectedness” (de Bot et al., 2007). Moreover, since the early states of dynamic systems are considerably influential on their development in the long run, the existence of the butterfly effect is another observable phenomenon in complex systems, causing small initial differences to have drastic long-term impacts.

From the perspective of Complex Dynamic Theory (CDST), second language development varies from person to person due to the multitude of impactful individual factors that vary among learners (de Bot et al., 2007; Larsen-Freeman & Cameron, 2008). Since this theory was introduced into the area of language acquisition about three decades ago (Larsen-Freeman, 1994), many researchers in different fields of applied linguistics have examined different aspects of CDST in their studies (Hiver et al., 2022). This theory brings to our attention the non-linearity of second language development in addition to the variability that exists between and within learners’ second language developmental patterns (van Dijk et al., 2011; Larsen-Freeman, 1997, 2006; Verspoor et al., 2008). According to CDST, language development is an emergent, context-dependent and dynamic process filled with complex connections (de Bot, 2008; Hiver et al., 2022). The developmental variability in the examined language subsystems and also the multitude of unpredictable interactions among them have been repeatedly observed in the literature (e.g., Spoelman & Verspoor, 2010; Verspoor et al., 2008; Verspoor et al., 2021; Verspoor and van Dijk, 2011). Even after averaging out a number of specific learner trajectories, Verspoor et al. (2011) as well as Larsen-Freeman (2006) have reported that the remaining developmental patterns of the examined groups were different from that of every group member. The development of different second language subsystems can take place in different orders and at different stages. In addition, the interactions between such subsystems are intertwined, adding to the complexity and unpredictability of the dynamic development. It should also be noted that while the dynamicity of second language development makes it unpredictable, it does not mean that this development is totally random (Larsen-Freeman & Cameron, 2008).

According to CDST, accurate observation of second language development requires longitudinal studies focused on the language development of individuals with enough data points (Lowie & Verspoor 2019; Verspoor et al., 2021). While the value of the contributions and findings of group studies even with few data points is undeniable, such studies are unable to keep track of the unique and flexible developmental trends of each learner. Some of the numerous cognitive and environmental factors that impact the dynamic development of a second language include motivation, anxiety, memory capacity, age, aptitude, intelligence, available learning time, available knowledge, level of education, maturity, and the amount of exposure to the new language (de Bot et al., 2007; Kliesch & Pfenninger, 2021; Zhang et al., 2022). For example, the study conducted by Piniel and Csizér (2015) indicated that learners with higher degrees of motivation and lower degrees of anxiety had more variable developmental patterns in comparison with others. Working memory capacity has been reported to correlate with the success of second language acquisition (Linck et al., 2014; Serafini, 2017). Additionally, previous studies that examined how age can influence the cognitive performance of second language learners have reported inconclusive findings. In two studies conducted by Bak et al. (2016) and Wong et al. (2019), the observations were supportive of better cognitive performance of older second language learners. However, such cognitive benefits were not reported by Berggren et al. (2020) and Ramos et al. (2016).

A multitude of previous studies have attempted to examine different areas of second language development from the CDST viewpoint. Such investigated areas include complexity, accuracy, and fluency (CAF) in speaking (Larsen-Freeman, 2006; Lowie et al., 2017; Sauer & Ellis, 2019), and writing (Larsen-Freeman, 2006; Spoelman & Verspoor, 2010; Verspoor et al., 2021; Zhang et al., 2022), in addition to vocabulary (Caspi & Lowie, 2013; Zheng, 2016), and pronunciation (Munro & Derwing, 2008). A number of previous studies have also examined how the dynamic development of a second language is influenced by issues such as self-concept (Mercer, 2011), motivation (Han & Hiver, 2018; Lowie & Verspoor, 2019; Nitta & Baba, 2015; Nitta & Baba, 2018; Zhang et al., 2022), individual differences (Lowie & Verspoor, 2019; Nitta & Baba, 2018; Pfenninger, 2022), and also corrective feedback (Fogal & Koyama, 2022). While these factors are indeed impactful on language developmental patterns, it does not mean that grouping learners by these variables can necessarily result in exactly similar learning patterns. For example, even after grouping the participants by their aptitude and motivation, no similar developmental patterns were observed by Lowie and Verspoor (2019).

Hiver and Al-Hoorie (2016) argued that the main goals of CDST research in the area of applied linguistics include (a) representations of definite complex systems at different scales; (b) identification of the outcomes of emergent systems and their dynamic patterns of change; (c) tracing and possibly modeling

the complex mechanisms of the emerging patterns; and (d) understanding how the behavior of the systems can be influenced by the relevant parameters. As a classic example among the studies that have attempted to examine and keep track of the dynamic development of second language, Larsen-Freeman (2006) examined the developmental patterns in speaking and writing of five Chinese second language learners of English at high-intermediate level of proficiency over the period of six months. The tasks that the participants were asked to complete were freely writing about past events in addition to retelling these stories orally three days after writing them. The participants completed four writing and four speaking tasks during the six-month time period. The developing linguistic subsystems under scrutiny were grammatical complexity (measured with average number of clauses per t-unit), vocabulary complexity (measured with word types per square root of two times the words), accuracy (measured with the proportion of error-free t-units to t-units), and fluency (measured with average number of words per t-unit). The results were indicative of the non-linearity of second language development, waxing and waning developmental patterns, and also inter- and intra-individual variations on the linguistic measures, all of which were supportive of CDST view of second language development. Although only four data points seem to be a small number considering the fact that tracking second language development calls for numerous data points over long periods of time (in comparison with group studies), the aforementioned study was one of the pioneering ones that examined second language learning through the CDST angle. Since then, different fields of applied linguistics have benefitted from the contributions of CDST (Hiver et al., 2022). Such areas include language acquisition (Lowie et al., 2010; Verspoor et al., 2008), educational linguistics (Hult, 2010), the evolution of language (Mufwene et al., 2017), planning and policies in language (Bastardas-Boada, 2013; Larsen-Freeman, 2018), language ecology (Kramsch & Whiteside, 2008), and sociolinguistics (Blommaert, 2014) among other realms.

While previous studies in the literature have shed light on different aspects of dynamic development of second languages, there seems to be the lack and need of systematic reviews that can offer bigger pictures of the current state of the art in this field. The aspects of the previous studies that can be investigated through systematic reviews are the general focus of previous studies, their designs, the regularities within and contrasts between their findings, and the gaps which have not received sufficient attention yet. Hence this systematic literature review attempts to address the abovementioned issues in the literature of second/foreign language development from the CDST point of view. To be more specific, the present study aimed to address the following questions:

RQ1. What are the design and methodological characteristics adopted by the CDST studies in the field (including their conduction place, duration, number of data points, number and characteristics of participants, contexts, language

skills, data collection methods, the examined subsystems, and the utilized measurements)?

RQ2. What are the important patterns in terms of findings?

## Method

### Design of Search Strategy

The online search for finding the relevant articles was done using Scopus. With the aim of finding thoroughly inclusive keywords for the search queries, after an initial scanning of the most cited papers in the relevant realm, the following string was chosen and searched for the titles, abstracts, and keywords of the articles in Scopus:

((("L2" OR "second language" OR "foreign language") w/2 (development OR acquisition OR learn\* OR complexity)) AND (longitudinal OR "case stud\*")) OR ((("L2" OR "second language" OR "foreign language") AND ("complex dynamic system\* theory" OR "dynamic development" OR "complex dynamic system\*"))

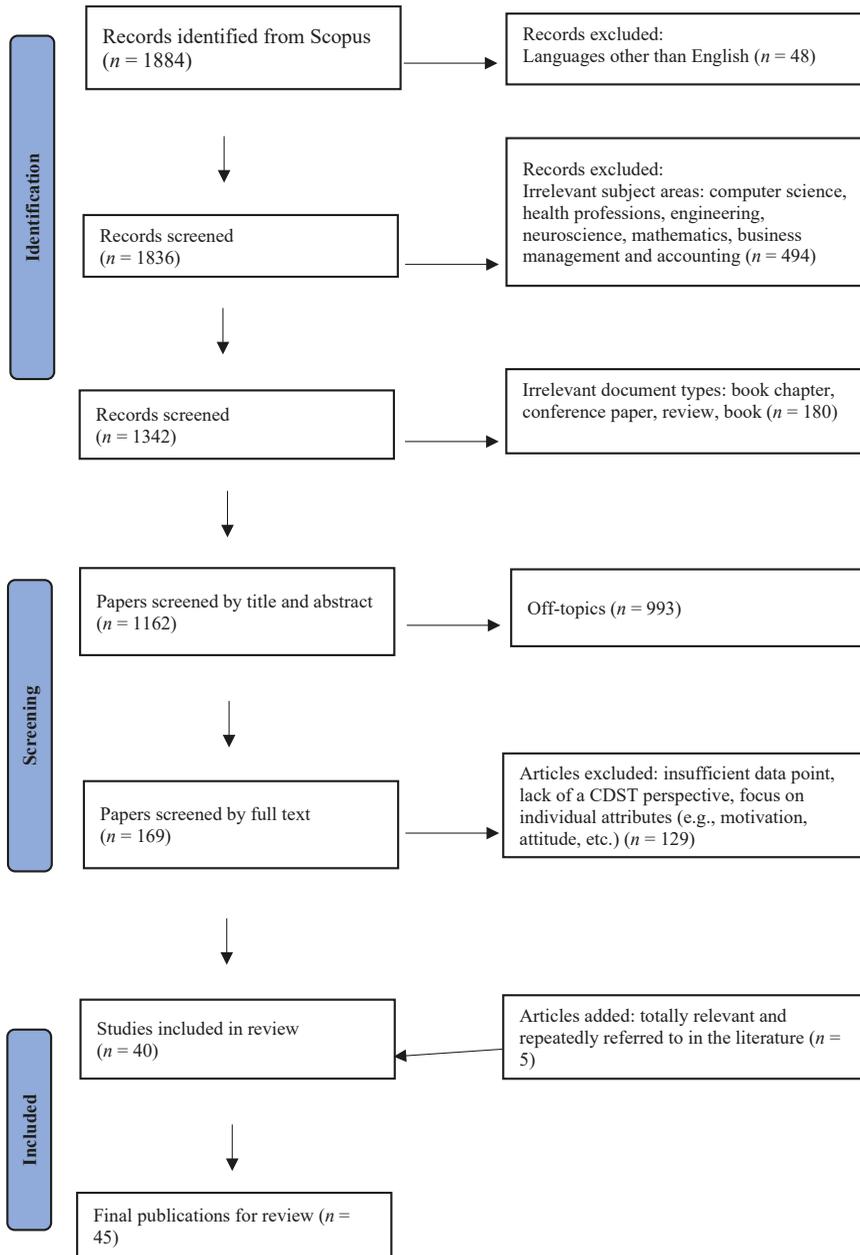
Figure 1 illustrates a PRISMA flowchart of the search process. After excluding languages other than English and also irrelevant document types and subject areas (such as "Computer Science," "Health Professions," Engineering," "Neuroscience," "Mathematics," and "Business, Management and Accounting"), 1162 was the number of articles left. As the next step, title and abstract filtering was conducted after which 169 articles remained. Forty articles passed the full text filtering of these papers. Moreover, five additional articles were identified as studies that were referred to in the literature, were totally relevant, and completely met the inclusion criteria but had not been detected in the Scopus database search. The searching strategy led to 45 final articles that are included in the present systematic review.

More detailed criteria for the selection of the articles during the manual filtering phase were the followings:

- Publications were included only if they involved empirical research.
- Publications were included only if they were longitudinal with enough data points (at least four) since these two criteria are the fundamental elements present in studies investigating second language development from CDST

**Figure 1**

*PRISMA Flowchart of the Search Process and Identification of Studies via Databases and Registers*



perspective. Studies with these two aforementioned traits were included even if “CDST” was not mentioned in them.

- Since the purpose of this systematic review was to examine previous studies that have investigated the development of second language skills and their subsystems from CDST perspective, publications were included only if they kept track of the language development, and the ones focused on other variables such as attitude, motivation, awareness, willingness and perceptions were excluded.
- Publications were excluded if they were only concerned with the effectiveness of teaching methods and strategies and developmental patterns were not investigated.
- Publications were excluded if they were focused on newborn bilingual children who were learning two languages simultaneously. The reason for this exclusion was the fact that since newborns have not acquired their first language completely, the two languages that bilingual newborns learn at the same time cannot be distinctively identified with respect to which one can be considered as the first language and which one can be the second.

## Results

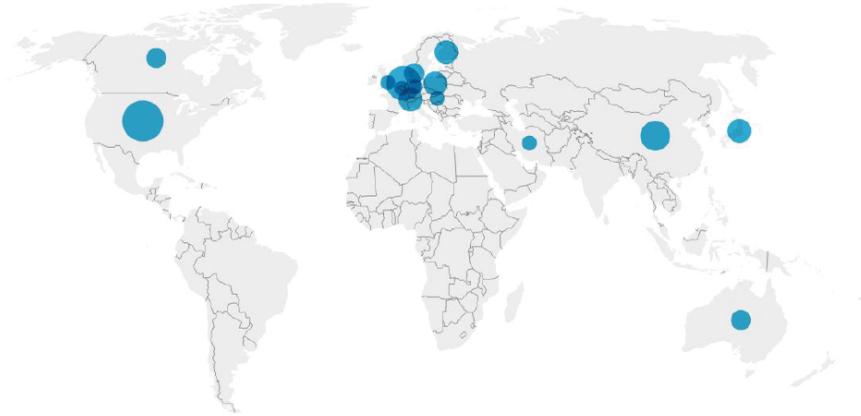
Table 1 presents an overview of the 45 reviewed studies providing information on their authors, years and places of publication, timespans, data points, the examined skills and subsystems, data collection methods, and measurement in addition to the number of participants and their backgrounds, first languages, second languages, and second language levels of proficiency. Additionally, different aspects and methodological characteristics of the articles are examined in this sections.

### Second Language Development and CDST around the World

The 45 studies examined in this review have been conducted in 16 countries. The most research was conducted in the US with ten studies, comprising almost one fourth of all of the studies. Mostly due to the works of Verspoor and Lowie, the Netherlands is the next country where the number of contributions to the literature has been significantly more than that of other countries (seven studies). Except China with five studies, the number of articles in any other country does not exceed three. Figure 2 illustrates the distribution of the conducted studies around the world.

**Figure 2**

*The Distribution of Previous Studies around the World*



### **Duration and Data Points**

Lengthier duration of the developmental tracking and the abundance of data points are two required fundamental characteristics of CDST in the field of second language development (Lowie & Verspoor, 2019; Verspoor et al., 2021). As can be seen from Figure 3, the duration of the longitudinal CDST studies varies from two, in Casillas (2020), to 72 months, which was the timespan of the study conducted by Pfenninger (2022). The average duration of all of the reviewed articles is almost 18 months. It is worth-mentioning that the duration of twelve studies was an academic year (nine months). Out of all of the 45 studies, 18 of them lasted at least for one year which is indicative of the importance of longer studies for the investigation of second language development from the CDST perspective.

The number of data collection points in the studies ranged from four (Larsen-Freeman, 2006) to 100 (Chan et al., 2015; Lowie et al., 2017). The approximate mean of data points among the reviewed studies was twenty-three. As can be seen in Figure 4, the number of data points that the studies had were rather evenly distributed and they were not densely clustered around a specific number. As the most repeated number of data points, participants were tested 30 times in five of the conducted studies (Baba & Nitta, 2014; Baba & Nitta, 2021; Chang & Zhang, 2021; Evans & Larsen-Freeman, 2020; Zhang et al., 2022).

**Table 1**  
*An Overview of the Studies*

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Zhang et al.	2022	9	30	Writing	Lexical complexity, syntactic complexity, accuracy, and fluency	Writing tasks with topics from IELTS	Lexical complexity: low-free- quency word types/total word types; Syntactic complexity: clauses/T-units; Accuracy: error free T-units/T-units, Fluency: word/T-units	A2	University students	22	Chinese	English
Pfenninger	2022	48 to 96	16 to 32	Speaking and writing	Lexical complexity, syntactic complexity, accuracy, and fluency	Writing a timed English narrative and a re-telling oral task	Fluency: written text length and pruned syllables per minute; Complexity: clauses/T- unit and clauses/analysis of speech units (AS-unit); Lexical richness: total word count/total factor count; Accuracy: total number of error-free T-units/ AS-units	B2	50/50 bilingual school students	71	Swiss German	English
Rokoszewska	2022	36	21	Writing	Lexical complexity, syntactic complexity, accuracy, and fluency	Written essays during classes	Syntactic complexity: clauses/ T-unit; Lexical complexity: complex type-token ratio; Accuracy: correct T-units/all T-units in a text; Fluency: the average number of words/T- units	C1	Secondary school student	100	Polish	English
Wind	2021	7	7	Writing	Syntactic and lexical complexity	IELTS type argumentative essays	Lexical complexity: the average word length (AWL) index; Syntactic complexity: finite verb ratio (FVR)	B2	An IELTS preparatory course student	6	Hungarian	English

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Routarinne & Ahlholm	2021	9	48	Speaking	Making requests	Video-recorded class sessions	Microanalysis of multimodal interaction	B1	Nine-year-old school student	1	Russian	Finnish
Gui et al.	2021	4	12	Academic reading	Academic reading ability in chemistry	Pen and paper test	The test question types: Vocabulary, True or False judgment, Syntactic parsing, Translation from English into Chinese, and Summary	B1	Chinese chemistry major undergraduates	27	Chinese	English
Lesonen et al.	2021	9	28 to 35	Speaking and writing	Verbal and adjectival constructs	Writing and talking about chosen topics	Normalized frequencies	A1	University students	4	Varied	Finnish
Kliesch & Pfenninger	2021	7	30 to 32	Speaking	Speaking and integrative L2 skills, lexical reception, grammatical reception, fluency, morphosyntactic accuracy, target-like use of lexical items, and lexical richness	Weekly multiple oral and written tests	Integrative L2 Skills: C-test, Lexical reception: odd-one-out task, Grammatical reception: a created test, and oral interviews for Fluency, Morphosyntactic accuracy, Target-like use of lexical items, and Lexical richness	A1	A 64-year-old learner	28	Dutch	Spanish
Baba & Nitta	2021	9	30	Writing	Fluency	A timed narrative-writing task	Number of words in an essay	B2 and C1	University students of an academic writing course	105	Japanese	English

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Chang & Zhang	2021	42	30	Listening	Listening performance	IELTS listening tests	IELTS listening test scores	C1	University students majoring in statistics	3	Chinese	English
Verspoor et al.	2021	5.5	23	Writing	Syntactic and lexical complexity, and fluency	Short writing tasks	Judgements of ten experience raters; Syntactic complexity: mean length of T-unit; Lexical complexity: Guiraud	Various levels	Secondary school student	23	Dutch	English
Evans et al.	2020	9	30	Speaking	Syntactic complexity	Narrative tasks and dialogic conversations about daily topics	Analysis of speech units (ASunits): independent clause, or sub-clausal unit together with any subordinate clause(s) associated with either	A2	A 27-year-old male immigrant	1	French	English
Casillas	2020	2	8	Pronunciation	Development of Spanish stop voicing contrasts	Delayed repetition and reading	Praat (Voice onset time: the difference (in ms) between the onset of voicing and the burst; Relative voice onset time: the aforementioned VOT value divided by the duration of the CV (stop + vowel) sequence)	A1	7-week domestic immersion program students	20	English	Spanish
Huang et al.	2020	9	12	Writing	Complexity, accuracy and fluency	Free writing	Holistic rating (Complexity: sentence structure; Accuracy: correctness of grammar use; Fluency: length of the text within a certain duration of time)	B1	University students	2	Chinese	English

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Fogal	2020	48	42	Writing	Authorial voice (writer, writing situation, disciplinary expectations, genre expectations, Lexico-rhetorical devices, and Dialogism)	University writing tasks	Three informed outside raters using certain developed criteria	C1	University student in Thailand studying actuary science in English	1	Thai	English
Khomeijani Farahani et al.	2020	9	10	Writing	Complexity, accuracy and fluency	English textbook writing tasks	Complexity: words/finite verbs ratio; Accuracy: error-free T-unit ratio; Fluency: words/T-unit	B2	A 17-year-old senior secondary school student	1	Farsi	English
Rokoszewska	2020	9	21	Speaking	Lexical complexity (lexical density, sophistication, variation, and frequency)	Descriptive and argumentative interviews about the topics covered in the coursebooks	Lexical density: the number of lexical tokens/total number of tokens; Sophistication: the number of more advanced tokens/total number of lexical tokens; Variation: sophisticated or complex type-token ratio (CTTR); Frequency: the percentage of words used by the learner at different frequency levels	"Good, average, and poor"	16-year-old secondary school learners	3	Czech	English
Yu & Lowie	2020	4	12	Speaking	Complexity and accuracy	Speaking for 2 to 3 minutes about IELTS Speaking Test topics	Complexity: mean length of speech units in words; Accuracy: number of error-free past tenses and speech units	B2	College learners of L2 English	10	Chinese	English



Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Lowie & Verspoor	2019	7	23	Writing	Syntactic and lexical complexity	Free writing tasks related to participants' lives or class topics	Experienced raters and analytical measures (Syntactic complexity: the mean length of T-Unit and the number of dependent clauses/T-unit; Lexical complexity: mean length of word and Giraud; Complexity of noun phrases: the number of dependents/nominal in noun phrases)	B1 and B2	Secondary school students	22	Dutch	English
Menke & Strawbridge	2019	33	11 to 17	Academic writing	Syntactic complexity in academic discourse	Written assignments completed during the major	Length-based measurement: a)mean length of clause, b) mean length of T-unit, c)mean length of noun; Inter-clausal relationship indices: a)clauses / T-unit, b)grammatical intricacy, c)simple sentence ratio; Phrasal/clausal variety: a)noun phrase accessibility hierarchy, b)noun phrase modification types	B1	University students of a Spanish major program	3	English	Spanish
Bulté & Housen	2018	19	11	Writing	Syntactic complexity	Writing tasks about chosen topics	Mean length of T-unit, sub-clause ratio (subclauses/clauses), Coordinate clause ratio (coordinate clauses/sentences), Mean length of finite clause (word tokens/finite clauses), Mean length of noun phrase (word tokens in NPs/NPs)	A1 and A2	Secondary school students	10	Dutch	English

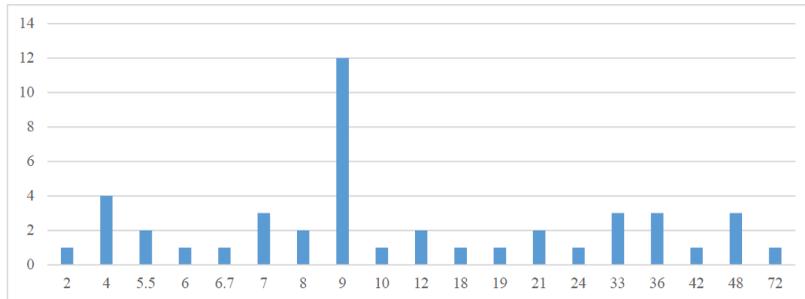
Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Lowie et al.	2017	8	100	Speaking and writing	Syntactic and lexical complexity	Speaking and writing tasks (TOEFL topics)	Syntactic complexity: mean length of T-units; lexical diversity, VocD (which is an adjusted metric for the type or token ratio)	A2	Identical 15-year-old twins	2	Chinese	English
Verspoor et al.	2017	48	22	Writing	Syntactic and lexical complexity	Academic papers	General (lexical: average word length, syntactic: finite verb ratio) and specific (lexical: finite adverbial, nominal and relative clauses, syntactic: non-finite constructions)	C1	University students	3	Dutch	English
Zheng	2016	10	8	Writing	Lexical complexity	Academic writing tasks	Uber's index (lexical diversity), VocabProfile software (the frequency-based lexical measures), AntConc (searching the target bundles)	B2	First-year university students enrolled in an English major program	15	Chinese	English
Vyatkina et al.	2015	21	17	Writing	Syntactic complexity	Writing essays as curricular tasks	ANNIS (used for pronominal (attributive) adjectives, cardinal numbers, predicative and adverbial adjectives, adverbs)	A1	University students enrolled in a basic German language program	12	English	German
Eskildsen	2015	24	5	Speaking	Yes/no and WH interrogatives	Audio-visual recordings of classroom interaction	Qualitative categorization and analysis	A1	University students	2	Spanish	English

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Chan et al.	2015	8	100	Speaking and writing	Syntactic complexity	Oral and written tasks (topics selected from TOEFL tests)	Mean length of T-unit, Dependent clauses/T-unit, Coordinate phrases/T-unit	A2	Identical twins	2	Taiwanese	English
Baba & Nitta	2014	9	30	Writing	Writing fluency	Compositions as assignments	The total number of words in a composition	A2 and B1	University students	2	Japanese	English
Kowal	2014	33	6	Writing	Writing fluency	Writing narratives of personal experiences	Mean transition time, Mean length of burst	A1	University students	15	Polish	Swedish
Rosmawati	2014	4	10	Writing	Complexity and accuracy	Academic argumentative essays on TOEFL topics	Complexity: word/finite verb, Accuracy: error-free clause/ clause	C1	A post-graduate student in Australia	1	Japanese	English
Polat & Kim	2014	12	24	Speaking	Lexical complexity, syntactic complexity, and accuracy	Unstructured, friendly inter-views	Syntactic complexity: mean length of AS-units, clauses/ AS-unit, mean length of clauses; Lexical diversity: D; Accuracy: errors/100 words	C1	An untutored immigrant	1	Turkish	English
Zhang & Lu	2013	4	6	Writing	Chinese numeral classifier system (fluency, diversity, and accuracy)	Written essays (coursebook topics)	Fluency: number of classifier tokens or types/hundred characters, Diversity: type-token ratio of classifiers, Accuracy of produced classifiers; accuracy rate of classifiers produced by learners, Accuracy of classifiers in obligatory environments; accuracy rate of classifiers in all obligatory environments	B1 and B2	University students	87	English	Chinese

Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Caspi & Lowie	2013	9	36	Writing	Academic vocabulary recognition, recall, controlled and free production	Longitudinal Academic Vocabulary Tests (LAVT) and written essays	Recognition, recall, and controlled production: the Longitudinal Academic Vocabulary Tests (LAVT), Free production: the ratio of correct academic (UWL and AWL) word families to the total number of academic word tokens divided by the total number of correct content words and multiplied by the general family or token ratio	C1	A university student	1	Portuguese	English
Vyatkina	2012	21	19	Writing	Complexity	Curricular writing tasks	General complexity: words/sentence (sentence length); Clausal complexity: clauses/sentences and words/clauses; Coordination and subordination: normalized coordinating conjunctions and subordinating conjunctions frequencies/100 words; Lexicogrammatical variety: type-token ratio	A1	University students	2	English	German
Spoelman & Verspoor	2010	36	54	Writing	Complexity and accuracy	Writing samples on academic topics	Accuracy: the difference between the total number of cases and the number of incorrect cases divided by the total number of cases. Word complexity: the average sentence length in morphemes and the average sentence length in words, NP complexity: averaging NP length in words, Sentence complexity: averaging number of dependent clauses per text	A1	A theoretical linguistics university student who took a minor in Finnish	1	Dutch	Finnish

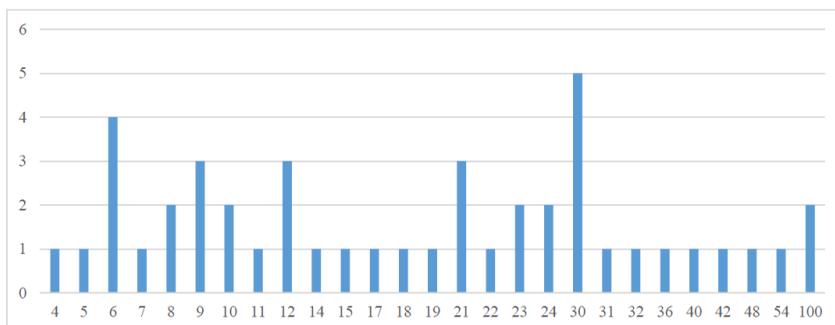
Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Eskildsen	2009	48	40	Speaking	Usage of "car"	Audio-visual recordings of classroom interaction	Qualitative and manual measurement	A1	Language school students	1	Spanish	English
Li & Schmitt	2009	9	9	Writing	Lexical phrases development and appropriateness	Academic writing essays	The judgement of a panel of raters	B2	An MA student in English Language Teaching program	1	Chinese	English
Munro & Derwing	2008	12	6	Pronunciation	Vowel intelligibility	Delayed repetition task	The judgement of four phonetically trained Canadian judges	A1	Immigrants in an ESL program	44	Chinese and Slavic	English
Mellow	2008	6.7	15	Writing	Complexity: acquisition of resolutions of the argument dependencies of verbs	Written narratives summarizing wordless picture books	Judgements of raters	A2	12-year-old immigrant school student	1	Spanish	English
Serrano & Howard	2007	33	9	Writing	Composition, grammar, and mechanics	Free narrative writing assignments	A developed analytic rubric: Composition (topic development, sentence formation, supporting details, descriptive language), Grammar (verbs, agreement, placement, prepositions), Mechanics (spelling, punctuation, capitalisation, paragraph formation)	B2	Language school children students	2	Spanish/English	English/Spanish

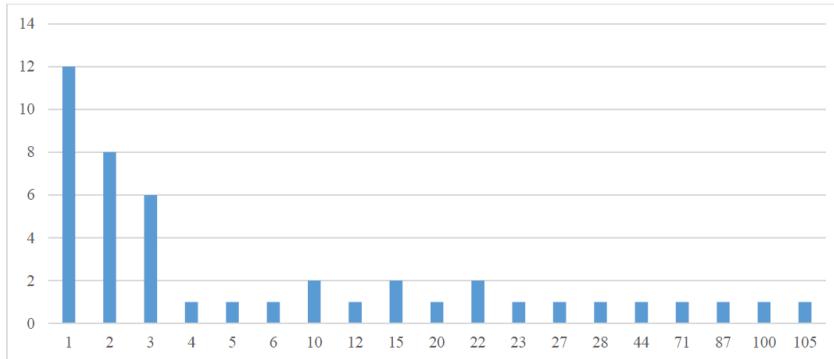
Author(s)	Year	Timespans (months)	Data Point	Skill(s)	Subskill(s)	Data Collection Method	Linguistic measurement	Proficiency	Setting	Participant(s) (number)	L1	L2/FL
Larsen-Freeman	2006	6	4	Speaking and writing	Lexical complexity, syntactic complexity, accuracy, and fluency	Writing a narrative about a past episode and telling the stories orally	(In addition to qualitative analysis) Grammatical complexity; average number of clauses/T-unit, Vocabulary complexity; word types/square root of two times the words, Accuracy: the proportion of error-free T-units to T-units, Fluency: average number of words/T-unit	B2	Immigrants	5	Chinese and Slavic	English
Zhang	2004	9	9	Speaking	Adjective marker -de in Chinese	Free-production speech samples	Emergence criterion (the adjective suffix -de(ADJ) was viewed as having emerged if there were a minimum of four tokens of it in a sample set), qualitative analysis.	A1	University students	3	English	Chinese
Skiba & Dittmar	1992	36	21	Speaking	Morphosyntactic development and grammaticalization	Video and cassette recordings	Qualitative measurement in addition to valency and co-occurrence of expressions	A1	Immigrants	3	Polish	German
Hanania & Gradman	1977	18	18	Speaking	General development	Tape-recorded conversations	Qualitative measurement of numerous linguistic components	A1	An immigrant	1	Arabic	English

**Figure 3***The Frequencies of Timespans*

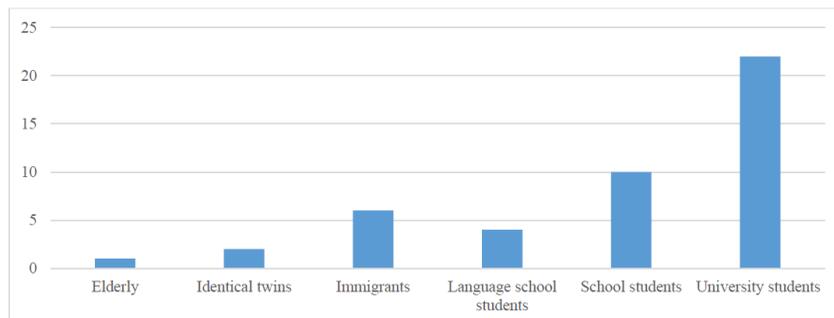
### Participants and Settings

Figure 5 illustrates the number of participants in the reviewed studies. Although in two of the studies one hundred or more participants were examined (Baba & Nitta, 2021; Rokoszewska, 2022), these are rare cases in the realm of longitudinal CDST investigations of second language development. More than half of the studies had at most three participants whose second language developmental patterns were tracked. A considerable portion of the studies focused on the second language development of only one participant. Regarding the background of the participants (Figure 6), 22 studies examined university students. Next, school students, such as high school, secondary school and elementary school students (in ten studies), and immigrants (in six studies) comprised the biggest portion of the studied participants. Two of the papers studied the second language developmental patterns of identical twins with the aim of having a clearer comparison between participants and minimizing the impact of extraneous variables on second language development as much as possible (Chan et al., 2015; Lowie et al., 2017).

**Figure 4***The Frequencies of Data Points*

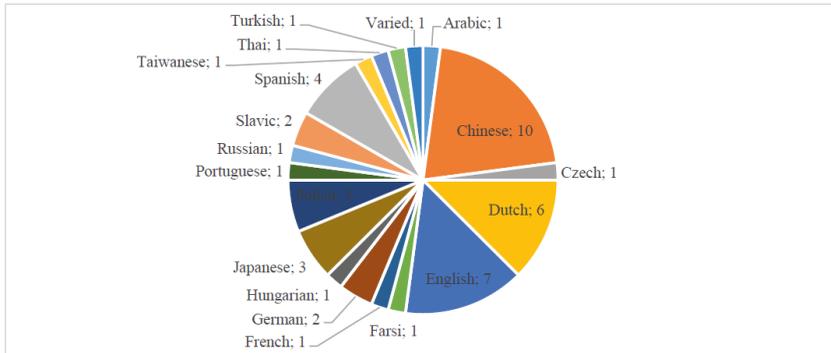
**Figure 5***The Frequencies of Number of Participants*

The participants of the reviewed studies had different first languages (Figure 7). Chinese (21%), English (15%), and Dutch (13%) were the most observed first languages. However, the second languages that the participants were learning were not so different (Figure 8). In more than two thirds of the studies, English was the language that participants were learning. The other learned second languages included Spanish, German, Finnish, Chinese, and Swedish. Among the 45 papers, 37 mentioned the second language proficiency levels of their participants. Figure 9 shows the distribution of participants' level of proficiency in the reviewed studies. The biggest portion of the studies were focused on dynamic second language development of participants at beginner (A1) proficiency level.

**Figure 6***The Frequency of Contexts*

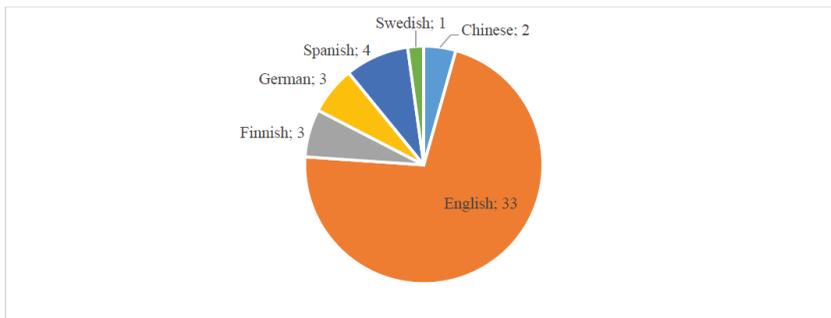
**Figure 7**

*The Frequency of Participants' First Languages*



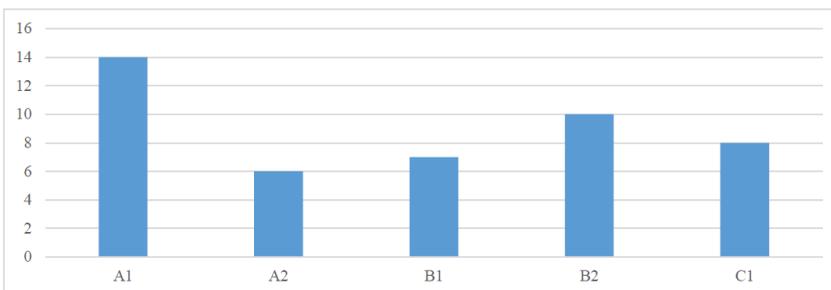
**Figure 8**

*The Frequency of Participants' Second or Foreign Languages*



**Figure 9**

*The Frequency of Participants' Level of Proficiency*

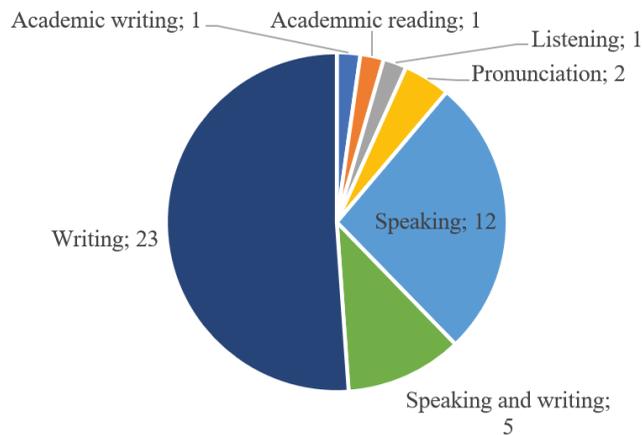


## Language Skills, Data Collection Methods

More than half of the studies used writing in order to observe the developing trend of second languages among participants. Speaking, and a combination of speaking and writing were respectively 11 and five times the focus of the reviewed studies. Only five studies were dedicated to the dynamic second language development of other language skills and subskills, including pronunciation, listening, academic reading, and academic writing. The fact that the development of a newly learned language can be more accurately traced in the production of utterances (as opposed to receiving utterances) explains the considerably abundant writing and speaking studies in comparison with the reading and listening ones. It should also be noted that in some studies the aim was to investigate the development of specific grammar structures while speaking or writing were just ways through which the structures were investigated. For example, Eskildsen (2015), Eskildsen (2009), and Zhang (2004) attempted to keep track of the development of interrogatives, use of the modal verb “can,” and adjective makers respectively. Figure 10 illustrates the proportions of each language skill among the reviewed papers.

**Figure 10**

*The Portions of Examined Language Skills*



A variety of ways were adopted to collect data from the participants (Table 2). In order to examine the development of participants' writing, the studies utilized class writing tasks, academic writing samples, and narratives essays in addition to writings about chosen topics, TOEFL topics, and IELTS topics. The studies focused on spoken language of the participants utilized interviews, recorded speech, narratives, and tasks about IELTS and TOEFL topics. As a technique that requires participants to repeat words and phrases

after hearing them, delayed repetition was used by Munro and Derwing (2008) for collection of data on participants' pronunciation. This method was combined with reading out loud in Casillas (2020). As the only two studies which focused on the developmental patterns of receptive skills, Gui et al. (2021) and Chang and Zhang (2021) used a pen and paper and IELTS listening tests respectively to investigate their participants' development in academic reading and listening.

## **Subsystems and Measurement**

Table 3 presents an overview of the subsystems which were investigated in the reviewed papers. Complexity (both syntactic or lexical), accuracy, and fluency (CAF), as three central aspects of language used for measuring second language development (Barrot & Agdeppa, 2021), have been the most commonly investigated facets of language production in the reviewed papers. Among the 41 studies where the development of writing and speaking were investigated, 28 kept track of the development of at least one of CAF elements. The rest of the examined subsystems include academic vocabulary, linguistic constructs, and interrogatives. The measurements utilized for each subsystem in each study can also be seen Table 1.

## **General Observations**

All of the reviewed papers reported findings that were supportive of the CDST principles in second language development. Inter- and intra-individual variability, non-linearity, and dynamicity were the characteristics that the 45 reviewed articles repeatedly used to describe the developmental patterns of the different examined subsystems among learners. It should also be noted that 20 papers reported the existence of at least some degree of similarities between how the examined subsystems developed among learners. Such similarities included general developmental trends, second language trajectories, and learning prototypes. The detection of such similarities does not contradict the dynamicity and variability of second language development since only some aspects of developmental patterns were reported to follow regularities. Finally, a few studies reported the significant role of individual characteristics like age and level of proficiency in second language developmental patterns.

**Table 2**  
*Data Collection Methods among Language Skills*

Skills	Data collection methods	Count
Writing	Academic writings	4
	Free writing	2
	Writing about specified topics	6
	Writing about TOEFL topics	3
	Writing narratives	6
	Writing tasks IELTS	2
	Written class essays	6
Reading	Pen and paper test	1
Listening	IELTS listening test	
Speaking	Interviews	4
	Recorded speech	5
	Speaking about IELTS topics	1
	Speaking about specified topics	1
	Speaking about TOEFL topics	2
	Telling narratives	3
Pronunciation	Delayed repetition and reading	1
	Delayed repetition	1

**Table 3**  
*The Examined Subsystems among Language Skills*

Skills	The examined subsystems	Count
Writing	Academic vocabulary recognition, recall, controlled and free production	1
	Authorial voice	1
	CAF	2
	Chinese numeral classifier system (fluency, diversity, and accuracy)	1
	Complexity	1

Complexity and accuracy	2
Complexity: acquisition of resolutions of the argument dependencies of verbs	1
Composition, grammar, and mechanics	1
Fluency	3
Lexical complexity	1
Lexical complexity, syntactic complexity, accuracy, and fluency	2
Lexical phrases development and appropriateness	1
Syntactic and lexical complexity	3
Syntactic and lexical complexity, and fluency	1
Syntactic complexity	3
<hr/>	
Speaking	
Adjective marker -de in Chinese	1
Complexity and accuracy	1
General development	1
Lexical complexity	1
Lexical complexity, syntactic complexity, accuracy, and fluency	1
Lexical complexity, syntactic complexity, and accuracy	1
Making requests	1
Morphosyntactic development and grammaticalization	1
Speaking and integrative L2 skills	1
Syntactic complexity	1
Usage of "can"	1
Yes/no and WH interrogatives	1
<hr/>	
Speaking and writing	
Lexical complexity, syntactic complexity, accuracy, and fluency	2
Syntactic and lexical complexity	1
Syntactic complexity	1
Verbal and adjectival constructs	1
<hr/>	
Listening	
General listening performance	1
<hr/>	
Reading	
Academic reading ability in chemistry	1
<hr/>	
Pronunciation	
Development of Spanish stop voicing contrasts	1
Vowel intelligibility	1
<hr/>	

## Discussion

The two research questions that this review has attempted to investigate were focused on the methodological characteristics, and also the patterns of findings in previous CDST studies in the field of second language development. Regarding the first research question, most of the studies were expectedly focused on productive language skills (speaking and writing) as their development can be more easily measured and traced. More than half of the papers examined the development of at least one of the CAF subsystems. Abundance of data points and lengthier timespans were other expected characteristics which were observed. The biggest portion of the studies (73%) had participants who were either university students or school students. Additionally, the participants in 14 studies were at beginner level with respect to the second language they were acquiring.

To address the second research question, the observations of the 45 reviewed studies conclusively support the dynamicity of second language development. The studies tracked the development of participants with various characteristics from different backgrounds and also in different contexts and, although some general regularities and patterns were seen, all of the studies were supportive of the CDST view in second language development. Consistently, patterns of participants' second language development were indicative of within and between individual variability. Patterns of development were non-linear and various between and within participants. The examined subsystems showed interconnectedness as well. The innumerable individual and extraneous variables that impact second language acquisition make this development a dynamic one. A few of such variables include affect, motivation, environmental factors (Khomeijani et al., 2020; Zhang et al., 2022), age of onset (Pfenninger, 2022), first language fluency, individual learner investment, amount of available free time, and competition among learners (Kliesch & Pfenninger, 2021).

While the majority of the reviewed studies were focused on development of subsystems (especially complexity, accuracy, and fluency) in writing and speaking, the dynamic nature of second language development was also supported in the few studies which were focused on the receptive language skills (reading and listening). Keeping track of the reading ability gains of 27 English learners 12 times over an academic semester, Gui et al. (2021) observed that the development among the learners was individual and non-linear. The performance of three Chinese learners of English in IELTS listening tests during 3.5 years was also supportive of the dynamicity of second language listening development in Chang and Zhang (2021).

Among the reviewed papers, two separate studies attempted to investigate and compare second language development among identical twins (Chan et al.,

2015; Lowie et al., 2017). Consistent with CDST, the findings of both studies showed that second language acquisition is idiosyncratic and patterns of development and even the degree of variability changed from twin to twin. The study conducted by Wind (2021) showed that even the development of self-reflection was individual and dynamic. The development of phonetic skills was similarly shown as dynamic and non-linear in Munro and Derwing (2008) and Casillas (2020).

Some of the studies have found ways of roughly categorizing the dynamic development of learners. In the research conducted by Zhang et al. (2022), three different prototypical patterns were observed, which included three groups of participants: participants with continuous stable development, participants with initial fluctuating development followed by steady development, and participants with constant fluctuating developmental patterns. Similarly, general patterns of development were shared among three groups of participants in Baba and Nitta (2021). Examining the narrative writing tasks of university students, collected 30 times over an academic year, showed three general developmental patterns including a stagnating, steadily growing, and markedly growing patterns. A number of common syntactic and lexical developmental patterns were also revealed among learners in Verspoor et al. (2021).

The two studies which focused on the age of second language learners, in different settings, did not find congruent findings. Among the 71 examined participants, Pfenninger (2022) observed that the ones who started learning English at the age of seven gained higher degrees of proficiency with different second language trajectories in comparison with the participants who started learning when they were nine years of age. Nevertheless, the study conducted by Kliesch and Pfenninger (2021) which tested 28 participants over the age of 64 did not show any significant impact of age and even some older participants performed better than their younger peers.

Different directions of development were observed with respect to the language subsystems in the course of second language acquisition. The development of CAF subsystems were reported to have different and inconsistent directions, fluctuating between supportive and competitive at different stages (Evans & Larsen-Freeman, 2020; Spoelman & Verspoor, 2010; Yu & Lowie, 2020; Wind, 2021; Zhang et al., 2022). Zhang et al. (2022) observed that accuracy had trade-off effects with the other subsystems (complexity and fluency). The study conducted by Rokoszewska (2022) was indicative of negative associations between the development of the syntactic and lexical complexity, accuracy, and fluency, while supportive within-subsystem relationships were observed (e.g., subordination, coordination, and nominalization). The varied and flexible associations between the development of linguistic subsystems are indicative of the interconnectedness as a characteristics of complex and dynamic systems.

An interesting observation which was reported several times is that more variability and fluctuation in second language development seems to correlate with increased proficiency (Gui et al., 2021; Huang et al., 2021; Lesonen, 2021). The study conducted by Zhang et al. (2022) showed that among the three prototypes of learners, the prototype with constant variability in the development had more overall progress compared to the other prototypes. Among the two remaining ones, the participants with only initial variability had improved more than the participants with constant steady developmental patterns. The correlation between variability and linguistic ability gains was also observed in the only study that focused on reading in second language development (Gui et al., 2021). In addition, variability was positively related with the measures of aptitude, motivation, and exposure in Lowie and Verspoor (2019). Moreover, Kliesch and Pfenninger (2021), Khomeijani et al. (2020), Spoelman and Verspoor (2010), and Bulté and Housen (2018) observed that the association between variability and second language development was strongest at lower levels of proficiency and it seemed to wane when certain degrees of progress occurred. Lower degrees of variability was also negatively associated with second language development among the examined identical twins in Lowie et al. (2017). Even in the development of pronunciation subskills, Munro and Derwing (2008) observed that more changes occurred during the initial stages of participants' developmental patterns and the variability decreased as participants grew more proficient in their pronunciation.

## Future Research

While previous findings are incontrovertibly indicative of the fact that second language development is dynamic, the next steps that need to be taken seem to be finding potential regularities, establishing dynamic models for second language development, and testing for optimized learning methods and activities in accordance with potentially predictable developmental patterns. From the perspective of CDST, second language learning is unpredictable although not random (Larsen-Freeman & Cameron, 2008). However, considering other complex systems, it can be observed that the unpredictability and uncertainty of complex systems are more pronounced in the long term. Different tools and methods have been established to investigate complex systems in different realms such as biology (Karr et al., 2012), climate science (Lau & Ploshay, 2013), chemistry (Lewars, 2011), and physics (Holovatch et al., 2017). Comparably, complex systems have not been sufficiently examined in agent-based systems where autonomous decision-making agents like people

exist (An et al., 2021; Hilpert & Marchand, 2018; Schulze et al., 2017). It is a shortcoming which stems from the fact that traditional mathematical modeling is significantly more difficult for agent-based systems. A new line of research can be the pursuit of appropriate mathematical models and analyses for approximate prediction of patterns and trends of second language development. As an example of the symptoms that can potentially predict the short-term trend of development, Evans and Larsen-Freeman (2020) observed that phases of instability, characterized by increase in the flow disruptions and also production bimodality, were indicative of phase shifts in the developmental patterns of their participants. More studies can attempt to identify trends in the dynamic development of second languages, like Zhang et al. (2022) and Baba and Nitta (2021) where three prototypes for individual learning patterns were observed. Additionally, Gui et al. (2021) reported seven developmental patterns among the reading development of participants over time.

In addition to providing deeper theoretical understanding about the dynamic patterns of second language development, being able to roughly predict these patterns has numerous implications. For example, considering the ups and downs and phase shifts that a learner experiences in their development of second language speaking subskills, being able to predict the short-term developmental fluctuations can help them set their IELTS speaking exam at a peak of this pattern of speaking development. While CDST emphasizes the uniqueness of language acquisition for each learner (Lesonen, 2021), attempting to find regularities in second language development is of great importance (Dornyei, 2014; Ellis, 2007; Lowie & Verspoor, 2019; Pfenninger, 2022; Zhang, 2022). Bulté and Housen (2018) propose the utilization of true dynamic methods and adding mathematical models in future studies. Leveraging big data, future studies can benefit from computational linguistics and data science for further investigation of the dynamic development of second language. They can also open the doors for future studies focused on adopting targeted instructions and feedback which would optimize second language development of learners.

There is a number of other issues on which future research can shed more light. While some previous studies have revealed that more variability in second language development leads to higher degrees of proficiency (Kliesch & Pfenninger, 2021; Lesonen, 2021), future studies specifically focused on this relationship can show the significance of such a correlation. Munro and Derwing (2008), Kliesch and Pfenninger (2021), Spoelman and Verspoor (2010), and Khomeijani et al. (2020) observed that the association between variability and proficiency gains is stronger at lower proficiency levels. Future experiments can clarify the waning of such correlation at higher proficiency levels.

A specific question that can be pursued is the degree of dynamicity in the development of different linguistic subsystems. As an example, Kliesch and Pfenninger (2021) observed that the group developmental pattern of fluency bet-

ter presented the individual trajectories of the participants in comparison with that of lexical richness. Also, another subject which needs further investigation are the relationships and interactions between different language subsystems during second language development (especially syntactic and lexical complexity). Different associations with variable directions have been reported in the literature (Khomeijani et al., 2020; Rokoszewska, 2022; Spoelman & Verspoor, 2010; Yu & Lowie, 2020; Zhang et al., 2022). Examining the existence of such supportive or competitive associations between and within the subsystems can further clarify them. The rate and speed comparisons of development in CAF subsystems can be another issue requiring further examination. Verspoor et al. (2017) emphasized the need for different linguistic measures for different proficiency levels in order to increase the accuracy of developmental measurements.

Comparing dynamic second language development between different ages, different L1 fluency levels, and different levels of second language proficiency are other topics which have not received specific but sporadic attention in the literature. Since initial states of complex systems can strongly influence the long-term conditions of these systems (an impact also known as the butterfly effect), the aforementioned learner characteristics can considerably impact the progression and outcome of second language acquisition. In addition, a number of previous studies have mentioned task effect an extraneous variable the impact of which could not be controlled (Lesonen et al., 2021; Menke and Strawbridge, 2019; Vyatkina et al., 2015). Controlling for this factor can add to the accuracy of the observations in future research. The effects of targeted feedback can also be tested on the subsystems with slower developing subsystems (Rokoszewska, 2022). Another issue is the summer gaps that occurred during the data collection of studies which examined the development of students during academic years. Such gaps stop the continuous tracking of participants' developmental patterns and can be delimited in future studies.

While the acquisition and development of specific second language subsystems, especially complexity accuracy and fluency, has been repeatedly investigated, few studies have focused on development of other aspects of second language (such as the use of "can," authorial voice, and interrogatives). Future research can delve into the development of such other subsystems from the CDST viewpoint. Additionally, the development of listening and reading has not received sufficient attention since only two studies (Chang & Zhang, 2021; Gui et al., 2021) have examined second language development in receptive language skills. Future studies can alleviate the shortage in this realm.

## Conclusion

From the perspective of CDST, second language development is a unique and different process for every single person. The present systematic literature review has attempted to examine previous studies which examined second language development from CDST viewpoint. The main three goals of this review were presenting an overview of the methodological characteristics of the previous studies in the field, providing a synthesis of their findings, and identifying the gaps and areas which require further investigation in future studies. The searching strategy of this review led to finding 45 articles in the literature. Since CDST studies need keeping track of learners' linguistic development over longer periods of time in a detailed manner, longitudinal design and abundance of data points were the two important characteristics of the reviewed papers. Speaking and writing were the most investigated skills, while complexity, accuracy, and fluency were the most examined subsystems. The observations of all of the reviewed papers supported the CDST principles in second language development. The reports indicated that developmental patterns were non-linear and variable between and within participants. A number of repeated, but yet inconclusive correlations were observed. Increased fluctuations in developmental patterns of a subsystem were associated with more development of that subsystem. Moreover, developmental fluctuations seem to decrease as a learner becomes more proficient. While such issues require further examination, other important areas which require research are modeling the dynamic development of second languages and searching for potential regularities, prototypes, and trajectories.

## References

- An, L., Grimm, V., Sullivan, A., Turner II, B., Malleson, N., & Heppenstall, A. et al. (2021). Challenges, tasks, and opportunities in modeling agent-based complex systems. *Ecological Modelling*, 457, 109685. <https://doi.org/10.1016/j.ecolmodel.2021.109685>
- Baba, K., & Nitta, R. (2014). Phase transitions in development of writing fluency from a complex dynamic systems perspective. *Language Learning*, 64(1), 1–35. <https://doi.org/10.1111/lang.12033>
- Baba, K., & Nitta, R. (2021). Emergence of multiple groups of learners with different writing-development trajectories in classroom: Growth mixture modeling. *Journal of Second Language Writing*, 54, 100856.
- Bak, T. H., Long, M. R., Vega-Mendoza, M., & Sorace, A. (2016). Novelty, challenge, and practice: The impact of intensive language learning on attentional functions. *Plos One*, 11, e0153485.

- Bastardas-Boada, A. (2013). Language policy and planning as an interdisciplinary field: Towards a complexity approach. *Current Issues in Language Planning*, 14(3–4), 363–381.
- Berggren, R., Nilsson, J., Brehmer, Y., Schmiedek, F., & Lövdén, M. (2020). Foreign language learning in older age does not improve memory or intelligence: Evidence from a randomized controlled study. *Psychology and Aging*, 35(2), 212–219.
- Blommaert, J. (2014). From mobility to complexity in sociolinguistic theory and method. *Tilburg Papers in Culture Studies*, 103, 1–24.
- de Bot, K., & D. Larsen-Freeman (2011). Researching second language development from a dynamic systems theory perspective. In M. H. Verspoor et al. (Eds.), *A dynamic approach to second language development* (pp. 5–23). Benjamins.
- de Bot, K., W. Lowie, & M. Verspoor (2007). A Dynamic Systems Theory approach to second language acquisition. *Bilingualism: Language and Cognition*, 10(1), 7–21.
- Bulté, B., & Housen, A. (2018). Syntactic complexity in L2 writing: Individual pathways and emerging group trends. *International Journal of Applied Linguistics*, 28(1), 147–164.
- Casillas, J. V. (2020). The longitudinal development of fine-phonetic detail: Stop production in a domestic immersion program. *Language Learning*, 70, 768–806. <https://doi.org/10.1111/lang.12392>
- Caspi, T., & Lowie, W. M. (2013). The dynamics of L2 vocabulary development: A case study of receptive and productive knowledge. *Revista Brasileira de Linguística*, 13(2), 437–462.
- Chan, H., Verspoor, M., & Vahtrick, L. (2015). Dynamic development in speaking versus writing in identical twins. *Language Learning*, 65, 298–325.
- Chang, P., & Zhang, L. J. (2021). A CDST perspective on variability in foreign language learners' listening development. *Frontiers in Psychology*, 12(601962), 1e14. <https://doi.org/10.3389/fpsyg.2021.601962>
- Dörnyei, Z. (2014). Researching complex dynamic systems: 'Retrodictive qualitative modelling' in the language classroom. *Language Teaching*, 47, 80–91.
- Ellis, N. C. (2007). Dynamic systems and SLA: The wood and the trees. *Bilingualism: Language and Cognition*, 10(1), 23–25.
- Eskildsen, S. W. (2009). Constructing another Language—Usage-Based Linguistics in Second Language Acquisition. *Applied Linguistics*, 30(3), 335–357. <http://doi.org/10.1093/applin/arn037>
- Eskildsen, S. W. (2015). What counts as a developmental sequence? Exemplar-based L2 learning of English questions. *Language Learning*, 65(1), 33–62.
- Evans, D. R., & Larsen-Freeman, D. (2020). Bifurcations and the emergence of L2 syntactic structures in a complex dynamic system. *Front. Psychol.* 11, 574603. <http://doi.org/10.3389/fpsyg.2020.574603>
- Fogal, G. G. (2020). Investigating variability in L2 development: Extending a complexity theory perspective on L2 writing studies and authorial voice. *Applied Linguistics*, 41(4), 575–600. <http://doi.org/10.1093/applin/amz005>
- Fogal, G. G., & Koyama, D. (2022). A study of co-adaptation through journaling. *Journal of Second Language Writing*, 55, 100873.
- Gui, M., Chen, X., & Verspoor, M. (2021). The dynamics of reading development in L2 English for academic purposes. *System*, 100, 102546. <http://doi.org/10.1016/j.system.2021.102546>
- Han, J., & Hiver, P. (2018). Genre-based L2 writing instruction and writing-specific psychological factors: The dynamics of change. *Journal of Second Language Writing*, 40, 44–59.
- Han, Z. (2021). Usage-based instruction, systems thinking, and the role of Language Mining in second language development. *Language Teaching*, 54(4), 502–517. <https://doi.org/10.1017/S0261444820000282>

- Hanania, E. A., & Gradman, H. L. (1977). Acquisition of English structures: A case study of an adult native speaker of Arabic in an English-speaking environment. *Language Learning*, 27, 75–91. <https://doi.org/10.1111/j.1467-1770.1977.tb00293.x>
- Hilpert, J. C., & Marchand, G. C. (2018). Complex systems research in educational psychology: Aligning theory and method. *Educational Psychologist*, 53(3), 185–202.
- Hiver, P., & Al-Hoorie, A. H. (2016). A dynamic ensemble for second language research: Putting complexity theory into practice. *The Modern Language Journal*, 100(4), 741–756.
- Hiver, P., Al-Hoorie, A. H., & Evans, R. (2022). Complex dynamic systems theory in language learning: A scoping review of 25 years of research. *Studies in Second Language Acquisition*, 44(4), 913–941.
- Holovatch, Y., Kenna, R., Thurner, S. (2017). Complex systems: Physics beyond physics. *European Journal of Physics*, 38, 023002. <https://doi.org/10.1088/1361-6404/aa5a87>
- Huang, T., Steinkrauss, R., & Verspoor, M. (2022). Learning an L2 and L3 at the same time: Help or hinder?. *International Journal of Multilingualism*, 19(4), 566–582. <https://doi.org/10.1080/14790718.2020.1779726>
- Hult, F. (2010). The complexity turn in educational linguistics. *Language, Culture and Curriculum*, 23, 173–177.
- Karr, J. R., Sanghvi, J. C., Macklin, D. N., Gutschow, M. V., Jacobs, J. M., Jr. Bolival, B., Assad-Garcia, N., Glass, J. I., Covert, M. W. (2012). A whole-cell computational model predicts phenotype from genotype. *Cell*, 150, 389–401. <https://doi.org/10.1016/j.cell.2012.05.044>
- Khomejani Farahani, A. A., Rezaee, A. A., & Moshtaghi Zonouz, R. (2020). Exploring the development of writing complexity, accuracy, and fluency in relation to the motivational trajectories: a dynamically-oriented case study. *English Teaching & Learning*, 44(1), 81–100.
- Kliesch, M., & Pfenninger, S. E. (2021). Cognitive and socioaffective predictors of L2 micro-development in late adulthood: A longitudinal intervention study. *The Modern Language Journal*, 105(1), 237–266. <https://doi.org/10.1111/modl.12696>
- Kowal, I. (2014). Fluency in second language writing: A developmental perspective. *Studia Linguistica Universitatis Jagellonicae Cracoviensis*, 131(2014), 229–246. <https://doi.org/10.4467/20834624SL.14.013.2321>
- Kramsch, C., & Whiteside, A. (2008). Language ecology in multilingual settings. Towards a theory of symbolic competence. *Applied Linguistics*, 29(4), 645–671.
- Larsen-Freeman, D. (1994). On the parallels between chaos theory and second language acquisition [Paper presentation]. Second Language Research Forum, McGill University, Montreal.
- Larsen-Freeman, D. (1997). Chaos/complexity science and second language acquisition. *Applied Linguistics*, 18(2), 141–165. <https://doi.org/10.1093/applin/18.2.141>
- 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, 590–619. <https://doi.org/10.1093/applin/aml029>
- Larsen-Freeman, D. (2017). Complexity theory: The lessons continue. In L. Ortega & Z. Han (Eds.), *Complexity theory and language development* (pp. 11–50). John Benjamins.
- Larsen-Freeman, D. (2018). Resonances: Second language development and language planning and policy from a complexity theory perspective. In M. Siiner, F. M. Hult, & T. Kupisch (Eds.), *Language policy and language acquisition planning* (pp. 203–217). Springer.
- Larsen-Freeman, D., & Cameron, L. (2008). *Complex systems and applied linguistics*. Oxford University Press.
- Lau, N.-C., & Ploshay, J. J. (2013). Model projections of the changes in atmospheric circulation and surface climate over North America, the North Atlantic, and Europe in the twenty-first century. *J. Clim.* 26, 9603–9620. <https://doi.org/10.1175/JCLI13-00151.1>

- Lesonen, S., Steinkrauss, R., Suni, M., & Verspoor, M. (2021). Dynamic usage-based principles in the development of L2 Finnish evaluative constructions. *Applied Linguistics*, 42(3), 442–472.
- Lewars, E. G. (2011). *Computational chemistry: Introduction to the theory and applications of molecular and quantum mechanics*. Springer.
- Li, J., & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18(2), 85–102.
- Linck, J. A., Osthus, P., Koeth, J. T., & Bunting, M. F. (2014). Working memory and second language comprehension and production: A meta-analysis. *Psychonomic Bulletin & Review*, 21(4), 861–883.
- Lowie, W. M., Caspi, T., Van Geert, P., & Steenbeek, H. (2011). Modeling development and change. In M. H. Verspoor, K. De Bot, & W. Lowie (Eds.), *A dynamic approach to second language development: Methods and techniques* (pp. 22–122). Benjamins.
- Lowie, W. M., Van Dijk, M., Chan, H. P., & Verspoor, M. H. (2017). Finding the key to successful L2 learning in groups and individuals. *Journal of Language Teaching and Learning*, 7, 127–148. <https://doi.org/10.14746/ssl1t.2017.7.1.7>
- Lowie, W. M., & Verspoor, M. H. (2019). Individual differences and the ergodicity problem. *Language Learning*, 69, 184–206. <https://doi.org/10.1111/lang.12324>
- Lowie, W. M., Verspoor, M. H., & de Bot, K. (2010). A dynamic view of second language development across the lifespan. In K. de Bot & R. W. Schrauf (Eds.), *Language development over the lifespan* (pp. 125–146). Routledge.
- Mellow, J. D. (2008). The emergence of complex syntax: A longitudinal case study of the ESL development of dependency resolution. *Lingua*, 118(4), 499–521.
- Menke, M. R., & Strawbridge, T. (2019). The writing of Spanish majors: A longitudinal analysis of syntactic complexity. *Journal of Second Language Writing*, 46, 1–15. <https://doi.org/10.1016/j.jslw.2019.100665>
- Mercer, S. (2011). Language learner self-concept: Complexity, continuity and change. *System*, 39(3), 335–346.
- Mufwene, S. S., Coupé, C., & Pellegrino, F. (Eds.) (2017). *Complexity in language: Developmental and evolutionary perspectives*. Cambridge University Press.
- Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning*, 58(3), 479–502.
- Nitta, R., & Baba, K. (2015). Self-regulation in the evolution of the ideal L2 self: A complex dynamic systems approach to the L2 motivational self-system. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational dynamics in language learning* (pp. 374–404). Multilingual Matters.
- Nitta, R., & Baba, K. (2018). Understanding benefits of repetition from a complex dynamic systems perspective. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 279–309). John Benjamins Publishing Company.
- Piniel, K., & Csizér, K. (2015). Changes in motivation, anxiety and self-efficacy during the course of an academic writing seminar. *Motivational dynamics in language learning*, 81, 164–194. <https://doi.org/10.21832/9781783092574-015>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., et al. (2021). The prisma 2020 statement: An updated guideline for reporting systematic reviews. *BMJ* 372:n71. <https://doi.org/10.31222/osf.io/v7gm2>
- Pfenninger, S. E. (2022). Emergent bilinguals in a digital world: A dynamic analysis of long-term L2 development in (pre)primary school children. *International Review of Applied Linguistics in Language Teaching*, 60(1), 41–66.

- Polat, B., & Kim, Y. (2014). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics*, 35(2), 184–207.
- Ramos, S., García, Y. F., Antón, E., Casaponsa, A., & Duñabeitia, J. A. (2017). Does learning a language in the elderly enhance switching ability?. *Journal of Neurolinguistics*, 43, 39–48.
- Rokoszewska, K. J. (2020). Intra-individual Variability in the Emergence of Lexical Complexity in Speaking English at Secondary School. *Theory and Practice of Second Language Acquisition*, 6(1), 107–142.
- Rokoszewska, K. (2022). The dynamics of monthly growth rates in the emergence of complexity, accuracy, and fluency in L2 English writing at secondary school—a learner corpus analysis. *System*, 106, 102775. <https://doi.org/10.1016/j.system.2022.102775>
- Rosmawati, R. (2014). Dynamic development of complexity and accuracy: a case study in second language academic writing. *Australian Review of Applied Linguistics*, 37(2), 75–100. <https://doi.org/10.1075/ara1.37.2.01ros>
- Routarinne, S., & Ahlholm, M. (2021). Developing requests in multilingual classroom interaction: A case of second language development in middle childhood. *Applied Linguistics*, 42(4), 765–790. <https://doi.org/10.1093/applin/amaa048>
- Sauer, L., & Ellis, R. (2019). The social lives of adolescent study abroad learners and their L2 development. *The Modern Language Journal*, 103(4), 739–762. <https://doi.org/10.1111/modl.12589>
- Schlüter, M., Baeza, A., Dressler, G., Frank, K., Groeneveld, J., Jager, W., Janssen, M. A., McAllister, R. R. J., Müller, B., Orach, K., Schwarz, N., Wijermans, N. (2017). A framework for mapping and comparing behavioural theories in models of socioecological systems. *Ecol. Econ.* 131, 21–35.
- Serafini, E. J. (2017). Exploring the dynamic long-term interaction between cognitive and psychosocial resources in adult second language development at varying proficiency. *The Modern Language Journal*, 101(2), 369–390.
- Shanker, S. G., & King, B. J. (2002). The emergence of a new paradigm in ape language research. *Behavioral and Brain sciences*, 25(5), 605–620.
- Skiba, R., & Dittmar, N. (1992). Pragmatic, semantic, and syntactic constraints and grammaticalization: A longitudinal perspective. *Studies in Second Language Acquisition*, 14(3), 323–349.
- Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics*, 31, 532–553. <https://doi.org/10.1093/applin/amq001>
- Serrano, R., & Howard, E. (2007). Second language writing development in English and in Spanish in a two-way immersion programme. *International Journal of Bilingual Education and Bilingualism*, 10(2), 152–170.
- Van Dijk, M., Verspoor, M. H., & Lowie, W. M. (2011). Variability analyses in language development. In *A dynamic approach to second language development: Methods and techniques* (pp. 55–84). John Benjamins Publishers.
- Van Geert, P., & van Dijk, M. (2002). Focus on variability: New tools to study intra-individual variability in developmental data. *Infant Behavior & Development*, 25, 340–374.
- Verspoor, M., Lowie, W., Chan, H., & Vahtrick, L. (2017). Linguistic complexity in second language development: Variability and variation at advanced stages. *Recherches en didactique des langues et des culture. Les Cahiers de l'Acedle*, 14(1), 1–27. <https://doi.org/10.4000/rdlc.1450>
- Verspoor, M., Lowie, W., & Wieling, M. (2021). L2 developmental measures from a dynamic perspective. In B. Le Bruyn & M. Paquot (Eds.), *Learner corpus research meets second language acquisition* (pp. 172–190). Cambridge University Press.

- Verspoor, M., Lowie, W., & Van Dijk, M. (2008). Variability in second language development from a dynamic systems perspective. *Modern Language Journal*, 92, 214–231. [https://doi:10.1111/j.1540-4781.2008.00715.x](https://doi.org/10.1111/j.1540-4781.2008.00715.x)
- Verspoor, M. H., de Bot, K., & Lowie, W. (Eds.) (2011). *A dynamic approach to second language development: Methods and techniques*. John Benjamins.
- Verspoor, M., & van Dijk, M. (2011). Visualizing interaction between variables. In M. H. Verspoor, K. de Bot, & W. Lowie (Eds.), *A Dynamic Approach to Second Language Development* (pp. 85–98). John Benjamins Publishing Company.
- Verspoor, M. & van Dijk, M. (2012). Variability in a dynamic systems theory approach to second language acquisition. In C. Chapel (Ed.), *The Wiley- Blackwell Encyclopedia of Applied Linguistics*. Blackwell Publishing.
- Vyatkina, N. (2012). The development of second language writing complexity in groups and individuals: A longitudinal learner corpus study. *The Modern Language Journal*, 96(4), 576–598. <https://doi.org/10.1111/j.1540-4781.2012.01401.x>
- Vyatkina, N., Hirschmann, H., & Golcher, F. (2015). Syntactic modification at early stages of L2 German writing development: A longitudinal learner corpus study. *Journal of Second Language Writing*, 29, 28–50.
- Wind, A. M. (2021). Nonlinearity and inter-and intra-individual variability in the extent of engagement in self-reflection and its role in second language writing: A multiple-case study. *System*, 103, 102672. <https://doi.org/10.1016/j.system.2021.102672>
- Wong, P. C., Ou, J., Pang, C. W., Zhang, L., Tse, C. S., Lam, L. C., & Antoniou, M. (2019). Language training leads to global cognitive improvement in older adults: A preliminary study. *Journal of Speech, Language, and Hearing Research*, 62(7), 2411–2424.
- Yu, H., & Lowie, W. (2020). Dynamic paths of complexity and accuracy in second language speech: A longitudinal case study of Chinese learners. *Applied Linguistics*, 41(6), 855–877. <https://doi.org/10.1093/applin/amz040>
- Zhang, Y. (2004). Processing constraints, categorial analysis, and the second language acquisition of the Chinese adjective suffix-de (ADJ). *Language Learning*, 54(3), 437–468.
- Zhang, J., & Lu, X. (2013). Variability in Chinese as a foreign language learners' development of the Chinese numeral classifier system. *The Modern Language Journal*, 97(S1), 46–60.
- Zhang, S., Zhang, H., & Zhang, C. (2022). A Dynamic Systems study on complexity, accuracy, and fluency in English writing development by Chinese university students. *Frontiers in Psychology*, 13, 787710.
- Zheng, Y. (2016). The complex, dynamic development of L2 lexical use: A longitudinal study on Chinese learners of English. *System*, 56, 40–53.

Soheil Behdarvandirad

## **Systematische Übersicht des Zweitspracherwerbs aus der Perspektive der Theorie komplexer dynamischer Systeme**

### **Zusammenfassung**

Aus der Sicht der Theorie komplexer dynamischer Systeme (CDST) weist die Zweitsprachentwicklung unvorhersehbare und nicht-lineare Muster auf, die von Lerner zu Lerner variieren können. Um diese dynamische Entwicklung nachzuvollziehen, sind Längsschnittstudien mit einer entsprechenden Anzahl von Datenpunkten erforderlich. Die

vorliegende systematische Literaturübersicht versucht, einen Überblick über die bisher durchgeführten Längsschnittstudien zu geben, welche die Entwicklung der zweitsprachlichen Teilsysteme im Hinblick auf die CDST-Theorie verfolgt haben. Ausgehend von Veröffentlichungen aus dem Jahre 1884 werden im Rahmen der systematischen Suchstrategie insgesamt 45 Artikel einer Analyse unterzogen, um den aktuellen Stand der Forschung darzulegen. Die untersuchten Studien sprechen eindeutig für die Umsetzung der CDST-Prinzipien in der Zweitsprachentwicklung. Es wird eine Synthese der Ergebnisse von Arbeiten vorgestellt und abschließend eine Vielzahl von Vorschlägen für weitere Forschung gegeben, die zukünftigen Studien helfen können, die bestehenden Lücken in der Literatur zu klären.

*Schlüsselwörter:* Zweitsprachentwicklung, Zweitspracherwerb, CDST, systematische Übersicht





**Robert Oliwa**

State Univeristy of Applied Sciences in Przemysł, Poland

 <https://orcid.org/0000-0002-5251-1086>

## Contextualising the Hyflex Model of Instruction for Language Classes

### Abstract

The paper discusses a possible, practical application of the hybrid flexible (*hyflex*) model of content delivery in a tertiary context. The model may as well be applied in foreign language instruction. It is the author's belief that there are convincing reasons for employing digital competences and skills acquired over the last two years to enhance teaching and learning. An appropriate contextualisation of the hyflex model may further encourage the retention of the already attained competences and skills in the new normal. It may also allow to combine the traditional, in-class instruction with the new, online interaction both synchronously and asynchronously. The possibility of hyflexing language courses can be a challenging, however responsible, prospect because the restoration of the education we used to know might not be easy. Although theoretical, the paper provides practical guidelines as well as offers a selection of online tools to contextualise the hyflex model in language teaching. The paper may also constitute a strong claim for the practical application of the hyflex model especially when combined with an in-depth analysis and research of the model in various educational contexts. As the future is hard to envisage, various models and approaches might be taken into account so as to appropriately respond to the challenges and requirements education is likely to be confronted with.

*Keywords:* hybrid, hyflex, platform, online, language, instruction

*That's one small step for [a] man, one giant leap for mankind.*  
Neil Armstrong, 20 July 1969.

The contemporary rewording of the famous words pronounced on July 20, 1969, by the American astronaut Neil Armstrong when he put his left foot on the lunar surface if we wish to address what the world of education has experienced might be *those were many giant steps for a man and one giant leap*

*for education.* A brief history of education may suggest that it has developed from a stage when the content was presented in a fashion the teacher decided and found appropriate to a stage when the content is available anytime and anywhere in a fashion students choose, remotely accessing resources they enjoy having a continuous access to teachers who flexibly and timely adjust the instructional process. Of course, prior to the formalised education, it was associated with pleasure (the original meaning of the word from Greek *skholē* is spare time, leisure, rest, that in which leisure is employed). Only later it came to be used for the place for learned discussion, also for lectures and school. Besides, the aim of education has developed too, starting from inflicting on students arduous memorisation chores, through equipping them necessary with competences and most recently to encouraging autonomy, self-education and discovery which we are witnessing now. The question which arises at this point in the evolution of education, after the most recent misfortunes, is how it will develop bearing in mind that it is in a constant state of flux. The contemporary education enables interactive content to be shared, personalised or augmented. Moreover, the array of available apps and environments allow the flexibility and customisation of the learning experience. Materials are available on mobile devices as WiFi technology became omnipresent and the learning process can take place anytime/anywhere. On the other hand, there is the issue of information quality, information which is plentiful and whose selection may be a vital skill. The contemporary education is also an area where new teaching methods, techniques, and tools are coming into use. These include direct instruction, flipped classroom, gamification, mind mapping, inquiry-based learning, webquest, project-based learning, VAK teaching, problem-based learning and most recently hybrid teaching. Education is becoming affordable with a myriad of massive open online courses which revolutionise the way students acquire skills and competences which the labour market requires. A synergy of the traditional with the modern along with their reciprocal interaction might be an interesting offer the coming years will bring. One way of dealing with the challenges may be the hyflex model which the following part aims to explain.

## **Background**

As shown by Bonk and Graham (2006) as well as Lockman and Schirmer (2020), the concept of hybrid teaching has been used interchangeably with distance, blended or online learning and denoted a combination of in-person instruction augmented with electronic, computer-based or online resources. As such, it did not take into consideration the concurrent delivery of the same content in

a traditional, online and asynchronous environments to different receptionists who may interact with each other and may retrieve the instructional content anytime and anywhere. There have been many attempts, such as multi-access learning, blendflex, hyflex course design, flexlearning or remote live participation to devise an approach which would combine in-person and distance, synchronous and asynchronous instruction. This is what the hyflex instruction has recently attempted to contribute. The hyflex course design was first introduced in 2005 at San Francisco State University as a reaction to the declining student enrolment. The model was implemented with a focus on the flexibility of student access but the advantages may also affect teachers who join classes remotely while students attend classes in-person from a room at home or on campus. The design was described as “[...] class sessions that allow students to choose whether to attend classes face-to-face or online, synchronously or asynchronously” (SFSU Academic Senate Policy S19–264, 2010) which results in “a student-directed, multi-modal learning experience” (Beatty, 2019, online). Beatty (2019) delineated four fundamental characteristics of HyFlex courses:

**Learner choice**—students are allowed to choose the mode of engagement (in-person, online synchronous, online asynchronous) that works best for them.

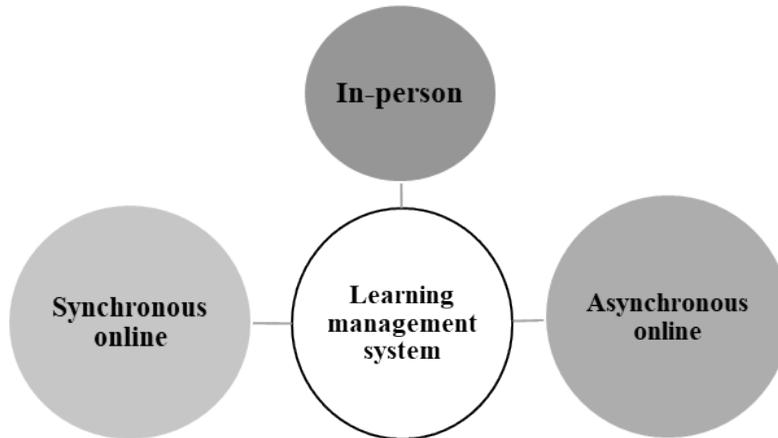
**Equivalence**—students are expected to reflect, contribute and interact with their peers in the process of learning and achieve equivalent learning outcomes.

**Reusability**—students have access to the same learning resources and the recorded output of students’ in-class activities is available online and may be used again by other students in other groups.

**Accessibility**—students are granted access to similar technology as well as content delivery channels and they have similar IT literacy.

All classes and learning tasks are offered in three modes, namely in-person, synchronous online, and asynchronous online (see Figure 1) and students themselves decide about how they want to take part. Central to this model is software application or web-based technology used to create and deliver learning content, monitor student participation and assess student performance. It may also offer students interactive features such as video conferencing and threaded discussions. Such a setup can encourage students’ liability for the process of learning and fashion it the way their needs, requirements, and differences permit. Accordingly, this may help them relate the learning style to their social and professional live settings and boost their motivation both intrinsically and instrumentally.

Despite its apparent resemblance and shared characteristics, there are certain dissimilarities that make the hyflex model stand out from the existing hybrid models of instruction. First of all, the hyflex model emphasises the concurrence of the learning modes on the student side and the delivery modes on the teacher side. Next, the delivery modes complement one another and depend on a learner’s, not the teacher’s choice (see Figure 1).

**Figure 1***Hyflex Model Setup*

Finally, it incorporates a more flexible study framework as far as attendance and task accomplishment are concerned (Parra & Abdelmalak, 2016). The above considerations as well as the presentation of the hyflex model in the available literature will be examined in the following part.

## Review of Literature

The review of the available research literature indicates varied, contradicting, outcomes as it often covers overlapping phenomena, including distance, mixed-mode, blended or online learning (Bonk & Graham, 2006; Oliver & Trigwell, 2005; Lockman & Schirmer, 2020). Similarly, the hyflex instruction cannot be analysed in a complete isolation from the other methods and approaches which either preceded it or occurred concurrently. First and foremost, the research indicates (Robertson & Kelly 2013; Alexander et al., 2014; He, 2015; Detienne et al., 2018; Beatty, 2019; Raes et al., 2020) that various kinds of hybrid, computer assisted learning have been implemented at the tertiary education level both before, during, and after the recent pandemic. The outcomes confirm either advantages or disadvantages of the process. The former include increased student satisfaction (Heilporn & Lakhali, 2021), time and place flexibility, immediate feedback, reference to the previous learning experience and cost efficiency (Smith & Hardaker, 2000; Dhawan, 2020; Singh & Matthees, 2021). The research carried out by Serdyukov & Hill (2013) confirmed student autonomy, engagement as well as better understanding of a course structure

and improvement of computer literacy. The latter include technical issues, such as the digital divide, students' and teachers' computer literacy and social skills (Dhawan, 2020; Shek et al., 2022). The limitation, however, which the available research has relates to little attention paid to appropriate scientific standards and quality assurance (Tallent-Runnels et al., 2006; Martin et al. 2020). Both in-person and online learning methods appeal to different types of learners and intelligence types. As the evidence from Bernard et al. (2014), Pinchot and Pullet (2014) and Idrizi et al. (2019) indicates, online channels of instruction attract learners who appreciate flexibility and autonomy and the onsite ones—those who enjoy structure and traditional social interaction. Finally, examination of the relationship between student satisfaction with online learning and intelligence types indicates an increase in satisfaction as a result of both interpersonal and visual-spatial intelligence types (Dziuban, 2015).

According to Klimova and Kacetla (2015) as well as Ramalingam et al. (2022), amidst the publications on the relationship between hybrid learning and the teaching of foreign languages there have been two main trends. The first one explains the hybrid approach and offers instruction manuals for those who wish to apply blended learning. The other one has focused on investigation and research discussing the outcomes of hybridity in language learning. On top of these, there was an attempt by Ying et al. (2019), Harun and Hussin (2018), Annamalai and Kumar (2020) which discovered that the integration of the latest technologies (game-like elements, social media, mobile applications) strengthens learners' motivation and engagement. Moreover, hybrid learning increases students' interaction and arouses their enthusiasm. As it can be discerned from the above considerations, the available research on hybrid teaching and learning in language instruction acknowledges the outcomes discovered in other areas. These include a more dynamic, interactive (language), learning environment which allows for a flexible course accessibility and reflection. The online environment encourages student communication, autonomy, and accountability for the learning process (Reynard, 2007; McBride & Fägersten, 2008). On the downside, the challenges include the problem of live interaction or social deficits, the amount and quality of teacher support, digital divide, output quality and management of the teaching/learning process (Bonk & Graham, 2006; Riel et al., 2016; Cheung et al., 2021).

The existing examination of the hyflex model may also imply often contradictory results (Miller et al., 2013; He et al., 2015; Verrecchia & McGlinchey, 2021). These include no substantial differences between in-person and hyflex models of instruction as regards academic achievements, no adverse impact on student performance, the advantage of the traditional model over the hyflex one or an achievement gap between students taking hyflex courses versus those enrolled in face-to-face classes. Nevertheless, as suggested by Parra and Abdelmalak (2016), students perceive the hyflex learning setting as helpful

and value the fact that it satisfies their needs, which was also acknowledged by Beatty (2009, 2012) and Gobeil-Proulx (2019). Furthermore, as revealed by Love (2015), hyflex courses support a broader range of learning types as it was discussed for the hybrid learning environment. Last but not least, on account of the autonomy the hyflex model offers, students' preferences are often unpredictable and can change over the course of studies.

From a teacher's perspective, the hyflex model can provide flexibility in delivering instruction and meeting the needs of all students, including those who may not be able to attend class in-person due to health or other reasons. However, it can also be challenging to manage and coordinate instruction for both in-person and online students at the same time, and may require additional planning and preparation. Additionally, teachers may need to adapt their teaching style and methods to effectively engage and connect with students in both settings. Overall, the hyflex model of instruction can present both advantages and challenges. These include:

Advantages:

- **Flexibility:** The hyflex model allows teachers to deliver instruction both in-person and online, which can provide more options for students who may not be able to attend in-person classes due to health or other reasons (Raes et al., 2019).
- **Meeting diverse needs:** The blended approach of hyflex model can cater to different learning styles and levels of proficiency among students.
- **Personalized learning:** Online resources and tools can allow for self-paced learning, which can benefit students at different levels of proficiency (Malczyk & Mollenkopf, 2019).
- **Authentic materials:** Online resources can provide access to authentic materials, such as videos, news articles, and podcasts, which can help language learners improve their listening and reading skills.

Challenges:

- **Coordination and management:** The hyflex model requires teachers to manage and coordinate instruction for both in-person and online students at the same time, which can be challenging and time-consuming (Leijon & Lundgren 2021).
- **Adapting teaching style:** The hyflex model may require teachers to adapt their teaching style and methods to effectively engage and connect with students in both in-person and online settings.
- **Technology:** Teachers may need to learn new technologies and tools for online instruction, which can be difficult and time-consuming (Zydney et al., 2019).
- **Student engagement:** Online learning can be isolating for some students, which may lead to less engagement and participation.
- **Teacher workload:** The hyflex model may require teachers to work longer hours and put in more effort, which can lead to burnout.

- **Equity:** Not all students may have equal access to technology or internet connectivity, which could lead to disparities in learning opportunities for online students (Koskinen, 2018).
- **Privacy:** Online learning raises concerns about student privacy, particularly with the collection and use of personal data.
- **Accessibility:** Some students with disabilities may face barriers to accessing online instruction and materials, and may require additional accommodations.
- **Student-teacher interaction:** The lack of direct interaction between teachers and students may impact the student's ability to understand the material, ask questions and receive feedback.

The research on the hyflex model in language teaching and learning, although very scant or in the early, nascent period of development, may indicate that the hyflex model of instruction can play a valuable role in language teaching and learning by providing flexibility and catering to the diverse needs of students. It offers:

- **Flexibility:** The hyflex model can allow students to attend class in person or online, which can be especially useful for language learners who may have different scheduling needs or face barriers to attending in-person classes (Abdelmalak & Parra, 2016).
- **Meeting diverse needs:** The hyflex model allows for a blended approach to teaching, which can be beneficial for students with different learning styles and levels of proficiency in the target language (Qayyum & Zawacki-Richter, 2018).
- **Customised learning:** The model can offer opportunities for a personalised learning pace, which can be valuable for different levels of language command (Malczyk & Mollenkopf, 2019).
- **Collaborative learning:** The hyflex model can facilitate collaborative learning through online discussions, group projects, and virtual language exchange programs, which can help language learners improve their communication skills (Miller et al., 2013).

An in-depth analysis of the Polish context including the study regulations at universities reveals that in the majority of higher education institutions, practicals, tutorials, seminars, foreign language courses, as well as laboratory and field classes are held on-campus (Adam Mickiewicz University, Jagiellonian University, Warsaw University of Technology, Catholic University of Lublin, Maria Curie-Skłodowska University, Warsaw University of Life Sciences). Precise indications as to which classes can also be held remotely are communicated by the deans of faculties and heads of organizational units after conducting surveys among students and academics (University of Warsaw). The Cracow University of Economics decided immediately to apply the hybrid model. Lectures are to be conducted only online. Other classes, that is, practicals, seminars, laboratory classes, and language courses will be conducted in

the stationary mode, unless the epidemiological situation forces the authorities to completely switch to remote mode. The most comprehensive attempt was made by University of Social Sciences and Humanities. It applied the innovative hyflex method of classes during which even several cameras simultaneously broadcast live lectures, practicals or workshops conducted in a university room with students physically on site. This allows students to partake either stationary or remotely at the same time depending on the requirements and circumstances. Moreover, an interesting initiative from University of Social Sciences and Humanities which obtained a prestigious grant from the National Agency for International Exchange to expand the English-language educational offer with International Intensive Education Programs. During two summer semesters, lecturers from foreign universities will conduct online and hyflex classes in the field of psychology, management, and English. The grant will cover the costs of conducting courses and equipping the classroom with tools for conducting hyflex classes on the LMS platform. The project will also include training for academic teachers in the field of conducting classes as part of International Intensive Education Programs online and in the hyflex formula. Lower Silesian University announced that the hyflex model remains the visionary goal of the development of e-learning strategy. Flexible competences of people and organizations to implement the hyflex model will be achieved in 2030, which will thus become the starting point for the implementation of education at selected directions according to the hyflex methodology. West Pomeranian Business School plans a further path of development and focus on the transformation of education and learning and the introduction of 5+ education model. The model is about many forms of synchronous and asynchronous learning: in the classroom, online in real time, hybrid learning (some in the classroom, some online), self-learning online and teaching hyflex, that is, in the classroom and online at the same time. It may be mandatory that the evaluation of the hyflex model in both foreign and Polish universities involves:

- student feedback (an insight into student experiences with the model, their level of engagement, satisfaction, and perceived learning outcomes);
- teacher feedback (an insight into the challenges and advantages, how they have adapted their teaching methods to accommodate both in-person and online students);
- learning outcomes (the model can be evaluated by comparing student learning outcomes in classes that use the model to those who do not, through assessments, quizzes, and exams);
- attendance and participation (monitoring student attendance and participation, both in-person and online, to get a sense of how well the model is working);
- technical and logistical issues (how well the technology, equipment, and logistics are functioning and if any adjustments need to be made).

Overall, the hyflex model of instruction can be a valuable tool for teachers but it also requires a lot of planning, preparation, and coordination. It is essential for schools and teachers to be aware of the ethical issues and work to address them in order to ensure that the hyflex model of instruction is equitable and accessible for all students. The model can provide a flexible and effective way to teach language, but it is important to keep in mind that it can also present some challenges, such as ensuring equity and accessibility for all students, and providing appropriate support for language learners. Finally, as the results indicate (Romero-Hall & Ripine, 2021) teachers are prepared to successfully engage in hyflex instruction that were significantly similar to skills required for in-person teaching. However, they confessed to being less prepared to cope with the characteristics exclusive to the hyflex model. Also, teachers believe various pedagogical strategies can be integrated into hyflex instruction. However, for those unskilled with this instructional modality, hands-on support and resources are required before planning and applying a course.

The limited body of available research on the hyflex model, also in the area of foreign language learning as well as its flawed implementation necessitate further examination and illustration of the medium to enhance a better understanding of how the method is constructed and what tools it entails. The following part of the paper will give some practical insight into the set up and technicalities of the hyflex model.

## **Practical Considerations**

A practical application of the hyflex model of instruction has to be founded on a current legal standing which in Poland was laid down by the Ministry of Education and Science in the Act of 20 July 2018 with its later amendments, including the COVID-related legal regulations. Moreover, the application process may also hinge on the type of the tertiary institution (general academic, vocational) as well as local conditioning. As already discussed, the hyflex model is based on three modes, namely in-person, synchronous online, and asynchronous. As Zajac (2020) suggests, when planning a hyflex course the reverse order may be put into place instead of planning a chronological order of stages a course comprises. First, the expected learning outcomes in the areas of knowledge, practical skills, and social competences are developed. Next, methods of the verification of the learning outcomes are planned. Finally, the means of content presentation are selected. Table 1 shows a sample design of an English language course.

**Table 1***Sample Design of an English Language Course*

<b>Stage 1</b>		
Areas	Access mode type	Sample online resources
<b>Learning outcomes</b>	<p><b>Knowledge</b> A student</p> <ul style="list-style-type: none"> <li>– knows vocabulary and grammatical structures at the required level;</li> <li>– knows basic vocabulary in the field of study.</li> </ul>	Asynchronous
	<p><b>Practical skills</b> A student can</p> <ul style="list-style-type: none"> <li>– understand functional texts and oral statements at various levels;</li> <li>– is able to construct oral and written statements in everyday and professional life situations;</li> <li>– independently acquire knowledge and develop their language skills using various sources.</li> </ul>	Synchronous in-person  Synchronous online
	<p><b>Social competences</b> A student is ready to</p> <ul style="list-style-type: none"> <li>– demonstrate the need to cooperate and work in a team, assuming various roles;</li> <li>– see the need to learn foreign languages to communicate in society and understand other cultures.</li> </ul>	
		Moodle* Google Classroom* Microsoft 365* Blackboard* Canvas Schoology Sakai
<b>Stage 2</b>		
Sample methods	Access mode type	Sample online resources
<b>Methods of verification</b>		EasyTestMaker
		Edpuzzle*
		Gimkit
		Google Forms
		Hot Potatoes*
		Kahoot
		Microsoft Forms
		Padlet
		Pear Deck
		Plickers
Quizlet		
Quizziz		
Testportal*		
Wordwall*		

Stage 3			
	Sample methods	Access mode type	Sample online resources
<b>Methods of content presentation</b>	Presentation	Asynchronous	Edpuzzle*
	Case Study		Mentimeter
	Project		Canva*
	Flipped Classroom		Animoto
	Gamification		ATutor
	Mind Mapping		BrainPOP
	Webquest		EdrawMind
		Synchronous in-person	ClassFlow
		Synchronous online	Coggle
			CourseSites
			EdApp
			Green Screen
			Mindmaster
			Nearpod
			Prezi

\*Note: Discussed in the following part.

The launch of the hyflex model and the subsequent running of its elements depend on the appropriate set up and interaction of three elements, namely a learning management system (LMS), progress assessment resources and content presentation tools. The three elements will be discussed in the following part.

## Learning Management Systems

The LMS supports the delivery of personalised e-learning content in the form of learning objects and the rules of using them. A growing practice is to refer to such environments as virtual learning environments as they facilitate interaction among its participants and contextualized materials in a virtual context. The last five years and the introduction of remote learning has created a shift to cloud-based environments, which is becoming the standard model for modern environments. The most popular LMSs (see Figure 2) include:

**Figure 2**

*Most Popular Learning Management Systems*



Such systems include the whole learning environments and not just individual applications or tools (e.g., MS Teams or Google Meet). Despite the fact that the decision to install a particular LMS may not be our managerial or financial capacity, its scope and functionalities are often similar and include the management, the type of instructional content, assessment, reporting, and (mobile) pedagogy it supports. Moreover, it may be applied to deliver any content including the language instruction.

## **Moodle**

Moodle offers an open, customisable, and multi-level structure which allows autonomy and administration of the platform to suit individual preferences as far as different courses, including the language ones, are concerned. The platform allows changing and allocating roles to course participants. It also permits the integration of different files, texts or documents, graphics and multimedia within or outside the environment. Moodle enables to offer online instruction and advanced instructional tools to present language content in a customisable manner to suit course participants depending on their needs. Moodle allows for integration of additional outside plug-ins when it does not offer a particular tool or functionality within the environment. The platform includes built-in modules for assessing content completion to provide for a variety of testing styles such as true/false, matching, multiple choices, multiselect, short answer or longer text. The assessment modules allow for giving feedback on student task accomplishment either quantitatively or qualitatively. Similar to the other LMSs, the environment supports social interaction as well as knowledge acquisition and construction through collaboration. On top of this, students and teachers can engage in synchronous and asynchronous communication such as email, forums, chat, discussion boards and, most importantly, social networking. Finally, regarding the asynchronous access to the platform, it supports recording video and audio directly within its interface. This can be achieved using a built-in recording tool or third party recording software to capture media of a particular course. As a result, a particular course can be accessed by the students who did not partake in the live classes or for a better course content retention.

## **Google Classroom / Microsoft 365**

Both are designed in a similar fashion to assist teaching and learning of any subject, including languages, by providing their own ecosystems. These include calendar, content sharing tools, live video streaming, group projects,

testing and grading tools as well as communication channels or cloud storage space. The most popular Google Classroom tools include: Docs, Drawings, Drive, Forms, Gmail, Jamboard, Maps, Meet, Sheets, Sites, Slides. The most popular Microsoft 365 tools include: Excel, Forms, Lists, OneDrive, OneNote, PowerPoint, Sway, Teams, Visio, Word.

Although the suites themselves are not considered stand-alone learning management systems, they regularly pilot and add new functionalities and applications which may lead to the fact that they begin to work and look LMSs. They resemble supermarkets which offer multitude of connected products or services to their users in one location despite the fact that their vendors are often different. From the managerial perspective, they allow the customisation of the content which is stored on a remote server in a cloud. This offers a multi-folder structure which is divided into app-related areas. Both systems permit the incorporation of different media types as well as linking to outside apps and resources available on the Internet. The modularity enables a teacher to revise and update available content. The acquisition of presented content can be evaluated with numerous assignment types whose results can be tracked and evaluated either automatically by the system or a teacher, which is visible to all the parties. Reporting on students' attainment depends on the internal or external tools and can take the form of digital annotations, video/audio comments, rubric assessment, and automatic quiz comments. The platforms offer collaboration and cooperation because students and teachers can engage in synchronous and asynchronous communication sharing the available content. Most importantly, both platforms are gaining power and influence because of their accessibility, affordability, ease of use and integration with their own and many other third-party apps and services for teachers, students, and admins alike. Regarding the asynchrony of the platform, all classes and meetings can be recorded and posted for students' later access. This is helpful when students are unable to take part or for a better retention of information.

### **Blackboard (and Other Commercial LMSs)**

It is a fully customisable multi-level course environment which allows turning on and off all aspects of the platform as well as its integration with different operating systems. It also permits linking its elements with external content and services. Functionalities which are not available on the platform can be accessed on-demand through the system support. Individual language settings are supported both on the instructor/teacher level as well as on the student level. All content can be structured into scalable folders, modules, and courses to cater for individual requirements and preferences. The setting includes a robust set of tools to deliver and organise instruction. The scope of the

system can be both its advantage and disadvantage especially for novice users. All types of instructional materials can be supplemented with different media types. The system includes a robust and powerful testing engine which offers a set of different quiz/exam types. Assessment includes customisable formats such as lists, graphs, rubrics, and written reports on students' progress. Both testing and assessment encourage the customisation of their formats including feedback, due dates, late submissions, and test access logs. Users have a social profile with a photo and information about themselves which they can share or hide in the security settings. Users can also communicate and post direct messages to students within the environment and externally. Finally, all major commercial LMSs support session recordings of real-time, virtual class with a student, a student group or entire class. A recorded computer screen, video, audio, a whiteboard, polls, or chat can be shared and accessed by students who missed a class due to illness or for review. Commercial learning management system like Blackboard whose pricing is not publicly available make the platform cost dependant on the number of licenses as well as their billing period.

### **Progress Assessment Resources**

The resources enable teachers to oversee and assess student progress with on-going formative and summative information regarding the outcome of instruction. They probably constitute the most subjective area because once chosen they stay one's favourite for years. In terms of the continuing changes, they may be the ones which come and go most often on the one hand. On the other, they are the most loved or hated resources teachers may put into place. The following online tools represent merely a drop in the ocean of options the Internet offers and because of this they are often a subject of modifications depending on local preferences. The most obvious choices with regard to progress monitoring resources include the built-in tools, for example, Microsoft and Google Forms or Moodle Docs and Test Creator which different LMSs offer. The presentation below will address stand-alone online application which also enable integration with the available LMSs.

### ***Hot Potatoes***

The old school, still free software first launched in 1998 offers a suite of five applications that can create online exercises. It is downloaded and set up on a PC or laptop and includes five applications: JCloze, JCross, JMatch, JMix, and JQuiz. The sixth application of the suite, The Masher combines different

Hot Potatoes exercises into one unit. As the very names suggest, Hot Potatoes suite enables the creation of fill-in-the-blank tasks (JCloze), jumbled-word exercises (JMix), crossword puzzles (JCross), matching exercises (JMatch), and text-entry quizzes (JQuiz). The simplicity of Hot Potatoes is both its blessing and its curse. Although the offered templates do not allow for a lot of customisation compared to the contemporary, interactive tools, Hot Potatoes is appreciated for designing and producing online, interactive language learning activities by the teachers who do not wish the “bells and whistles” to overshadow its instructional content or the assessment value or whose IT skills are still developing. Moreover, the comparison with the contemporary online tools delineated in the following parts indicates that the interaction is reduced to the interaction between the student and the automatic feedback generated by the programme. As such, Hot Potatoes activities may be viewed as interactive in the poorest sense. Nevertheless, it may be applied in any learning environment including the language learning one. Hot Potatoes’ user-friendly modules provide teachers with flexible, easy-to-use Web-based assessment tools which students can work on in class and at home obtaining the feedback that guides them towards answers. The continuing presence of the suite makes it one of the most popular tools available accompanied by numerous video tutorials, fora, and interest groups.

### ***Wordwall***

The online, game-based application is intended to help teachers create and draw from a variety of interactive and involving progress assessment resources. It offers three pricing plans including the free one which allows for the use of 18 types of interactive resources. It provides various templates for teachers to select from to enhance their classes and generate online assessment. The adaptable templates include multiple choice quizzes, flip tiles, group sort, match up, word search, brainstorm, rank order and crossword. However, the greatest asset of Wordwall is the fact that once an activity is created, it is shareable in different ways. The content from one template can be transferred to another by changing the format of a presentation. Moreover, the resources are open to use and modification by other platform users which can increase student interest and engagement while also providing a rapid feedback on performance. Wordwall may allow to set up a fun web-based (language) learning environment which can boost student engagement and interaction while providing both formative and summative feedback on their achievement.

## ***Testportal***

An online platform for assessing knowledge and competences in business and education. It offers both free and charge-based pricing plans. The free EDU plan is offered to the teachers of kindergartens, primary schools, secondary schools, and universities for education-related purposes only. Testportal includes a wide range of functionalities and can be used as a standalone online platform or as an LMS (e.g., Microsoft Teams) add-in app. The test design process is user-friendly and intuitive. It guides the user through following screens to configure different test settings, that is, question types, test kits, access method, start page, assessment, summary, and finally the time settings. The question manager allows an extensive configuration of all question types, from closed tasks, through multiple choice, open or true/false questions to surveys. It provides a highly advanced feedback panel to adjust detailed analysis of the results which the system will verify by itself. It also includes an option which prevents students from exiting the test to check their answers on the Internet. This solution may effectively discourage students from cheating. The undoubted advantages of Testportal is collecting students' results in the result database where they are stored for future reference. In language instruction Testplatform has a great potential regarding an effective both formative and summative assessment setting that produces positive pedagogical effects and increasing the engagement of students and content retention. The progress assessment tools presented above besides an abundance of other tools available online including: EasyTestMaker, Edpuzzle, Gimkit, Google Forms, Kahoot, Microsoft Forms, Padlet, Pear Deck, Plickers, Quizlet, Quizziz offer various types of tests, quizzes, and questionnaires that teachers can administer to evaluate students' progress in different areas including language learning. Online assessment tools enable teachers to create tests based on curriculum content or various teaching requirements. Furthermore, they allow tracking student progress against past tests for a better picture of individual students potential and to respond to a classroom setting. Together with the abovementioned LMSs and the content presentation tools described in the following part, they may complement the array of tools which streamline the instructional and assessment processes ensuring their reliability.

## **Content Presentation Resources**

The below resources constitute only a fraction of the limitless presentation tools the web offers. They can be used either as stand-alone tools or combined with one another to follow an order of content presentation.

### *Edpuzzle*

The basic and free pricing plan helps to get started with video presentations and allows a storage space for twenty video lessons. Edpuzzle includes integrations with the major LMSs which is of the utmost importance when embedding within an already functioning system. The integration includes connecting with an outside application and selecting the appropriate account when prompted. At first, an online video is selected or uploaded from a teacher's computer. Alternatively, a video lesson created by another teacher can be used. Then, the video is edited and augmented with own content. Besides, the application allows to record the voice to personalise the presented content. Finally, the video is assigned to your students and progress is monitored in real-time while students are learning at their own pace. Once a class is set up, students can be invited either by sharing the class code or importing students directly from, for example, Google Classroom. Assignments can be set individually or multiple videos can be assigned at once to single students or groups. Edpuzzle offers four different options for seeing students' progress which include student progress across multiple assignments, entire classroom progress for one assignment, individual student progress for one assignment, and grade questions. In language instruction Eduportal may be a tool for the flipped classroom as it facilitates students' interaction with the visual content at their own pace. Moreover, several versions of the same video can be created for different levels and ages. Instant formative assessments can be introduced along with new topics or revisions of the concepts already presented. Additionally, students can also create their own video content, which may be stimulating for technology-driven youth and can increase their engagement and accountability.

### *Mentimeter*

Mentimeter offers a collection of tools which enable making dynamic presentations with an online editor. The free version allows to reach an unlimited audience with unlimited presentations and include question and quiz slides. The built-in slide templates can be adapted to suit any type of presentation. The available editor features a simple and easy-to-navigate layout accompanied by help bubbles. It offers a variety of tools to share a presentation. With the available tools, Mentimeter converts passive viewing into active acquisition, where in-class and online students can discover, analyse, and apply concepts through augmented learning content. Mentimeter enables designing a wide variety of interactive polls, word clouds, quizzes or questions and answers, which can be the most effective follow-up to increase engagement, make a presentation memorable and add elements of both competition and interaction. Mentimeter

can be integrated seamlessly with the existing LMSs. In a language classroom, apart from its demonstration value, Mentimeter is a powerful and flexible tool which can improve the dynamics of the large classroom by promoting active learning and student participation. Instant formative assessment can be more interactive and exciting using Mentimeter. Recorded data can also be used further to design and supervise the available course. Both students and teachers appreciate the engagement, group involvement, collaboration as well the real-time feedback on mobile devices.

### *Canva*

Canva is meant for anyone wanting to design instructional content on their own or share it with others. Canva for Education, a variation of the primary platform, which has all the characteristics of Canva Pro but with additional benefits specifically tailored for teachers. It is appreciated for its simple interface and vast library of ready-to-use templates and resources. Design experience is not essential. It offers pre-designed templates for creating presentations to demonstrate an array of topics through individual or collaborative development of posters, flyers, infographics, book covers, newsletters, programmes, reports, and social network postings. The platform encourages communication between teachers and students through their collaboration during the learning process. Students get alerts, submit papers and engage with the assigned instructional resources. Canva integrates the accounts with social media accounts and allows teachers to post comments, integrate videos, blogs, wikis, and other instructional materials. Canva offers built-in audio and video recording function which supports the publication of course materials, setting homework assignments and tests. The platform also integrates with the major LMSs. In language teaching, apart from the content presentation, introduction of topics or follow-ups to previously published resources, it can be incorporated in teaching to write and share functional texts, that is, brochures, announcement, letters, and curricula vitae by (co-)editing the existing text according to the assignment set by the teacher. The content presentation tools mentioned above along with a wealth of other tools available online including: Animoto, ATutor, BrainPOP, EdrawMind, ClassFlow, Coggle, CourseSites, EdApp, Green Screen, Mindmaster, Nearpod, Prezi may help to set up an environment which assists presentations in-person, online synchronously and on-line asynchronously as well as assigning students video lessons to watch both in-class and at home. This may result in using class time more effectively with more meaningful activities. Students view the basic concepts of the class accessing the available content at their own pace, anytime and anywhere. This gives the teacher more time to work on accompanying and further support students' understanding of the topic. As regards

teaching a language any visual or video available online can be embedded to enrich the presented content and later check its acquisition. Moreover, sharing an interesting class-related topic using a video or going on a virtual field trip may enable exposing students to a language as it is spoken or discovering it in a cultural setting. Video presentations of grammar rules or vocabulary may help to augment the language experience and create opportunities for further language practice.

## Conclusions

The author's principal intention in this paper was to recommend to language teachers the hyflex model of instruction because it may enable them to make good use of the digital, 21st-century competences and skills acquired over the last two years to enhance their teaching methods and techniques. Such an enhancement is the evolution of what has been already taken place and a step outside the comfort zone. The hyflex model allows to tailor the educational offer in such a manner so as to suit the requirements of contemporary students. Sadly, these students may seem to perceive education as a product or service they can acquire rather than a journey they undertake to discover certain values education used to embody. Education, whether we accept it or not, has become an achievement-oriented set of competences, skills or numbers and percentages which stand for how much students know. It may help the young participants of the educational process to realise the needs they have which include living and working anywhere in the world, flexitime, self-fulfilment, approval, and comfort. The hyflex model although still requiring solid, empirical evidence and research either in its present shape or adjusted may take language learning and teaching to a next level where the parties involved decide how, where, and when education takes place. This resembles the mother tongue acquisition when language is acquired in the most natural environment where all the elements, namely knowledge, skills, and social competences acquired at the own pace. Moreover, the participants of the instructional process, regardless of how, where, and when they learn, may achieve comparable results through meaningful reflection, interaction, and response to similar stimuli. The hyflex model of instruction which undoubtedly calls for research and evidence to prove its value may help to answer some of the above considerations bearing in mind what has been achieved to date. The scaling and contextualisation of the hyflex model may merge the traditional language teaching and learning with the online interaction both synchronously and asynchronously. The transformed, hyflexed language education may allow the introduction of tools which offer concurrent

design and delivery of instruction in different environments, assessment of student progress and finally the demonstration of content on the basis of the undertaken evaluation.

## References

- Abdelmalak, M. M. M., & Parra, J. L. (2016). Expanding learning opportunities for graduate students with HyFlex course design. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, 6(4), 19–37.
- Alexander, M. M., Lynch, J. E., Rabinovich, T., & Knutel, P. G. (2014). Snapshot of a hybrid learning environment. *The Quarterly Review of Distance Learning*, 15(1), 9–21.
- Annamalai, N., & Kumar, J., A. (2020). Understanding smartphone use behavior among distance education students in completing their coursework in English: A mixed-method approach. *The Reference Librarian*, 61(3–4), 199–215. <https://doi.org/10.1080/02763877.2020.1815630>
- Beatty, B. (2012). HyFlex course design: The advantages of letting students choose the blend. *Proceedings of the 9th Annual Sloan Consortium Blended Learning Conference & Workshop*. <http://dx.doi.org/10.4018/978-1-5225-5466-0.ch015>
- Beatty, B. (2009). Student self-reflections on learning in a hybrid course environment: Do participation mode differences lead to differences in reflections? Peer-reviewed paper presented at the *Association for Educational Communication and Technology International Conference*, Louisville KY.
- Beatty, B. (2019). *Hybrid-Flexible course design* (1st ed.). EdTech Books. Retrieved from <https://edtechbooks.org/hyflex>
- Bernard, R. M., Borokhovski, E., Schmid, R. F., Tamim, R. M., & Abrami, P. C. (2014). A meta-analysis of blended learning and technology use in higher education: From the general to the applied. *Journal of Computing in Higher Education*, 26(1), 87–122. <https://doi.org/10.1016/j.chb.2009.04.007>
- Bonk, C. J., & Graham, C. R. (2006). *The handbook of blended learning: Global perspectives, local designs*. John Wiley & Sons.
- Cheung, S., Kwok, L., Phusa, K., & Yang, H. (2021). Shaping the future learning environments with smart elements: Challenges and opportunities. *International Journal of Educational Technology in Higher Education*. <http://dx.doi.org/10.1186/s41239-021-00254-1>
- Detienne, L., Raes, A., & Depaepe, F. (2018). Benefits, challenges and design guidelines for synchronous hybrid learning: A systematic literature review. In T. Bastiaens, J. Van Braak, M. Brown, L. Cantoni, M. Castro, R. Christensen, G. Davidson-Shivers, K. DePryck, M. Ebner, M. Fominykh, C. Fulford, S. Hatzipanagos, G. Knezek, K. Kreijns, G. Marks, E. Sointu, E. Korsgaard Sorensen, J. Viteli, J. Voogt, P. Weber, E. Weippl, & O. Zawacki-Richter (Eds.), *AACE. Proceedings of EdMedia: World Conference on Educational Media and Technology*, 2004–2009. Retrieved from <https://www.learntechlib.org/primary/p/184440/>
- Dhawan, S. (2020). Online learning: A panacea in the time of COVID-19 crisis. *Journal of Educational Technology Systems*, 49(1), 5–22. <https://doi.org/10.1177/0047239520934018>
- Dziuban, C., Moskal, P., Thompson, J., Kramer, L., DeCantis, G., & Hermsdorfer, A. (2015). Student satisfaction with online learning: Is it a psychological contract? *Online Learning*, 19(2). <http://dx.doi.org/10.24059/olj.v19i2.496>

- Gobeil-Proulx, J. (2019). La perspective étudiante sur la formation comodale, ou hybride flexible [What do university students think about hybrid-flexible, or HyFlex courses?] *Revue internationale des technologies en pédagogie universitaire*, 16(1), 56–67. <https://doi.org/10.18162/ritpu-2019-v16n1-04>
- Harun, F., & Hussin, S. (2018). Speak through your mobile app changing the game: English language in education 4.0. In *Proceedings of the 27th Melita International Conference, Johor Bharu, Malaysia*, 186–196.
- He, W., Gajski, D., Farkas, G., & Warschauer, M. (2015). Implementing flexible hybrid instruction in an electrical engineering course: The best of three worlds? *Computers & Education*, 81, 59–68.
- Heilporn, G., & Lakhal, S. (2021). Converting a graduate-level course into a HyFlex modality: What are effective engagement strategies? *The International Journal of Management Education*, 19(1), 100454. <https://doi.org/10.1016/j.ijme.2021.100454>.
- Idrizi, E., Filiposka, S., & Trajkovik, V. (2019). The discourse on learning styles in online education. *27th Telecommunications Forum (TELFOR)*, 1–4. <http://dx.doi.org/10.1109/TELFOR48224.2019.8971204>
- Klimova, B., F., & Kacetla, J. (2015). Hybrid learning and its current role in the teaching of foreign languages. *Procedia Social and Behavioral Sciences Journal*, 182, 477–481.
- Koskinen, M. (2018). Understanding the needs of adult graduate students: An exploratory case study of a Hyflex learning environment. Doctoral thesis, Northeastern University, Boston, MA. Retrieved from <https://repository.library.northeastern.edu/files/neu:m044c738v/fulltext.pdf>
- Leijon, M., & Lundgren, B. (2019). Connecting physical and virtual spaces in a HyFlex pedagogic model with a focus on teacher interaction. *Journal of Learning Spaces*, 8(1). Retrieved from <https://libjournal.uncg.edu/jls/article/view/1640/1338>
- Lockman, A. S., & Schirmer B., R. (2020). Online instruction in higher education: Promising, research-based, and evidence-based practices. *Journal of Education and e-Learning Research*, 7(2), 130–152.
- Love, S. (2015). A quantitative inquiry into first generation students' readiness for distance education. *ProQuest Dissertations Publishing*. Ann Arbor, MI, 2015.
- Malczyk, B. R., & Mollenkopf, D. (2019). Student-centered blended Learning: The HyFlex approach to blended Learning. *Innovation in Pedagogy and Technology Symposium, 2019 (Selected Conference Proceedings)*, 86–87.
- Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education*, 159/104009. <https://doi.org/10.1016%2Fj.compedu.2020.104009>
- McBride, K., & Fägersten, K. B. (2008). Student's role in distance learning. In S. Goertler & P. Winke (Eds.), *Opening doors through distance language education: Principles, perspectives, and practices* (pp. 43–66). CALICO.
- Miller, J., B., Risser, M., D., & Griffiths, R., P. (2013). Student choice, instructor flexibility: Moving beyond the blended instructional model. *Issues and Trends in Educational Technology*, 1(1), 8–23.
- Oliver, M., & Trigwell, K. (2005). Can 'blended learning' be redeemed? *E-Learning and Digital Media*, 2(1), 17–26. <http://dx.doi.org/10.2304/elea.2005.2.1.17>
- Parra, J., L., & Abdelmalak, M. (2016). Expanding learning opportunities for graduate students with HyFlex course design. *International Journal of Online Pedagogy and Course Design*, 6(4), 19–37. <http://dx.doi.org/10.4018/IJOPCD.2016100102>
- Pinchot, J., & Pullet, K. (2014). Different keystrokes for different folks: Addressing learning styles. *Online Education Information Systems Education Journal (ISEDJ)*, 12(2). Retrieved from <https://eric.ed.gov/?id=EJ1140787>

- Qayyum, A., & Zawacki-Richter, O. (Eds.) (2019). *Open and Distance Education in Asia, Africa and the Middle East National Perspectives in a Digital Age*. Springer Open Singapore. Retrieved from <https://library.oapen.org/handle/20.500.12657/23272>
- Raes, A., et al. (2019). A systematic literature review on synchronous hybrid learning: Gaps identified. *Learning Environments Research*, 23, 269–290. Retrieved from <https://link.springer.com/article/10.1007/s10984-019-09303-z>
- Raes, A., Vanneste, P., Pieters, M., Windey, I., Van Den Noortgate, W., & Depaepe, F. (2020). Learning and instruction in the hybrid virtual classroom: An investigation of students' engagement and the effect of quizzes. *Computers & Education*. <https://doi.org/10.1016/j.compedu.2019.103682>
- Ramalingam, S., Yunus, M., M., & Hashim, H. (2022). Blended learning strategies for sustainable English as a second language education: A systematic review. *Sustainability*, 14, 8051. <https://doi.org/10.3390/su14138051>
- Reynard, R. (2007). Hybrid learning: Challenges for teachers. *T.H.E. Journal*. Retrieved from <http://thejournal.com/Articles/2007/05/17/>
- Riel, J., Lawless, K. A., & Brown, S. W. (2016). Listening to the teachers: Using weekly online teacher logs for ROPD to identify teachers' persistent challenges when implementing a blended learning curriculum. *Journal of Online Learning Research*, 2(2), 169–200.
- Robertson, B., & Kelly, K. (2013). Operating a very large-section, hybrid principles of marketing class at a public university: Lessons learned over ten years. *Atlantic Marketing Journal*, 2(3), Article 10. Retrieved from <https://edtechbooks.org/-jVE>
- Romero-Hall, E., & Ripine, C. (2021). Hybrid flexible instruction: Exploring faculty preparedness. *Online Learning*, 25(3), 289–312. <https://doi.org/10.24059/olj.v25i3.2426>
- Serdyukov, P., & Hill, R. A. (2013). Flying with clipped wings: Are students independent in online college classes? *Journal of Research in Innovative Teaching*, 6(1), 53–65.
- SFSU Academic Senate Policy S16–264, 2010. Retrieved from [https://senate.sfsu.edu/policy/online-education-policy-1\\_](https://senate.sfsu.edu/policy/online-education-policy-1_)
- Shek, D. T. L., Zhu, X., Li, X., & Dou, Diya. (2022). Satisfaction with HyFlex teaching and law-abiding leadership education in Hong Kong University students under COVID-19. *Applied Research Quality Life*, 17, 2833–2858. <https://doi.org/10.1007/s11482-022-10040-4>
- Singh, J., & Matthees B. (2021). Facilitating interprofessional education in an online environment during the COVID-19 pandemic: A mixed method study. *Healthcare*, 9(5), 567. <https://doi.org/10.3390/healthcare9050567>
- Smith, D., & Hardaker G. (2000). E-Learning innovation through the implementation of an internet supported learning environment. *Educational Technology & Society*, 3(3), 1–16.
- Tallent-Runnels, M. K., Thomas, J. A., Lan, W. Y., Cooper, S., Ahern, T. C., Shaw, S. M., & Liu, X. M. (2006). Teaching courses online: A review of the research. *Review of Educational Research*, 76(1), 93–135. <http://dx.doi.org/10.3102/00346543076001093>
- Verrecchia, P., J., & McGlinchey, M. J. (2021). Teaching during Covid: The effectiveness of the HyFlex classroom in a 300 level statistics class. *Journal of Education and Training Studies*, 9(3), 23–27. Redfame publishing. Retrieved from <https://ideas.repec.org/a/rfa/jetsjl/v9y2021i3p23-27.html>
- Ying, C., L., J., Embi, M., A., & Hashim, H. (2019). Students' perceptions toward gamification in ESL classroom: KAHOOT! In *Proceedings of the International Conference on Creative Teaching, Assessment and Research in the English Language*, Melaka, Malaysia, 129–132.
- Zajac, M. (2020). Kursy typu HyFlex – możliwości i wyzwania. *e-mentor*. Retrieved from <http://www.e-mentor.edu.pl/blog/wpis/id/99>

Zydney, J. M., McKimm, P., Lindberg, R., & Schmidt, M. (2019). Here or there instruction: Lessons learned in implementing innovative approaches to blended synchronous learning. *TechTrends*, 63, 123–132. <https://doi.org/10.1007/s11528-018-0344-z>

Robert Oliwa

## **Kontextualisierung des Hyflex-Modells für den Sprachunterricht**

### **Zusammenfassung**

Der Artikel befasst sich mit einer möglichen, praktischen Anwendung des hybriden flexiblen (hyflex) Modells zur Vermittlung von Inhalten in dem tertiären Kontext. Das Modell kann auch im Fremdsprachenunterricht angewendet werden. Nach Meinung des Autors gibt es überzeugende Gründe dafür, die in den letzten zwei Jahren erworbenen digitalen Kompetenzen und Fähigkeiten zur Verbesserung der Lehr- und Lernprozesse einzusetzen. Eine angemessene Kontextualisierung des Hyflex-Modells kann die Beibehaltung der bereits erworbenen Kompetenzen und Fähigkeiten in der neuen Normalität weiter fördern. Sie kann es auch ermöglichen, den traditionellen Klassenunterricht mit der neuen Online-Interaktion zu kombinieren, sowohl synchron als auch asynchron. Die Möglichkeit der Flexibilisierung von Sprachkursen kann eine herausfordernde, aber zugleich auch eine verantwortungsvolle Perspektive sein, weil eine Rückkehr zum Unterricht in bisher bekannten Form nicht einfach ist. Obwohl der Artikel theoretisch ausgelegt ist, bietet er praktische Leitlinien und eine Auswahl an Online-Tools, um das Hyflex-Modell im Sprachunterricht zu kontextualisieren. Darüber hinaus wird darin stark für die praktische Anwendung des Hyflex-Modells plädiert, insbesondere in Verbindung mit einer eingehenden Analyse und Untersuchung des Modells in verschiedenen Bildungskontexten. Da die Zukunft schwer vorhersehbar ist, könnten verschiedene Modelle und Ansätze in Betracht gezogen werden, um angemessen auf die Herausforderungen und Anforderungen zu reagieren, mit denen die Bildung wahrscheinlich konfrontiert sein wird.

*Schlüsselwörter:* hybrid, hyflex, Plattform, online, Sprache, Unterricht





**Sabina Nowak**

University of the National Education Commission, Krakow, Poland

 <https://orcid.org/0000-0001-9761-8365>

**Aleksandra Rążewska**

Polish Nobel Laureates Primary School, Luborzycza, Poland

 <https://orcid.org/0000-0002-7025-5995>

## **Challenging but Rewarding Aspects of Telecollaboration: The Case of Virtual Israeli-Polish (VIP) Project**

### **Abstract**

Starting from the premise that implementing telecollaboration can be a challenging enterprise, it is assumed that such virtual exchanges bring positive outcomes. Despite rich body of research on telecollaboration, few studies to date have explored *group dynamics* in the context of online exchange. The current research examined an Israeli-Polish ( $N = 100$ ) telecollaborative intercultural experience. The Israeli participants were students from a BEd program in informal education and Polish students from a BA program in translation. The main purpose of the telecollaborative project was to develop English linguistic skills and intercultural communicative competence (Byram, 1997) of the students. However, the present study aimed to investigate how the group dynamics influenced the telecollaboration process and the project outcomes. The students engaged in the exchange of asynchronous video recordings (Vlogs), collaborative synchronous meetings on Zoom, and completed an on-line questionnaire. This paper outlines some of the challenges related to the international telecollaboration, including technological and institutional difficulties, context-specific demands, and some individual differences of the participants. Through “thematic analysis” (TA) (Braun & Clarke, 2012) the authors of the paper explored levels of “failed communication” (O’Dowd & Ritter, 2006) and how they impacted the ongoing development of the collaboration both technically and in terms of content and development of interaction skills to uncover certain themes that referred to learners’ and instructors’ challenges. The study concludes with pedagogical implications for more effective implementation of telecollaboration in higher education.

*Keywords:* telecollaboration, group dynamics, vlogging, reticence

Prior to the COVID-19 pandemic, telecollaboration was already lauded as a significant tool for foreign language learning because of its possibility to engage learners in authentic communication and meaningful collaboration across national and cultural boundaries (Godwin-Jones, 2019; O'Dowd, 2016). The COVID world-wide pandemic, however, has further legitimized the use of this tool. And yet, as with all technological tools, the virtues of telecollaboration may be tempered by multiple sources of potential challenge: from technological to contextual relating to content or participant characteristics. The aspects of telecollaboration as a process based on social interactions during a virtual exchange will be developed further in the paper. The importance of conducting the present study assumed that the complex nature of collaboration may be facilitated by online intercultural exchanges. In line with that, the constructivist worldview shall be a starting point for the introduction of communication to classroom settings. It explains the importance of social interactions in language development and encourages the most natural way of learning language through meaningful discourse.

## **Constructionism and Communicative Language Teaching**

The constructivists provided an insightful analysis of social aspects of learning processes connected with ways of acquiring a language which help to understand the importance of social relations in teaching and learning English as a second language. The general assumption of constructivism is based on the idea that people gain knowledge through their beliefs and individual experiences, in this sense, people are active learners, and they explore new concepts for themselves. This perspective advocates provision of opportunities to be actively engaged in exploring new problems, for instance, through social interaction. However, constructivism cannot be interpreted as a unified framework, in fact, three major perspectives can be distinguished: *exogenous*, *endogenous*, and *dialectical* (Schunk, 2012, p. 232). By the *exogenous* perspective we understand that one's knowledge is influenced by the external circumstances, exposure to models and teaching. Contrary, *endogenous* perspective assumes that knowledge reflects one's previous experiences and is not directly linked to external world. The last one exhibited within the constructivism is a *dialectical* view. According to this view, the source of knowledge can be found in interactions between people and their environments.

When in the 1970s language learning faced a paradigm shift, the outlook on language learning was revised and a new method emerged which addressed successful communication in authentic social contexts. Richards and Rodgers

(2014) put forward two main aims of Communicative Language Teaching (CLT), namely fostering communicative competence and establishing a set of effective practices for the development of all four language skills. The authors also mention three elements that may serve as principles of CLT. Firstly, to stimulate language learning, it is necessary to introduce tasks that engage learners in authentic communication. Secondly, L2 learning can be enhanced when elements of language are incorporated in meaningful activities. Lastly, language learning can be enhanced when linguistic material is meaningful for learners. According to these principles, meaningful communication is the core of CLT. Therefore, to facilitate interaction between students it is advisable to incorporate topics which learners are familiar with and which they take an interest in. Larsen-Freeman and Anderson (2011) explain that communicative interactions, which strongly encourage cooperative interchange, are at the centre of the above-mentioned approach. CLT introduces more authentic ways of using language in the classroom settings, which enable learners to develop fluency especially in oral production. It strives to develop communicative proficiency rather than focus on structural accuracy (Richards & Rodgers, 2014, p. 84). Moreover, Communicative Language Teaching and Learning is a gateway to meaningful language learning as it facilitates collaboration and cooperation between learners. Better understanding of the two concepts, which will be presented in the succeeding part of the paper, will allow us to indicate how communication and interaction can be enhanced in ESL.

## **Cooperation and Collaboration in EFL**

As indicated in Common European Framework of Reference for Languages (Council of Europe, 2020), the ability to communicate and collaborate is important in language education. Cooperation and collaboration are believed to be valuable and desired life skills, mainly as they are prerequisites for successful project work. It is essential to contextualise the two concepts in the context of L2 learning since they can be confused as interchangeable terms. Some researchers use the notions as synonyms, nonetheless, certain differences are present in those two forms of learning.

Following the equivocal definition provided by Johnson and Johnson (2008), *cooperative learning* is described as “students working together to maximise their own and each other’s learning” (p. 402). The given explanation is rather general and does not concern the mechanism which is used by the groups working together. A more detailed definition is constructed by Olsen and Kagan (1992), who conceptualise Cooperative Language Learning as “group learning

activity organised so that learning is dependent on the social structured exchange of information between learners in groups and in which each learner is held accountable for his or her own learning and is motivated to increase the learning of others” (p. 8). This understanding of cooperation suggests that activities implemented within the framework are directly linked to social context of learning and the main purpose is to obtain necessary information and process it to complete a task. Moreover, it implies that participants exhibit a sense of responsibility for the success of a project.

In *collaborative learning*, the authority is given to the group members, which means that they are empowered to achieve their own goals (Rockwood, 1995). As clarified by De Florio (2016), collaborative learning concentrates on an artefact or a product of teamwork, where the group members decide on a joined endeavour such as project or a problem. Rockwood (1995) indicates that collaboration knowledge is viewed as a social construct and consequently learning is interpreted as a social process. Collaboration requires “[1] managing one’s role and contributions to group communication, [2] active orientation of teamwork by helping to review key points and consider or define next steps, [3] use of questions and contributions to move the discussion forward in a productive way, [4] use of questions and turn taking to balance contributions from other group members with their own contributions” [...] [5] cognitively framing collaborative tasks by deciding on aims, processes and steps, [6] co-constructing ideas, solutions, [7] asking others to explain their thinking and identifying inconsistencies in their thought processes, [8] summarising the discussion and deciding on next steps” (Council of Europe, 2020, p. 109). Beatty (2010) claims that collaboration has numerous advantages in addition to the social comprehension of learners, arguing that it is the solution to urge spontaneous discourse as it enables to absorb learners in the direct communication. Nokes-Malach et al. (2019) recognise two groups of factors, namely cognitive and social, for both benefits and costs. As mentioned by Beatty (2010), peer collaboration is the core of cognitive and cultural development, whereas the category of social factors concerns the nature of interactions among groups members. According to Storch (2002), the most beneficial type of interaction is collaborative, as it allows for equal contribution in a task and mutual engagement in each other’s work. The idea of collaborative and cooperative learning may be especially visible by examining group dynamics.

## Group Dynamics

The idea of “group dynamics” is an underresearched concept in telecollaboration, despite its significance in L2 learning and teaching. The concept belongs to the field of social science and poses two main objectives (a) members of a group exhibit different behaviours compared to the people not associated with the group, and (b) despite the immense diversity of groups characters, they share universal features (Dörnyei & Murphey, 2003). Group dynamics, as defined by Forsyth (2019), “are the influential interpersonal processes that occur in and between groups over time” (p. 18). The above-mentioned processes determine how members of the group react, what group’s aims are, and what actions they adapt (Forsyth, 2019). Consequently, being more efficient than individuals on their own, group dynamics may influence effectiveness of learning.

One of the first stages of working in groups usually entails their selection. It can be the crucial point in organisation of a project since it determines attitudes of learners and their motivation to contribute. Jolliffe (2007) indicates three methods of assigning learners to groups, namely, “random selection,” “pupil selection,” and “teacher selection” (p. 50). Random and teacher selections can be conducted in various ways, for instance, by numbering students and assigning them to the respective group. Pupil selection allows students to select partners on their own, which can have positive or negative consequences on group dynamics.

Forming groups might be crucial in successful telecollaboration. Forsyth (2019) identifies different stages of group development (Figure 1). The first stage, called “forming,” concerns the initial recognition of group members. Members may be reserved or may prefer to observe others to gain information. In the “storming” stage, communication is rather limited, and participants may be hesitant to share opinions. The next stage, “norming,” refers to the phase when group cohesion is established, simultaneously members establish group norms and agree on the disciplined participation. “Performing” is the stage in which roles of members are utilised in a collective decision-making or problem-solving activity. The last stage is referred to as “adjourning” (Tuckman & Jensen, 1977, pp. 1419–1426) in which a post-project summary or key points of the lesson-learned are formulated. At this stage, learners may evaluate their own contribution, feelings and experiences of working with others. It is usually followed by a celebration of group’s achievements.

Forsyth (2019) lists the following group dynamics processes: *formative*, *influence*, *performance*, *conflict*, and *contextual processes* (pp. 18–19). *Formative processes* concern the initial relation within the group. Since members do not share personal information, the group is reserved. Later, as members realise the need to participate together, they overcome the inhibitions and foster group

cohesion. The next types of processes are *influence processes*. Being a part of a group requires the participants to follow standards set by their members, fulfil the assigned roles and agree to respect leader's directions. Essentially, all members are affected by the group and simultaneously they also influence the other members. *Performance processes* refer to group's activity towards their goal. Other significant elements of group dynamics are *conflict processes*. Since a group is the collection of people with different opinions and visions, for this reason conflicts are natural to occur. The sources of disagreements may be various, for instance, competition, power struggle, disaccord in terms of decision making or individual antagonisms. As the last component Forsyth (2019) lists *contextual processes*. They address the physical environment of the group and its purpose. These types of processes explain the context of environment and society in group dynamics. Working in a group increases learners' autonomy by providing learners with the opportunity to make their own decisions with whom they want to collaborate. Each group needs to establish its internal rapport and communicate effectively to reach a shared goal of collaboration.

## **Telecollaboration in Language Learning**

As noticed by Boss and Krauss (2007), technology has become an integral instrument used by learners in a language classroom to explore online resources, collaborate, and communicate (p. 12). Although online collaboration was broadly discussed in CEFR (Council of Europe, 2020), telecollaboration is "a relatively recent teaching tool" (Pfungsthor, Kramer, Czura, & Stefl, 2019, p. 160). Therefore, there is a need to explain what the term entails in more detail and what its main characteristics can refer to in the context of language learning. O'Dowd (2018) proposes the following definition of telecollaboration, "the engagement of groups of learners in online intercultural interactions and collaboration projects with partners from the other cultural context or geographical locations as an integrated part of their educational programmes" (p. 1). Telecollaborative learning involves group or pair work and largely depends on group interactions. Ware (2018) notices that the term can denote a myriad of activities and may refer to any combination of text-based, multimodal-enhanced, asynchronous, synchronous, monolingual, bilingual, and multilingual collaboration. Dooly (2017) defines the term as "the process of communicating and working together with other people or groups from different locations through online or digital communication tools (e.g., computers, tablets, cell phones) to co-produce a desired work output" (p. 169). Both definitions point out to the communication between learners coming from distant geographical

locations and representing various cultural backgrounds. An alternative definition was provided by Guth and Helm (2012), who refer to telecollaboration as “Internet-based cultural exchange between people of different cultural/national backgrounds, set up in an institutional context with the aim of developing both language skills and intercultural communicative competence” (p. 14).

Success in telecollaboration depends on proactive attitudes and digital competences as well as organisational and pedagogical skills (O’Dowd, 2015). Telecollaborative projects provide numerous opportunities to improve language skills in meaningful settings. O’Dowd (2018) notices that virtual exchanges can introduce cross disciplinary perspective on language learning. As a result, participants of such exchanges are provided with an opportunity to utilise their language skills in meaningful interactions with their non-native partners. Moreover, online exchanges open learners to explore new cultures as learning takes place in a contextualised way. Also, the ability to work in a team is currently a vital asset for employees. It is also believed that features of telecollaborative work can contribute to reduction of stress that is present in more traditional methods of learning (Taskiran, 2019).

A substantial body of literature exists concerning the benefits and challenges of telecollaboration (Table 1). Among the main advantages of telecollaboration for *language learning* researchers list vocabulary and grammar improvement (Chen & Eslami, 2011), building fluency in pronunciation (Kabata & Edasawa, 2011), increased quality of production (Jin, 2013), as well as instant linguistic feedback (Kötter, 2003). Additionally, participation in the multilingual social context can lead to better cultural knowledge and sensitivity (Canto et al., 2013). Through intercultural meetings with “the other” participants may begin to question previously held beliefs (O’Dowd, 2003). Researchers have examined telecollaboration across a variety of contexts using several methodological approaches that cluster around three main areas of inquiry: (a) language development, (b) intercultural communication, and (c) identity (Pfungsthorn et al., 2019).

**Table 1**

*Benefits and Challenges of Telecollaboration for Language Learning (TlcLL)*

Benefits <i>also referred to as claims of success</i>	Challenges <i>also referred to as tensions, failure, difficulties, pitfalls</i>
cultural awareness raising (Muller-Hartmann, 2000)	mismatches between individual learning outcomes and collaborative online activities (Greenfield, 2003)
'communication' skills development (Egert 2000; Lee, 2004)	institutional and professional misalignments (Belz & Müller-Hartmann, 2003; O’Dowd, 2005)

linguistic and social linguistic advances (Kötter, 2003)	workload (Rösler, 2004) task design (Dooly, 2010)
questioning previously held beliefs (O'Dowd, 2003)	age differences (Lee, 2005)
personal and cultural benefits (Itakura, 2004; Jin & Erben, 2007)	practical constraints (e.g., different time zones, semester dates, assessment arrangements) (Ware, 2005)
teacher professional development (Arnold & Ducate, 2006; Lund, 2006)	differences in interactional styles (regarding communication and negotiation) (Ware, 2005; Ware & Kramersch, 2005)
vocabulary / grammar (Chen & Eslami, 2011)	incompatibilities in world views (Ware & Kramersch, 2005) or cultural diversity (Weller & Conole, 2008).
pronunciation fluency (Kabata & Edasawa, 2011)	psychological pressures due to synchronicity (mainly being recorded, video) (O'Dowd, 2006)
cultural knowledge and sensitivity (Canto et al., 2013)	linguistic challenges (Lee, 2006) negative grammatical-pragmatic transfer (Blake & Zyzik, 2003)
quality of productive skills (Jin, 2013) improve the speaking skills (Mandasari & Aminatun, 2020)	technological challenges (Helm, 2015)

*Note:* The compilation of challenges was based on Lamy & Goodfellow (2010, pp. 109–110).

As presented in Table 1, a highly respectable compilation of challenges (also referred to as *tensions, failure, difficulties, pitfalls*) of telecommunication was identified by Lamy and Goodfellow (2010). Kohn & Hoffstaedter (2017) recognized challenges of telecollaboration related to “sustainable pedagogical implementation” (p. 14), including learner preparation, individual learning flexibility at home (not as regular class time), technological infrastructure, IT support, and lingua franca pedagogy. Some obstacles can be the result of age or interactional styles differences, which can hinder learners’ engagement in such online exchange. As pointed by Lamy and Goodfellow (2010), some behaviours can be interpreted as face threatening for the other culture. In fact, O’Dowd and Ritter’s (2006) recognised that instances of “failed communication” can occur on various levels, namely: *the individual* (learners’ intercultural knowledge, motivation, stereotypes, expectations), *classroom* (task design, teacher-to-teacher relationship, group dynamics, matching of the participants), *socio-institutional* (course organisation, workload and time constraints, assessment), and *interaction* levels (cultural differences, learners’ engagement). From the perspective of a teacher one can notice a challenge in managing institutional and technological aspects of online exchanges. Not only must the teacher navigate multiple

tools, but also reassure students' progression through a task design (Dooly, 2010). O'Dowd and Eberbach (2004) list tasks that teachers, who embark on a telecollaboration project, are required to complete. These include rising awareness among learners, teaching how to contribute to a project, or establishing partnerships with teaching facilities. Oftentimes, teachers may be overwhelmed with the duties connected with designing and conducting telecollaboration.

## **Rationale for the Study**

In this part of the paper, the rationale of the telecollaboration is outlined, followed by an explanation of the methodology and research design. The main aim of the telecollaborative project was, on the one hand, to develop the English linguistic skills and intercultural communicative competence (Byram, 1997) of the students, and on the other, to explore the challenges of the telecollaboration by examining group dynamics. Even though there have been many research projects which highlighted the challenges of telecollaboration (Table 1), not much has been written about what role group dynamics plays in influencing the process.

The main rationale for this study is to discuss challenges and benefits of telecollaboration seen from the perspective of the participants and instructors. This study aims to obtain data concerning group work in telecollaboration as it seeks to investigate what role group dynamics might specifically play in the process. Dooly (2017) advocates "a need for more research into political and social implications of telecollaboration" (p. 177). O'Dowd (2015) claims that "the literature and tools related to teacher training and CALL have not paid great attention to the challenges of establishing and running telecollaborative exchange projects" (p. 64). The potential challenges may range from technological to contextual relating to content or participant characteristics. The examined literature suggests that project work and group work activities are effective methods in fostering collaboration among learners; however, Le, Janssen, and Wubbels (2018) claim that collaboration at university may be hindered by "students' lack of collaborative skills, free-riding, competence status and friendship" (p. 1).

As the main framework to study instructors' challenges, we adopted "experiential modelling approach" (Guichon & Hauck, 2011, p. 195; O'Dowd, 2017, p. 38), which entails having a hands-on experience and gaining experience in telecollaborating oneself. By studying learners' behaviours and attitudes (group dynamics) while participating in the project, we observed the impact of the online environment to distinguish the aspects that may be challenging

and beneficial for learners. We aimed to examine the main challenges of telecollaborative learning for instructors and explore pedagogical implications, which can be applied in tertiary education. To develop intercultural competence, we used the *Cultura* model<sup>1</sup> developed by Furstenberg, Levet, English, and Maillet (2001), in which a learner language is viewed not as the end goal for instruction, but rather as a means by which intercultural understanding can be developed. We also adopted Byram's model (1997) of intercultural competences as the framework for the study which refer to dialectical aspect mentioned by Schunk (2012).

## Methodology and Research Questions

We used “convergent mixed methods” (Creswell & Creswell, 2018, p. 51) which indicate that both quantitative and qualitative data support the research. As defined by Leavy (2017) “convergent or concurrent designs involve collecting both quantitative and qualitative data, analysing both datasets, and then integrating the two sets of analyses in order to cross-validate or compare the findings” (p. 175). The study included several sources of data, in which qualitative instruments were Vlogs posted via online platform Flipgrid, accompanied with students' comments posted under the videos and a quantitative instrument was a questionnaire. Additionally, two synchronous online meetings on Zoom were observed and analysed to determine the way the Polish and Israeli students managed their work in real time during online telecollaboration. The above-mentioned instruments allowed for the triangulation of data sources which assured the findings to be reliable and accurately applicable in the context of the study. With the use of Vlog recordings as well as comments posted by the students on Flipgrid, observation of Zoom meetings and the questionnaire, the researchers aimed to obtain information necessary to answer the following research questions.

RQ 1. What behaviours do Polish and Israeli students exhibit while interacting asynchronously via Vlogs?

RQ 2. What group dynamics can be observed when Israeli and Polish students meet synchronously online?

RQ 3. What are the benefits and challenges of the Israeli-Polish telecollaboration?

---

<sup>1</sup> The *Cultura* model aims to develop better understanding between the students coming from the Muslim and Western worlds. Retrieved from: <http://cultura.mit.edu>.

The first research question aimed at observing behaviours of Polish and Israeli students during asynchronous interaction on Flipgrid. To successfully break down students' participation in the activity, a "thematic analysis" approach (Braun & Clarke, 2012) of video logs content was conducted. The Vlogs analysed in this part of the research provided information about group dynamics exhibited by the students. The second question sought to investigate how students managed to work in cross-national groups and what difficulties they encountered in the online environment. To answer this question, the researchers decided to observe an online synchronous meeting on Zoom and analyse group dynamics and interaction patterns between students in groups. This question aimed to explore participants' problems they encountered when fulfilling the given tasks. Language used by the participants and their reactions towards the events encountered would serve as hints to formulate an explanation to their problems. Lastly, the third research question focused on students' opinions of group work, as well as their perceptions of the challenges faced during the telecollaboration they participated in. It was possible to identify participants' opinions through the questionnaire, which aimed to identify learners' preferences concerning various aspects of group work to unveil possible restraints and identify the differences that Polish and Israeli students exhibited. It was expected to determine the challenges and benefits the participants noticed when working telecollaboratively. This objection aimed to explore the aspects that were the most difficult to deal with from students' perspective, as well as verify if students recognised positive features of telecollaboration.

## **Context and Participants**

Implementing the Polish-Israeli project required the cooperation of two instructors, one from Israel and another from Poland, as well as one Pedagogical University student, who at the time of the telecollaboration was conducting her master's degree based on the project. The Israeli instructor sent an invitation to the telecollaboration to the Polish University which specified the aims of the virtual exchange. The total number of participants was 100, including 40 Polish and 60 Israeli students (Table 2). The Polish group ages ranged from 18 to 25; whereas the majority of Israeli group ranged from 25 to 35 years old. The Polish participants were second year students of undergraduate programme studying translation in English philology department. Two groups of Polish students (each consisting of ca. 20 students) were randomly selected by the Polish teacher. At the time of the research, the students were enrolled in Discussion course as part of their Bachelor of Arts degree in applied lin-

guistics programme at the Polish university.<sup>2</sup> The students from Israel came from different professional backgrounds and they enrolled to an English course. They were a more diverse group consisting of people from various professions. The proficiency level of English represented by the participants was estimated by the instructors between B1–C1 (CEFR). The Israeli participants exhibited mix-ability level of English proficiency, ranging from B1–C1. The groups of Polish students were more linguistically homogenous (B2/C1). English was the language of instruction in the project.

**Table 2***The Israeli and Polish Participants of the Telecollaborative Project*

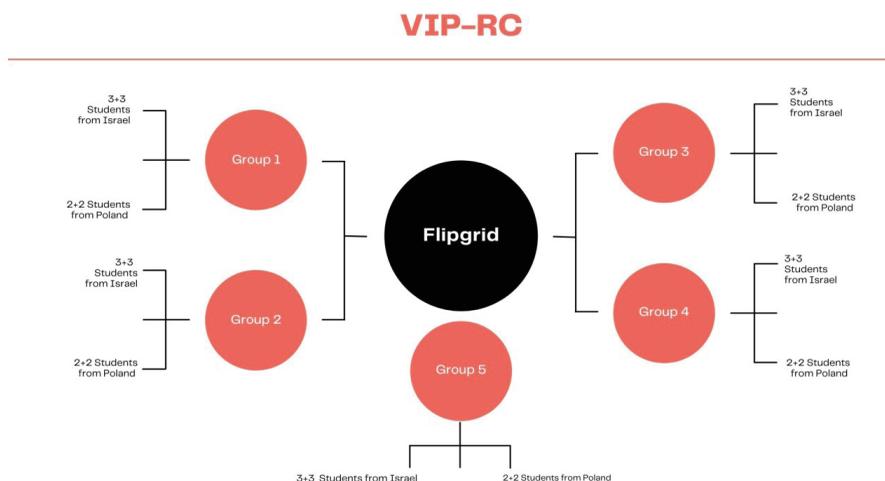
Nationality	Number of students	Age	Field of study	English language proficiency	Learner characteristics
Israeli	60	25–35	Informal education with a teaching certificate in History	CEFR B1–B2/C1	Youth movement working, studying, and living together
Polish	40	18–25	Translation studies conducted in English	CEFR B2/C1	English studies (translation)

Note:  $N = 100$  (total number of students in the study).

Mapping of the groups (Figure 1), which was prepared by the instructors for the participants, entailed pairing the Israeli and Polish students. The random assignment of the students was aimed to increase the generalizability of the results. However, the participants opted for “pupil selection” (Jolliffe, 2007) and chose their national partners on their own. The students ( $N = 100$ ) assigned themselves into 18 available groups, each consisting of four to eight students. The groups were first formed by the Israeli students themselves and Excel file with their names was then shared with the Polish students, who added their names to the already formed groups. As there were more students from Israel, they outnumbered the Polish participants in each of the groups.

<sup>2</sup>The structure of the program can be found here: <https://anglistyka.up.krakow.pl/wp-content/uploads/sites/31/2022/06/BACHELORS-DEGREE-PROGRAMME-2022-2023.pdf>.

**Figure 1**  
*Mapping of the Groups (Numbers 1–5).*



For the purpose of the research, every group received a coded name, for example, Gr18IS, in which the last element indicated either Polish or Israeli origins. Each student was given their individual code based on the following pattern S12P (Student number 12 from Poland) or S22I (Student number 22 from Israel).

## Procedure

Data for this study were collected during three stages (Table 3). Stage 1 related to students' behaviours of a vlogging task. Stage 2 addressed classroom dynamics and interaction of students who participated in an online synchronous meeting. Stage 3 aimed to examine participants' and instructors' challenges of the telecollaboration. To categorise the data obtained through Vlogs and comments, tables comprising of the main lists of categories and themes selected in the study were utilised. A qualitative analysis was adapted to investigate participation of the Polish and Israeli students in the online synchronous meeting. Similarly, to the vlogging task, tables enabled to systematise the data. To collect the data, the researchers obtained "informed consent" (Creswell & Creswell, 2018, p. 147) from the participants (Appendix 1).

**Table 3**

*Stages of the Data Collection in the study including research instruments, techniques, and corresponding research questions*

Stage 1	Stage 2	Stage 3
Vlogs	Synchronous meeting on Zoom	Questionnaire
Content analysis	Observation	Questionnaire: part I – group work, part II – telecollaboration
RQ1: What behaviours do Polish and Israeli students exhibit while interacting with non-native group of students via Vlogs?	RQ2: What group dynamics can be observed when Israeli and Polish students meet synchronously online?	RQ3: What are the benefits and challenges of the Israeli-Polish telecollaboration?

### **Stage 1: Vlogs**

In stage 1 of the project, students were asked to present a series of three Vlogs to all group members. However, to save time of the participants, they were requested to watch and leave comments only under the Vlogs recorded by the non-native partners they were paired with (Figure 2). This part of the vlogging project was set to establish the partnership within the national and cross-national groups. The content was shared via platform Flipgrid, which is a tool supported by Microsoft that was approved by both universities participating in the telecollaboration. The access to the recordings was possible only after login in with the university e-mail address. Introducing this platform ensured security and guaranteed that the shared video material was visible only to the participants of the project. Students were asked to prepare 3–4 minutes long Vlogs on the topics provided by the teachers. Following Helm (2015), we purposely omitted too difficult (political issues) or too easy themes (music, sport, travel) and urged the participants to reflect on the following themes outlined in Table 4. The scope of the questions was broad enough for the participants to elaborate on the meaning they wanted to convey.

The participants were working with their native group while recording the three Vlogs. After publication of the video material on the Flipgrid platform, the non-native groups were to watch the Vlogs and leave a comment underneath. The platform enabled the students to comment on the shared material. Contrary to the Israeli students, who were obligated to provide written feedback (comments) for their non-native partners, the Polish students were only encouraged to do so. The main reason for the difference in the written feedback students

were to provide stemmed from the course requirements. The Israeli instructor planned the participation of the telecollaborative project as an obligatory assignment for her students. The Polish students, who were invited to participate in the project, were not assessed for the outcomes of the project.

**Table 4***The Instructions for the Vlogs*

Vlog 1	Vlog 2	Vlog 3
The aim is to get to know each other better both personally and as part of the group. Where do you live? What are your lives like? What do you do? Tell something more about your teaching practice/study. What are your informal education activities like?	How would you like to influence the world? What is important to you, related or unrelated to what you study or do in life? Why did you choose your field of study? What is special about your studies? What is it like to be young in Europe now in general and also regarding the political and economic and covert situation? Expand on what you said in the first blog regarding your work. Provide examples of what you do and how you do it. Tell us more about your work by giving examples of what you do.	Looking ahead at your future and your continued professional development (You and the future) What would you like to do in the future? Describe how you see your professional self in 5–10 years from now. What would you say to your future self?

The video material and the written comments were the main sources of data for this stage. We used “thematic analysis” (TA) (Braun & Clarke, 2012) to identify lists of categories (Appendix 2) which gave “insights into patterns of meaning (themes)” (p. 57) in the data. The decision to use TA in the study was justified by the fact that we were not interested in “quantitizing” the data (Kawulich, 2005) by looking for frequencies in the lists of codes and themes, but we wanted to focus on the observed data of the verbatim conversation to explore the group dynamic processes. By distinguishing the main themes that were addressed by the students, we were able to describe the discourse of the students performing the asynchronous online task. The comments were coded similarly by distinguishing the main patterns of behaviours.

**Stage 2: Online Synchronous Meeting on Zoom**

Stage 2 of the project involved two online synchronous meetings via platform Zoom. The participants were divided into two teams according to the groups they belonged to (Team 1: groups 1–9, Team 2: groups 10–18). After the introductory part, each group was delegated to a separate breakout room. The researchers could freely visit every room during the process of completing the

task. To facilitate more effective communication, each group received a link to their individual Padlet board on which three assignments were displayed, which were to help guide group discussions. (1) **Discuss:** What are the similarities and differences that you have discovered between yourselves through this project? (2) **Explain:** What is the key takeaway of the telecollaborative project? (3) **Evaluate:** Summarise the outcomes and celebrate the end of the project. Use filters to express your emotions. The last part of the meeting was devoted to conclusions and “the celebration part,” during which the students and the teachers could use filters as a way of expressing their emotions after the meeting.

Some selected groups were randomly observed to distinguish how students managed group work in the virtual meeting. Unstructured observation was used as it could provide the researchers with “rich contextual information” (Cohen, Manion, & Morrison, 2018, p. 542). It was conducted with the premise that some patterns of group interaction in the online environment could be detected. The aim was to reveal learners’ behaviours, making it possible to form conclusions on the challenges learners faced. The observation, recorded in field notes (Appendix 2), aimed to establish the appropriate categories and themes. It was done mainly on a “descriptive level” (Cohen, Manion, & Morrison, 2018, p. 552) by focusing on “reconstructions of conversations” (p. 553). It was believed that the researchers would be able to gather *in situ* data. The “reflective part” of the field notes was filled with the instances of “failed communication” and the instructors’ comments.

### Stage 3: Questionnaire

A quantitative approach was employed as the method which allowed for summative analysis of students’ perception of telecollaboration. It was expected that the questionnaire would provide insights into students’ reservations and perceived benefits of the project. In total, there were 27 questions in the survey developed by the M.A. student-researcher, for which a Google Form was used. The design of the questionnaire (Appendix 3) was based on Cantwell and Andrews’ (2002) research instrument used in their study investigating factors underlying students’ feelings towards group work. Apart from that, Helm’s (2015) survey on students’ beliefs about values of telecollaboration was exploited. Twelve questions, adapted from Helm (2015), assessed students’ perception of the telecollaboration. The researchers highlighted the anonymity of the students and voluntary participation in the survey.

Apart from the introduction, in which the purpose of data collection and its use had been stated, the questionnaire consisted of three parts. The first part was an introductory section in which students provided their bio data. In the second section, participants were requested to indicate their preferences towards

group work. This part consisted of 12 questions, which were grouped by the factor they investigated. There were three factors that were studied (1) Preference for group learning, (2) Discomfort in group learning, and (3) Preference for individual learning. Students were asked to rank the given statements using the Likert scale (where 1 meant “not at all true of me” and 5 was described as “very true of me”). The third section addressed students’ perception of the telecollaboration. Specifically, students were asked about challenges and benefits of this form of learning, areas of language learning that improved in the telecollaboration, and their views on the ICT tools utilised in the project. The questionnaire was shared with the students after the synchronous meetings on Zoom via chat box. Additionally, the link was also sent via e-mail to the Polish and Israeli participants. All answers submitted were fully anonymous.

## Findings

As mentioned previously, the study considered three data sources gathered in the project process between November and February 2022. The researchers investigated the Polish Israeli telecollaboration in the context of higher education. Firstly, the results of the video recordings on Flipgrid are presented. Secondly, observations of the group dynamics during an online meeting on Zoom supported by Padlet are summarised. As the last stage the study outlines the results of the questionnaire.

### **Stage 1: Thematic Analysis of Vlogs on Flipgrid. Focus on levels of ‘Failed Communication’**

By means of “thematic analysis” (TA) (Braun & Clarke, 2012), the decision was made to identify lists of categories and develop themes which corresponded to instances of “failed communication” (O’Dowd & Ritter, 2006). We were interested in whether or how the participants managed to overcome their communicative barriers in the virtual exchange. As displayed in Table 5, Vlogs #1 generated significantly more views and comments, which was probably due to the novelty of the online task and increased interest in knowing about the foreigners. The following Vlogs #2 and #3, spawned less online interchange of ideas. The overall time of the video material was also longer than in the next two Vlogs, which might have been caused by providing the students with a longer period for interaction than two other Vlogs. After the publication of the Vlogs, students communicated their feedback through comments.

**Table 5***A Summary of the Vlogging Activity on Flipgrid*

	Vlog #1	Vlog #2	Vlog #3
Hours of discussion	73.3	51.2	28.6
Number of views	2299	935	519
Number of comments	111	90	45
Number of responses	43	43	44

*Note:* The statistics were automatically generated by the Flipgrid platform.

The first, *individual*, level of failed communication was noticed in Vlogs #2. The first theme that emerged here concerned misunderstandings caused due to lack of intercultural knowledge. In the vlogging task, the students indicated the differences and similarities of living in the two cultures. As far as the differences were concerned, they were related to the Israeli community and movement. One Polish participant explained: “your lifestyle isn’t something we’re familiar with (actually, the first time we’ve come across it), but the idea of equality and socialism seems fun.”<sup>3</sup> This comment sparked a short conversation, in which an Israeli student sent a link to the Wikipedia page ([https://en.wikipedia.org/wiki/Hashomer\\_Hatzair](https://en.wikipedia.org/wiki/Hashomer_Hatzair)) that provided general information about the youth movement. Another student (S29P) added: “I would love to know more about these communities you live in—I don’t think such a thing is available or popular in Poland and it sounds really interesting.” On the other side, the Israeli group found it difficult to understand the reasons for studying English. Some Israeli students expressed their confusion with the specific course of study chosen by Polish students (S13IS): “hey guys, I don’t understand why you don’t want to be teachers? and what would you like to do with tourists?.” Further, another Israeli student commented (S16IS): “it sounds like you chose your studies not by passion but by default, please correct me if I’m wrong.” In their responses, the Polish students were not that sure why they chose English. Some mentioned the fact that they study English because they “have always liked the language,” others that they were “good at it” or got “good grades,” some mentioned the fact that they did “not know what to do in their future.”

Another level of failed communication occurred at *classroom level*. It indicated how matching of the participants influenced group dynamics. The Israeli participants more frequently decided to record their Vlogs #1 as a whole group, while the Polish participants preferred to record the clips separately and combine them into one video material. The fact that most of the Vlogs were presented in the latter form could be ascribed not only to the social distancing

<sup>3</sup> Respondents' remarks are quoted in the original.

restrictions introduced in both countries because of COVID-19 pandemic, but also to the participants' preferences for group work. For the similar reason, many Polish students recorded their clips indoors individually, unlike some of the Israeli participants who chose to work outside. To provide insights into the level of failed communication between the students, two most commented vlog entries were chosen to report on the behaviours appearing in the interactions. In the Polish group (Gr6PL), consisting of four students, each person recorded the videos separately to combine the three clips in 1 minute 34 seconds long Vlog. The students mentioned their specialisations and described their interests. They were mainly mentioning literature, music, cinematography, and pets. As far as comments were concerned, the students received several questions from their non-native partners: "Why did you choose to study languages?," "What do you like about your studying?," "What are your dreams after you finish studying?," "Did you guys always knew you would study in the university after finishing high school?," To these questions the Polish students answered in the comments section three days after the feedback had been posted. The themes that emerged in the Vlogs #1 were related to norms and ideas implicitly conveyed within both groups, which were not understood by the opposite nation. Conversely, in case of the Israeli's (Gr4IS) most commented Vlog, the group recorded the video altogether sitting outdoors and listening to each other's utterances. This Vlog was significantly longer, lasting 4 minutes and 30 seconds. The Israeli students first described their free time activities (rock climbing, doing various sports, learning science, philosophy, cooking, watching movies, listening to music). Then the students explained that their current activities related to "the movement" they belong to. Some questions from the Polish students were added, however, no answers were provided by the Israeli group: "You sound passionate about the movement :D Are there any frustrating setbacks along the way that you have to overcome?".

The third level of failed communication signified workload and denoted challenges at the *socio-institutional* level. Contrary to the Israeli students, who commented sections in all groups, the Polish students tended to leave comments only under the group they were assigned to. In Vlogs #1, 54% of the Polish students wrote a comment under the Vlog of their Israeli group members, whereas 94% of the Israeli students commented under the Polish students' Vlog, excluding responses to the comments. In Vlogs #2, on average, 87% of the Israeli students left a comment under the Polish videos, while 73% of the participants from Poland posted feedback for the Israeli groups. The last Vlogs #3 generated only 45 comments, including 22 comments from the Polish students, 20 Israeli posts and three comments from the coordinators. This means that almost 48% of the Polish students posted a comment under their non-native group Vlog, while almost 31% of the students from Israel provided written feedback. Comments often referred to the content of the Vlogs. There were

usually requests for further elaboration on their hobbies (music, movies), the studies (English), professional life (the Israeli community, youth organisations, non-formal education), or the world-wide situation (the pandemic restrictions). Most of the questions concerned asking for more details and finding the things that the participants had in common. The Polish students especially, preferred to provide a written reply to those questions, whereas the Israeli often recorded their answers.

The fourth, *interaction* level, referred to hidden messages that could have been understood only when the participants were acquainted with the cultural differences or socio-political situation in Israel. The Polish students could not understand what is meant by living in the Kibbutz community. The Israeli discourse itself was unclear as it was not supported by examples or justification of the Israeli students' utterances. A few illustrative examples are given below:

S22IS: we do it together because we think that when we do something together it makes you feel better, and you feel that you have a lot of power

S63IS: we are very serious about it

S51IS: we are trying to pass on these things that we established here as the movement

For the reasons, the Polish students started commenting or asking additional questions (S04P): “[w]e have never heard of a community like yours, it sounds very interesting. The idea behind informal education seems intriguing as well. Looking forward to getting to know more about you and your culture!,” Another Israeli student (S06IS): commented on different characteristics of their countries and juxtaposed travelling opportunities, stating: “I’m a little bit jealous of the easiness you can cross borders in the EU. In our country it is more complicated but last year I’ve been in Sinai dessert, and it was great, but I don’t think it’s similar to your travelling options.” The utterance was not supported by any example or justification of what makes travelling “complicated.” Unless the Polish students had known about the unstable political situation in Tel Aviv, they would not have understood what was meant by the message.

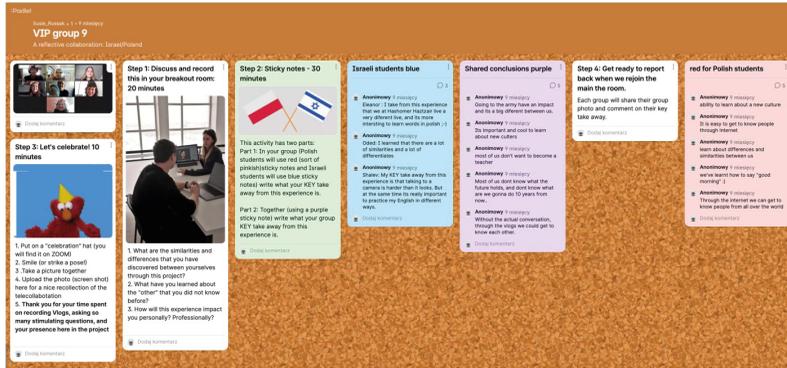
## **Stage 2. Observation of Online Synchronous Meetings via Zoom and Tasks on Padlet**

The second data source concerned online synchronous meetings, which provided information on challenging aspects of the online environment concerning students' behaviours while communicating directly in real time. The observations presented below indicate different challenges each of the groups faced.

Observation 1 concerned reticence which generally refers to not revealing one's thoughts or feelings readily (willingly without difficulty). During the synchronous online meeting, there were a few participants who appeared to be confident speakers while others preferred to refrain themselves from talking in the forum of the group. Being asked about the specialisation of their studies, the Polish students in group 8 were reluctant to elaborate on the topic. Similarly, one of the participants from Poland evaded a question about their hometown with a short response: "there's nothing much to talk about our town" to which Israeli student answered that "our towns are even smaller." In group 12, one Israeli student highlighted the problem of expressing oneself in the foreign language: "I was surprised by how hard it is to explain my thoughts in English. I think I have good English and I thought it would be easy but it's really hard to do the Vlogs to explain myself. I'm not sure if I had said the things I really wanted to say." The Polish student responded that they "do not face similar challenges" and the conversation was finished after the statement. Interestingly, some of the observed groups felt the need for introducing small talk during the online meeting via Zoom. They conducted an icebreaking game which allowed them to build team bonds. The activity that reappeared in the discussions functioned as a way of establishing relaxed atmosphere. On the other hand, the groups who immediately began to complete the tasks in Padlet seemed to feel unsure how to proceed with tasks and were rather apprehensive about expressing one's opinion.

Observation 2 referred to the lack of ICT skills. During the meeting on Zoom certain unanticipated problems appeared that influenced students' participation in the online meeting. The first concerned the use of Padlet. The fact that some students were not familiar with the virtual board, impeded the completion of the task. In group 8, to manage workflow in Padlet (Figure 2), one student in the group edited the posts and administered the board in the application. However, the group started to write their answers to the tasks in the Padlet intended for a different group. Another major obstacle concerned the use of a particular feature that Zoom offers, namely the use of filters. Many students could not use this embedded feature. Only after the help of other group members and the administrating teachers, some students managed to understand the filters option. The lack of basic ICT skills hindered the group activities.

Figure 2

*An Exemplary Padlet of the VIP Project*

The following examples show how technology can mediate for worse participation in the group. During the meeting in group 9, certain technical difficulties occurred. Firstly, due to poor Internet connection some students faced problems with connecting to Zoom and switching their cameras on. It influenced the group discussion as the students were not able to see each other's reactions during the task. Another technical problem that hindered the communication and the participation process related to the quality of the Internet connection in group 12. Poor Internet stopped group work when the video recording of the meeting ended midway. One of the Israeli students could not fully participate in the meeting as the sound was interrupted by external voices, consequently, his microphone was muted, and the camera was disconnected.

Observation 3 concerned participant responsibilities within the groups and their levels of engagement. Disconnected cameras caused problems in group 17 and group 18. It was especially visible in the case of the latter group, in which one of the Israeli students did not turn the camera on and did not participate actively in the meeting, leaving only one Israeli student to contribute. The visible awkwardness in the interaction between the students was emphasised by long pauses (e.g., no volunteer to initiate the discussion) and nervous laughter of participants. What is more, the group finished the discussion tasks relatively quickly compared to the rest of the groups. The work on Padlet also generated other challenges, for example, one student who appeared to be a leader commented that he would have preferred all the members to contribute to the posts on Padlet instead of him individually writing the answers. However, the most important observation was noticed thanks to the minutes of the meeting in Group 8 (Appendix 2). Even though the students were willing to communicate, for some reasons the discussion was devoid of effective language functions, mainly *justifying*, *explaining*, and *exemplifying*.

## Challenges of the Telecollaboration from the Students’ and Instructors’ Point of View

Although the telecollaborative project was successful, challenges at various levels remained to cause problems both to the participants and the instructors. Table 6 presents the challenges of the online virtual exchange that were collected from the questionnaire and observed throughout the telecollaboration process. The data were collected by means of *direct observation* which was based on the criteria of (1) introducing non-intrusive observation, (2) providing transparent information to the participants, (3) including subjective and objective information, (4) monitoring the online environment to increase ecological validity, (5) conducting observation over an extended period of time. To fully exploit *field notes* (Appendix 2) taken during the telecollaboration process, coding categories and themes that emerged during the first stage of the project were juxtaposed and then compared with descriptive and reflective notes. The observers’ comments in the reflective notes section usually concerned questions like “What are the challenges students can find in everything,” “What are the expectations of life?,” which were directing the researchers’ attention to a low quality of the spoken discourse. It was evident that lack of well-developed communicative skills, mainly language functions, prevented the participants from effective interaction and mediation of meaning (Council of Europe, 2020).

**Table 6**

*Challenges of the Telecollaboration from the Students’ and Instructors’ Point of View*

Learners’ challenges	Instructors’ challenges
Personal differences PL/I (age difference, sense of community, different studies, language level differences B1/B2 vs B2/C2).	Technological and institutional restraints (MS Teams—no chat for visitors, obtaining permission to use Zoom, Padlet—the Polish instructor had access to only 3 virtual boards, Zoom—no filters for guests).
Extra work (Antoniadou, 2011) – time constraints; – other commitments (children, jobs); – group responsibility.	Partner matching (the instructors could have matched students studying similar subjects, having similar interests, similar age groups/proficiency levels, 6–8 students in one group).
The tasks (Vlogs, Comments, Task no. 3 “Me and the future,” Padlet). The choice of adequate comment/questions.	Size of the population (difficult to manage a group of 100 participants, e.g., Zoom breakout rooms).

---

Inability to learn from work-related experience Lack of activation of reflective skills.	The choice of tasks (when you do not know the participants well, e.g., what their interests are, future work plans).
Disengagement or various levels of engagement (Yang, 2020) Shyness (Buss, 1985) Reticence (King, 2013) Lack of self-confidence.	Time constraints — Differences in time zones for synchronous meetings; — Challenges with scheduling meetings; (too close to the exam session for the Polish students); — Making efforts to adjust time frames (assigning tasks to keep the students occupied); — Amount of time preparing the telecollaboration (from the learning outcomes to the specific tasks).
A need for f2f, personal contact, live communication rather than Vlogs.	ICT support (not possible to find one universal platform, which would include the Zoom, Padlet and Flipgrid functionalities).
The need to identify self-initiated topics (Kohn & Hoffstaedter, 2017).	Monitoring students' on-task behaviour (Gillies & Boyle, 2010).
Lack of deep reflective skills; Ineffective use of academic language functions or perfunctory interactions.	Engaging students in "diversity surfing" (Kramsch, 2014, p. 98).
Psychological tension connected with recording oneself (O'Dowd, 2006).	Establishing teamwork beliefs and behaviours (Gillies & Boyle, 2010).
No individual learning flexibility (Kohn & Hoffstaedter, 2017).	Establishing group norms (Ruys, Van Keer, & Aelterman, 2012) (the Zoom etiquette).

---

## Learners' Challenges

When looking at the content of Table 6, it seems that most of the challenges the students faced related to their needs to feel safe and comfortable while participating in the virtual exchange. The first troubling issue emerged when recording the video logs on Flipgrid. The Polish participants felt tension and voiced their concerns with the fact that "there will be other people who will see them [the videos]." The learners often found it difficult to follow instructions for clarity. As it was in the case of vlogs #1 and #2, some participants hurried in answering the questions from the next task instead of recording them under the vlogs provided. This might have been because they were stressed due to video recording (O'Dowd, 2006), could not elaborate on their utterances

because of lack of linguistic competences, had nothing to say, or were not able to control the cognitive demands of the discourse. All students opted for more personal and face-to-face communication and wanted to discuss other topics than the ones chosen by their instructors. Other problems usually concerned extra work (Antoniadou, 2011) and various levels of engagement (Yang, 2020). As noted by Belz (2002), the differences in university schedules and other academic activities not related to the project could have limited students' contribution to telecollaborative tasks which resulted in frustration of some group members.

Another challenge concerns stepping out of the comfort zone. Some students avoided or did not elaborate on real-life world problems connected with political, ethical, and social dilemma, even though some of the topics appeared in the online discussion (human rights, gender roles, environmental sustainability, climate change). Some students questioned the rightfulness of the tasks prepared by the instructors, claiming that they should be allowed to talk about trivial matters, instead of their future or profession. Written reflections also appeared to be of a questionable depth and quality, especially when the Polish students rightfully protested “[h]ow to reflect on experience if we do not have one.” The lack of work-related activities in the academia prevented the students from elaborating on their “professional self.” Some students might have felt threatened by being asked to talk about experience not yet known. However, the missing aspect of experiential learning in the higher education context might be an important constraint of the university programme objectives.

The Polish students often admired high confidence of the Israeli students. The re-occurring statement, “I admire your confidence in speaking English” may be linked to their feeling of lack of self-confidence when speaking English that may have its roots in shyness or reticence. This finding is not consistent with that of Kohn and Hoffstaedter (2017) who studied “speaker identity,” “including a growing sense of speaker satisfaction and trust in one’s own creativity and strategic resourcefulness” (p. 1). Focusing extensively on one’s shyness in the context of the telecollaboration might have produced inaccurate perception of the intercultural encounter that accounted for limited understanding of the others.

## Instructors' Challenges

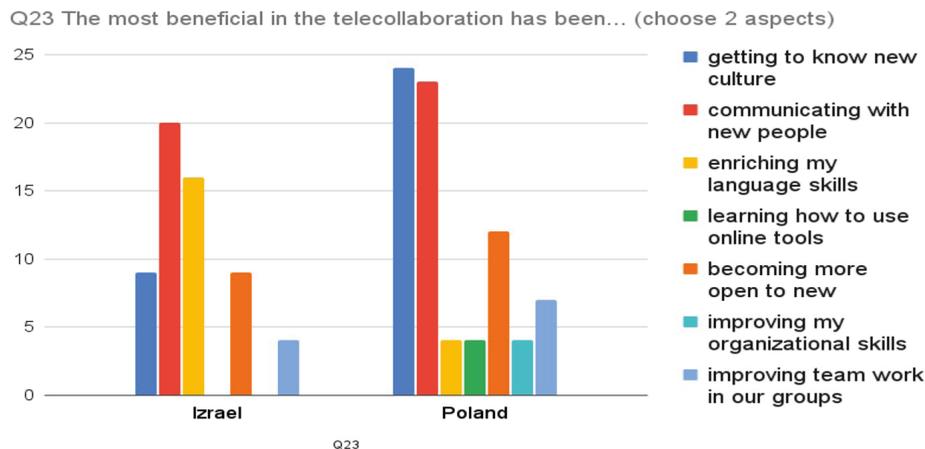
Both instructors encountered considerable difficulties in the telecollaborative process. First, *scheduling* seemed to play an important role. The instructors started planning in the summer 2021, however, they did not manage to coordinate the class schedules. Even though it did not impact the Vlog stage,

it did impact the synchronous meeting. Timing during the semester (too close to exams for the Polish students) decreased the quality of interaction in Vlog #3. On the other hand, time differences for the synchronous meetings were causing many problems for the Israeli students, who at the time of the meetings, were busy engaged in other obligations. Second, *institutional* (mainly administrative) *challenges*, which resulted from the Polish university policies in getting permission from its authorities to use Zoom. The Polish University uses MS Teams platform on which guests are denied access to the chat discussion. The inability to fully integrate guests from the outside of the Pedagogical University caused organisational burden. Third, *technological restraints* prevented the instructors from providing the participants with good quality virtual exchange. Apart from such obvious problems of echo, frame-freeze, muted microphones in the online synchronous meetings (Helm, 2015), there were those connected with functional disability of using certain ICT features on (disabled MS Team's chat and no filters on Zoom for guests). Finally, all *organisational* problems, often unanticipated, made the telecollaboration even more perplexing. They related to managing the course and the syllabus, choosing tasks that would be engaging and cognitively challenging for the participants, or establishing group norms online, including the Zoom etiquette.

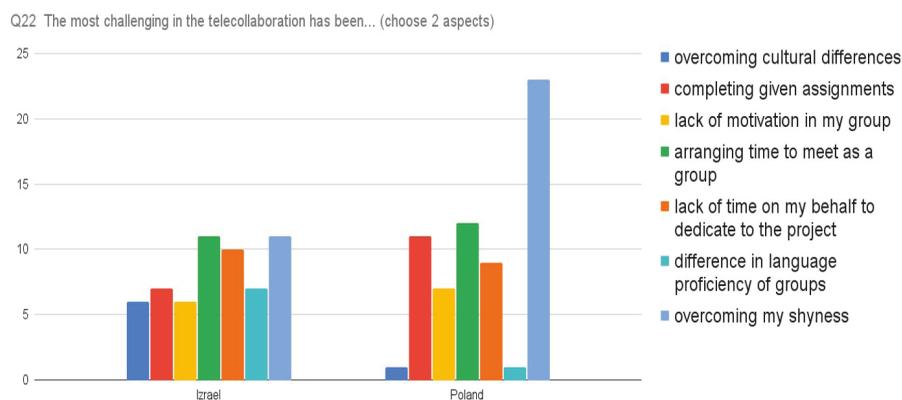
### **Stage 3. Questionnaire Answers Concerning Benefits and Challenges of Telecollaboration**

The researchers obtained  $N = 63$  responses from the questionnaire.  $N = 34$  belonged to the Polish students and  $N = 29$  to the Israeli participants. In the following paper, we decided to show the data that referred only to research question three. The complete summary of the study findings can be found in the final report of the project (Rążewska, 2022).

As regards the benefits of telecollaboration (Figure 3), both Israeli and Polish students recognised that communicating with new people was an asset. This component was mentioned by 67.6% responses of the Polish and 68.9% of the Israeli students. For 70.5% of the Polish students, the advantage of telecollaboration was getting to know new culture, whereas for the Israeli students this answer was voiced by 67% of the respondents. Over 55% of the Israeli students, who filled in the questionnaire, suggested that telecollaboration was beneficial because it enriched their language skills. Conversely, this advantage was indicated by only 11.8% among the Polish group.

**Figure 3***Perceived Benefits of Telecollaboration (Question no. 23)*

In addition to the indicated benefits, the students recognised skills that had improved after the telecollaboration. A common response among both Israeli and Polish students was that speaking advanced during the telecollaboration. Over 50% of all the answers mentioned this language competence. Similarly, students from both countries observed improvement in listening skills, as almost 40% indicated the skill to be improved. In the next question, number 24, there were some students ( $n = 12$ ) who saw “no significant changes” in their language development.

**Figure 4***Challenges of Telecollaboration (Question no. 22)*

On the other spectrum, Figure 4 illustrates the most challenging elements in the telecollaborative project. Over 53% of the students who filled in the questionnaire indicated that the most difficult aspect of the telecollaboration was to overcome their shyness. What is interesting about the data in Figure 4 is that over 67% of the Polish students, who voiced their opinion in the questionnaire, mentioned overcoming shyness as the most challenging factor in the telecollaboration. The second obstacle, expressed in the answers by 36% of students, was their difficulty in arranging time to meet as a group. Overcoming cultural differences appeared to be a challenge for the least number of participants, representing 11% of all the answers.

## Discussion

The results of the study indicate that there were challenging but rewarding aspects of the telecollaboration. With respect to the first research question certain levels of “failed communication” (O’Dowd & Ritter, 2006) were distinguished. The video material allowed the researchers to formulate conclusions about students’ group dynamics in telecollaboration and how they managed group work in the online environment. As suggested by the previously mentioned research conducted by Jauregi and Bañados (2010) *Get to Know Each Other* (known as GTKY activities) should not be omitted in the opening part of telecollaboration. Based on the examples of the project, the importance of this step was confirmed by students’ engagement in Vlogs #1 which was fully devoted to personal introductions. In fact, throughout the course of the whole project, students tended to insert some questions that were related to the personal lives of non-native partners. Additionally, the number of views for the GTKY task was the highest compared to the subsequent materials. It was observed that students naturally led to the conversations concerning hobbies, free time activities, favourite movies, books, and pets. On the other hand, the last task (Vlog #3), concerning plans for the future, posed a visible difficulty to mainly Polish and a few Israeli students who frequently mentioned that they had never contemplated on their future before. The length of the videos posted also reflected the struggle with the task. Especially the Polish students tended to limit their answers to expressions of uncertainty.

Comments posted under the Vlogs concerned the first direct interaction between the Israeli and Polish students. Most of the students posted their questions or their thoughts regarding the content of Vlogs #1. The Polish students left fewer comments under the Israeli Vlogs in the first round of the Vlogs; however, they provided elaborated written answers for the questions posted

under the Vlogs. Another noticeable difference was that the Israeli students preferred to provide their feedback in the video format. It may indicate that the Polish students preferred written communication; conversely, the Israeli students favoured oral communication. This difference could also be related to the *socio-institutional* level (workload or time constraints), differences in proficiency levels or students' individual preferences for using spoken or written discourse. This may be further discussed in the light of the fact that the Polish group was composed of people whose main intention was studying English whereas the Israeli students were enrolled for different courses.

It was found that looking for similarities and differences strengthened a sense of belonging to the groups as the comments developed a form of information exchange. Moreover, discovering shared roots may have destigmatized "the others." However, the most visible difference between the Polish and Israeli students appeared in the "willingness to communicate" (WTC) (MacIntyre, Dörnyei, Clement, & Noels, 1998) and their self-confidence (Altunel, 2021). The finding confirmed previous studies that level of engagement (Yang, 2020) observed among the participants was hindered by individual students' shyness or reticence (King, 2013; Shea, 2017), which might have had an impact on the collaborative learning. It may imply that the Poles felt less comfortable than the Israeli in voicing their ideas and communicating within the established groups. One unanticipated finding was that the instances of "failed communication" may have been caused by inability to use language more effectively due to lack of academic language functions or rhetorical devices. For these reasons some of the participants may have been more reserved than responsive during the telecollaboration.

The second phase of the study sought to answer research question two which addressed the group dynamics of the Polish and Israeli students. The part of the study referred to online telecollaboration during the synchronous Zoom meeting and focused on actions, reactions, and obstacles encountered by the students. Meeting synchronously and collaborating in the online environment required structured and open tasks, but above all the icebreaking activity. The importance of *Get to Know Each Other* (or GTKY) was highlighted by Jauregi and Bañados (2010) as this type of task introduces students to the online environment of collaboration. Furthermore, this finding is consistent with that of Helm and Guth (2010) who address this stage as Information Exchange, which allows for building connection between participants.

As indicated by Helm (2015), *technology* creates unprecedented opportunities for success or failure of telecollaboration. From the observation it was implied that possessing a good command of ICT skills was crucial in the successful telecollaboration development. Both participants and administrating teachers were required to employ their knowledge of Flipgrid, Zoom, and Padlet. It appeared that some students were not familiar with the virtual board

and were not used to using some features on Zoom. Similarly, past research on the topic indicated that the complexity of telecollaboration often discouraged students from contributing (Fuchs, 2016; Turula & Raith, 2015). Consequently, the use of three different platforms (Flipgrid, Zoom, Padlet) may have had a detrimental effect on the participants' feelings of comfort and online safety.

One of the key aspects influencing telecollaboration in the online environment could be linked to the students' level of engagement (Yang, 2020) within the groups online. It appeared that participants' digital presence in the meeting was crucial in the successful communication within a group. There were students whose cameras were not connected, and these students tended to be withdrawn from the discussion. Therefore, it can be assumed that being visible in front of the camera during the synchronous telecollaborative meeting allowed for a more effective communication between group members. However, when the number of students from Israel and Poland was equal or comparable (as it was in the case of groups 18, 8, 6), the groups tended to be visibly more active in the written discussion. These cross-national groups usually consisted of six to eight members. The same groups had no problems with participating in the online discussion on Zoom or performing the tasks on Padlet. This may signify that the most optimal number of group members is six to eight, provided that there is equal or similar number of representatives of two nationalities. This finding is consistent with that of Harmer (2015) who claimed that the most optimal number of learners in a group is five to eight. However, in the current study equal number of cross-national students allowed for stronger contributions from the participants and simultaneously introduced a variety of opinions.

With the reference to the previously mentioned research, on collaboration (De Florio, 2016; Olsen & Kagan, 1992) it appeared that some students when dividing the work within the groups depended on a single leader-like member who usually was in charge of sharing the screen and editing posts on Padlet. This, as reported by Greenfield (2003), may have shown mismatches between collaborative online activity and individual outcomes. Following Forsyth's (2019, p. 267) comment on leadership, groups nominate a leader in four situations: (a) when group members realise that success of group task is reachable, (b) the success of the task is considered gratifying, (c) for the assignment to be completed group endeavour is necessary instead of individual effort, (d) one group member is experienced in terms of being a leader. It needs to be stressed that the Israeli students came to the project with a sense of group well established, yet the majority experienced difficulties to share duties in the group. In case of the Zoom meeting, some Israeli students became leaders who tried to coordinate the work by suggesting ways of navigating the platform. This unequal division of labour hindered group work. Burdening one team member with the responsibility of writing all answers instead of collaboratively editing the posts may have negatively influenced group dynamics. This finding is consistent with

that of Forsyth (2019) who observed that the group members who sensed that they contributed to the group more than they received from other participants may have begun to limit their engagement. This might have been the case in the present study that the students preferred to delegate all the responsibilities to one participant and limit their contribution to only voicing their opinions.

With the reference to the research question three, it was discovered that the most challenging aspect of the telecollaboration indicated by the students was overcoming their shyness. Both Polish and Israeli students mentioned this difficulty explicitly; however, significantly more Poles chose shyness as the biggest obstacle which points out to the difference between these two learners' groups. According to Buss (1985), shyness is defined as a feeling of uneasiness, restraint, and clumsiness in communicative situations with unfamiliar people. This phenomenon was observed in the case of two groups during the online meeting on Zoom where the Polish students seemed to feel uncomfortable expressing their thoughts in the group discussion and often struggling to recall the suitable vocabulary items. Although it was assumed that their English proficiency level was higher than the majority of the Israeli students, some Poles did seem to struggle in face-to-face interactions frequently, answering with a single word only or in very short sentences. The Israeli respondents also indicated that overcoming one's shyness posed a difficulty, but not nearly unanimously as the Poles. This finding, which is in line with studies conducted by Crozier (1997), revealed that shy learners avoided participation in groups as well as tended to underperform compared to less shy learners in the use of vocabulary and fluency assignments.

From the results obtained in the questionnaire, one of the biggest assets of the telecollaboration was the opportunity to communicate with new people. This is an important insight mentioned by Helm (2015) who claimed that communication was the core aim of telecollaborative projects. In the same study it was established that one positive aspect of telecollaboration was the improvement of intercultural communication skills and online communication skills (Helm, 2015). As far as our study is concerned, when asked to define the area of improvement, the students differed in their answers significantly. It was found that the Israeli students enriched their vocabulary, whereas the Polish students benefited in self-monitoring skills. This outcome is contrary to that of Tsakiran (2019) who found that the participants improved vocabulary, as well as writing and reading skills via telecollaboration. Our study does not confirm the findings. Similarly to the study by Hartwell and Zou (2013), video recording on Flipgrid allowed the students to compare their language performance with peers. Therefore, it may be assumed that vlogging helped improve the speaking skills of learners, as it was indicated in research findings of Mandasari and Aminatun (2020). In our study, self-monitoring skills of one's performance was the main benefit for the Polish participants. It may be assumed that the

students who operated on lower proficiency levels benefited in basic language skills such as speaking, listening, and vocabulary. For students of higher proficiency levels, the improvement concerning self-monitoring skills (e.g., error correction) seemed more relevant.

For the Israeli students the component concerning “getting to know new culture” seemed less important compared to the more linguistic aspect of “enriching one’s language skills.” For the Polish students, the improvement in linguistic aspect was not so significant. The discrepancy between the two nationalities can be attributed to the differences in English proficiency levels. In case of the Israeli group vocabulary range and grammatical structures were in accordance with standard of B1–B2/C1 whereas the proficiency level of Polish students was estimated to be at B2–C1. This gap could be the cause of less perceivable improvement in language skills of the Polish students since they were exposed to less challenging language structures and vocabulary. As suggested by Dörnyei and Ushioda (2011), maintaining optimal degree of challenge is one of the components of motivation in language learning.

### **Limitations to the Telecollaborative Project**

There are certain limitations to the telecollaborative project that need to be addressed. The first limitation concerns peer feedback to Vlogs. Students, who came from different environments, both educational and professional, may not have been ready or willing to engage in the spoken and written discourse. The participants were asked to give feedback for Vlogs in a form of comments; however, commenting was not an obligatory part of the project for the Polish group, as opposed to the Israeli. As stated in the Procedure section of the paper, the differences in the course requirements between the two universities may have had an impact on the validity of eventual outcomes concerning the agency of the students, which was either internal or external. We may speculate whether we made the right decision to confront and compare groups which were not comparable or confrontable because their members came from different educational and professional environments. However, according to Mezirow (2012), “‘finding one’s voice’ becomes a prerequisite for free full participation” (p. 78)—irrespective of communicative events—and it is dependent on one’s inclusion in the discussion. We agree with other researchers who claimed that without such willingness to communicate (MacIntyre, Dörnyei, Clement, &

Noels, 1998; Altunel, 2021), no “communicative learning”<sup>4</sup> (Hubermans, 1984) can take place. However, we do agree that matching learners with comparable linguistic competences may have improved the quality of the telecollaboration. The second limitation refers to the term ‘group’ used in the questionnaire which did not indicate the influence of collaboration in native or non-native group. For this reason, the students voiced their confusion with the reference to unclear terminology. Future studies may explain or specify which notions are examined. Another major limitation of this study was linked to the synchronous online meeting. It was acknowledged that the virtual board Padlet used during the online encounter might not have been fully introduced to the participants before launching of the task. Following the study by Turula and Raith (2015), when students are provided with more elaborate introduction and assistance in the use of Internet Communication Tools, their participation becomes more meaningful. In retrospect, had we included more specific instruction, additional data could have been revealed. Future studies may conduct a preliminary study to evaluate learners’ digital competences through the examination of their ICT tools. The Internet tools used in the project ought to have been explicitly explained in terms of their utility in each task. Even though the study at hand had its limitations, it provided an overview of the recent findings concerning observable patterns of group dynamics in a specific context of telecollaboration. Among the many variables of the research, different levels of language proficiency, different reasons for studying the language as well as computer literacy, had an impact on the study and its results.

## Implications for More Effective Telecommunication

Based on the literature review and in reference to the limitations of the conducted study, the following implications of telecollaboration can be formulated. They may serve as guidelines to establish patterns that represent effective practices in telecollaboration. The first implication refers to telecollaboration to be conducted “in order to support social interaction, dialogue, debate, and intercultural exchange” (Belz, 2002, p. 3) which may be realised by introducing content-specific learning objectives and offering the participants opportunities for negotiation of meaning. The second implication refers to the promotion of positive educational outcomes of such online projects in which partici-

---

<sup>4</sup>Hubermans (1984) distinguishes two domains of learning. The first one is *instrumental learning*. It is the kind of learning to control and manipulate the environment or other people. In contrast, *communicative learning* entails learning to learn what others mean when they communicate with us. This type of learning involves values, moral issues, feelings, and intentions.

pants feel safe and supported within their “Zone of Proximal Development” (Vygotsky, 1978). Third, effective task design of telecollaboration (Hauck & Warnecke, 2013) is crucial, which shall include elements of peer-negotiated tasks to provide participants with a heightened sense of their autonomy and agency. Instead of forming the groups and developing teacher-generated questions, it may be suggested to allow students more autonomy in deciding on group members as well as fulfilling the need to identify self-initiated topics (Kohn & Hoffstaedter, 2017). By allowing the participants to decide on the topics they would like to discuss and choose the groups they would like to work in, not to mention the choice of the media, the participants will feel more empowered (Rockwood, 1995). Fourth, within the scope of any projects, students should be provided with instructions on how to use ICT tools which will support them with challenging aspects of technology. Before introducing technology in the telecollaborative project, the teacher shall consider the functionality and privacy settings of ICT tools. Online apps allow for more authentic exposure, but on the other hand, the teacher shall recognise whether limited privacy settings would be a more suitable option (Guth & Thomas, 2010).

As it was in the case of our telecollaboration, successful or challenging virtual exchange may depend on the following:

- (1) introducing icebreaking activities and finding similarities and differences between the cross-national groups (task in Vlog #2; Zoom meeting);
- (2) inducing change in attitudes of curiosity and openness by dealing with problems explicitly (questions and comments to Vlogs);
- (3) discovering the complexity of the online interaction (the use of various platforms and their functionalities);
- (4) training in the effective use of academic language functions (*justifying, explaining, exemplifying, persuading, classifying, inferring*);
- (5) raising critical and cultural awareness connected with the current situation in the world (COVID-19, the Israeli youth movement);
- (6) raising self-confidence of the participants to help them overcome their shyness.

As far as political and social implications of the telecollaboration are concerned, the virtual experience proved in line with Byram’s model (1997) of ICC as the project allowed for the development of knowledge of “the other” social group as well as skills of interpreting misconceptions between the participants related to various aspects of life.

## Conclusions

The purpose of the current study was to determine challenges of the Israeli-Polish telecollaborative project. The data confirmed that communication barriers occurred at various levels, namely individual, classroom, socio-institutional, and interaction. The study contributes to our understanding of the group dynamics by discovering specific instances of “failed communication.” These findings suggest that in general English proficiency levels determine the area of improvement. This observation, however, sparks a further question of how a more effective use of BICS and CALP<sup>5</sup> (Cummins, 1979) fosters collaborative learning among different proficiency levels and age of participants. Especially, more attention could be given to training in an effective application of language functions used by the learners. Further studies are needed to estimate to what extent development of oratory skills through training in BICS and CALP may improve the quality of such online interactions and to conclude the quality of effective communication in telecollaborative projects.

The research has shown that for the Polish students overcoming their shyness was the most challenging aspect of the telecollaboration. Although various studies conducted so far discussed reticence in ESL in relation to affective variables like anxiety (Lee & Lee, 2020) or learners’ actual or perceived proficiency (Amengual-Pizarro, 2018), there were no findings concerning ways of overcoming shyness and reticence in online environment. This result suggests that further research on shyness and reticence as factors influencing group work in telecollaboration is required. To the best of the authors’ knowledge, there have been no studies to date on reticence of Polish students in an ESL classroom (Zarrinabadi & Pawlak, 2021). As advocated by Crozier (2005), such studies could “explore methods of helping students overcome their lack of self-confidence” (p. 33) as it is likely that connections exist between shyness, reticence or self-confidence of the students.

Through the observation of the online meetings, it appeared that students enjoyed synchronous work in groups more than video recordings. Consequently, teachers need to consider implementing the ICT tools that allow for synchronous communication. Giving students more confidence through face-to-face contacts instead of Vlogging may reduce the psychological pressures caused due to both a/synchronicity (mainly being recorded, shared videos). Furthermore, all tools introduced during telecollaboration ought to be explicitly explained before the beginning of the task. Also, monitoring group activities and reacting immediately to learners’ struggles is crucial in the process.

---

<sup>5</sup>BICS refers to *basic interpersonal communication skills*, whereas CALP to *cognitive academic language proficiency*.

The study contributed to our understanding of effective ways of raising intercultural awareness between the nations, as it was in the case of the Israeli-Polish project. Our study confirmed the importance of *Get to Know Each Other* element. It was noticed that to establish rapport between group members, it was essential to initiate the group task with such an icebreaking activity. The results of the study indicate that the encounters with the “other” culture allowed the learners—to a certain extent—discuss plans for the future (VLOG #3 “Me and the future”). It is assumed that the recorded Vlogs may have led to a retrospective process of self-reflection, provided that the participants had some previous work-related experience. It is, however, unclear whether and to what extent the misconceptions that were addressed during the intercultural encounter were in fact dispelled. A further study could assess the long-term effects of telecollaboration, especially in relation to breaking down stereotypes about socio-cultural learning environment. These results could add to the rapidly expanding field of mediation in language learning (Council of Europe, 2020).

This present study provided a deeper insight into the complex notion of group dynamics within telecollaboration in the context of learners’ and instructors’ challenges. It has shown that when language learning is accompanied with immersion in a new intercultural experience, stepping outside of one’s comfort zone can be empowering, provided that the participants are not inhibited by individual constraints, such as shyness or reticence, as it was in the case of the Polish students. Telecollaboration does not only develop the vast array of skills but to be successfully implemented, it requires trusted and reliable partners, institutional and ICT support, positive attitudes, and above all reflective approach to learning about oneself and from others.

## References

- Amengual-Pizarro, M. (2018). Foreign language classroom anxiety among English for Specific Purposes (ESP) students. *International Journal of English Studies*, 18(2), 145–159.
- Akiyama, Y. (2017). Vicious vs. virtuous cycles of turn negotiation in American-Japanese telecollaboration: Is silence a virtue? *Language and Intercultural Communication*, 17(2), 190–209.
- Altunel, I. (2021). Insights into EFL learners’ willingness to communicate in online English classes during the Covid-19 pandemic: A case study from Turkey. *Language and Technology*, 3(1), 13–20. <https://dergipark.org.tr/en/download/article-file/1666143>
- Antoniadou, V. (2011). Using Activity Theory to understand the contradictions in an online transatlantic collaboration between student-teachers of English as a foreign language *ReCALL*, 23(3), 233–251.
- Beatty, K. (2010). *Teaching and researching computer-assisted language learning* (2nd ed.). Pearson Education Limited.

- Belz, J. A. (2002). Social dimensions of telecollaborative foreign language study. *Language Learning & Technology*, 6(1), 60–81.
- Boss, S., & Krauss, J. (2007). *Reinventing project-based learning. Your field guide to real-world projects in the digital age*. International Society for Technology in Education.
- Buss, A. B. (1980). *Self-consciousness and social anxiety*. Freeman.
- Braun, V., & Clarke, V. (2012). *Thematic analysis*. In H. Cooper (Ed.), *APA handbook of research methods in psychology. Research designs*. Vol. 2 (pp. 57–71). The American Psychological Association.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.
- Chun, D., & Turula, A. (2015). Cognitive and social presence in task-based telecollaboration [Conference session]. In J. Colpaert, A. Aerts, M. Oberhofer, & G. C. Plana (Eds.), *Task Design & CALL. Proceedings*. Universiteit Antwerpen.
- Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education* (8th ed.). Routledge.
- Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications, Inc.
- Crozier, W. R. (1997). *Individual learners: Personality differences in education*. Routledge.
- Crozier, W. R. (2005). Shyness and students' perceptions of seminars. *Psychology Learning & Teaching*, 4(1), 27–34. <https://doi.org/10.2304/plat.2004.4.1.27>
- Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research*, 49, 22–51.
- Dooly, M. (2010). Teacher 2.0. In S. Guth & F. Helm (Eds.), *Telecollaboration 2.0. language literacies and intercultural learning in the 21st century* (pp. 277–304). Peter Lang.
- Dooly, M. (2017). Telecollaboration. In C. C. Chapelle & S. Sauro (Eds.), *The handbook of technology and second language teaching and learning* (pp. 169–183). John Wiley & Sons, Inc.
- Dörnyei, Z., & Murphey, T. (2003). *Group dynamics in the language classroom*. Cambridge University Press.
- Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Pearson Education Limited.
- Forsyth, D. R. (2019). *Group dynamics* (7th ed.). Cengage Learning.
- Fuchs, C. (2016). “Are you able to access this website at all?”—Team negotiations and micro-level challenges in telecollaboration. *Computer Assisted Language Learning*, 29(7), 1152–1168. <https://doi.org/10.1080/09588221.2016.1167091>
- Furstenberg, G., Levet, S., English, K., & Maillet, K. (2001). Giving a voice to the silent culture of language: The CULTURA project. *Language Learning & Technology*, 5(1), 55–102.
- Gabryś-Barker, D. (2012). *Reflectivity in pre-service teacher education. A survey of theory and practice*. Wydawnictwo Uniwersytetu Śląskiego.
- Gillies, R., & Boyle, M. (2010). Teachers' reflections on cooperative learning: Issues of implementation. *Teaching and Teacher Education*, 26, 933–940.
- Godwin-Jones, R. (2019). Telecollaboration as an approach to developing intercultural communication competence. *Language Learning & Technology*, 23(3), 8–28. <http://hdl.handle.net/10125/44691>
- Godwin-Jones, R. (2020). Building the porous classroom: An expanded model for blended language learning. *Language Learning & Technology*, 24(3), 1–18. <http://hdl.handle.net/10125/44731>
- Guichon, N., & Hauck, M. (2011). Teacher education research in CALL and CMC: More in Demand than Ever. *ReCALL*, 23(3), 187–199.

- Guth, S., & Helm, F. (2010). *Telecollaboration 2.0. Language literacies and intercultural learning in the 21st century*. Peter Lang.
- Guth, S., & Helm, F. (2012). Developing multiliteracies in ELT through telecollaboration. *ELT Journal*, 66(1), 42–51.
- Guth, S., & Thomas, M. (2010). Telecollaboration with web 2.0 tools. In S. Guth & F. Helm, (Eds.), *Telecollaboration 2.0. Language, literacies, and intercultural learning in the 21st century* (pp. 39–68). Peter Lang.
- Harmer, J. (2015). *The practice of English Language Teaching* (5th ed.). Pearson.
- Hartwell, M., & Zou, B. (2013). *A Chinese-French case study of English language learning via Wikispaces, Animoto and Skype* [Conference session]. In I. Bradley & S. Thouesny (Eds.), *20 Years of EUROCALL: Learning from the past, looking to the future*. Evora, Portugal (pp. 106–111). EUROCALL.
- Hauck, M., & Warnecke, S. (2013). Materials design in CALL: Social presence in online environments. In M. Thomas, R. Hayo, & M. Warschauer, *Contemporary computer-assisted language learning* (pp. 95–115). Bloomsbury.
- Helm, F. (2015). The practises and challenges of Telecollaboration in higher education in Europe. *Language Learning and Technology*, 19(2), 197–217.
- Helm, F., & Guth, S. (2010). The Multifarious Goals of Telecollaboration 2.0. Theoretical and Practical Implications. In S. Guth & F. Helm (Eds.), *Telecollaboration 2.0. language literacies and intercultural learning in the 21st century* (pp. 69–106). Peter Lang.
- Jauregi, K., & Bañados E. (2010). An international video-web communication project between Chile and The Netherlands. In S. Guth & F. Helm (Eds.), *Telecollaboration 2.0 language, literacies, and intercultural learning in the 21st century* (pp. 427–436). Peter Lang.
- Johnson, D. W., & Johnson, R. T. (2008). Cooperation and the use of technology. In J. M. Spector, M. D. Merrill, J. van Mreerienboer, & M. P. Driscoll (Eds.), *Handbook of research on educational communications and technology* (pp. 401–423). Taylor & Francis.
- Jolliffe, W. (2007). *Cooperative learning in the classroom. Putting it into practice*. Paul Chapman Publishing.
- Kawulich, B. B. (2005). Participant observation as a data collection method. *Forum Qualitative Sozialforschung/Forum: Qualitative Social Research*, 6(2), article 43.
- King, J. E. (2013). Silence in the second language classroom of Japanese universities. *Applied Linguistics*, 34(3), 325–434.
- Kohn, K., & Hoffstaedter, P. (2017). Learner agency and non-native speaker identity in pedagogical lingua franca conversations: Insights from intercultural telecollaboration in foreign language education. *Computer Assisted Language Learning*, 30(1), 1–17.
- Kramsch, C. (2014). Teaching foreign languages in an era of globalisation. Introduction. *The Modern Language Journal*, 98(1), 296–311.
- Krauss, J., & Boss, S. (2013). *Thinking through project-based learning. Guiding deeper inquiry*. Corwin.
- Lamy, M.-N., & Goodfellow, R. (2010). *Telecollaboration and learning 2.0*. In S. Guth & F. Helm (Eds.), *Telecollaboration 2.0. language literacies and intercultural learning in the 21st century*. Peter Lang.
- Lan, Y. J. (2020). Immersion, interaction, and experience-oriented learning: Bringing virtual reality into FL learning. *Language Learning & Technology*, 24(1), 1–15. <http://hdl.handle.net/10125/44704>
- Larsen-Freeman, D., & Anderson, M. (2011). *Techniques and principles in language teaching* (3rd ed.). Oxford University Press.

- Le, H., Janssen, J., & Wubbels, T. (2018). Collaborative learning practices: Teacher and student perceived obstacles to effective student collaboration. *Cambridge Journal of Education*, 48(1), 103–122. <https://doi.org/10.1080/0305764X.2016.1259389>
- Leavy, P. (2017). *Research design. Quantitative, qualitative, mixed methods, arts-based, and community-based participatory research approaches*. The Guilford Press.
- Lee, J. S., & Lee, K. (2020). Affective factors, virtual intercultural experiences, and L2 willingness to communicate in in-class, out-of-class, and digital settings. *Language Teaching Research*, 24(6), 813–833. <https://doi.org/10.1177/1362168819831408>
- Luchikhina, L., & Elfimova N. (2019). Telecollaborative project as a method of intercultural communicative competence development in a non-linguistic university (pp. 472–476). *Advances in Social Science, Education and Humanities Research*. Vol. 333. Atlanta Press.
- MacIntyre, P. D., Dörnyei, Z., Clement, 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.
- Mandasari, B., & Aminatun, D. (2020). Improving students' speaking performance through vlog. *English Educational Journal of English Teaching and Research*, 5(2), 136–142.
- Mezirow, J. (2012). Learning to think like an adult. Core concepts of transformation theory. In E. W. Taylor & P. Cranton (Eds.), *The handbook of transformative learning. Theory, research, and practice*. Jossey-Bass.
- Nokes-Malach, T., Zepeda, C., Richey, J., & Gadgil, S. (2019). Collaborative learning: The benefits and costs. In J. Dunlosky & K. Rawson (Eds.), *The cognitive handbook of cognition and education*. Cambridge Handbooks in Psychology (pp. 500–527). Cambridge University Press.
- O'Dowd, R. (2012). Intercultural communicative competence through telecollaboration. In J. Jackson (Ed.), *The Routledge handbook of language and intercultural communication* (pp. 340–356). Routledge.
- O'Dowd, R. (2015). Supporting in-service language educators in learning to telecollaborate. *Language Learning & Technology*, 19(1), 64–83. <http://llt.msu.edu/issues/february2015/odowd.pdf>
- O'Dowd, R. (2016). Emerging trends and new directions in telecollaborative learning. *CALICO Journal*, 33(3), 291–310.
- O'Dowd, R. (2017). Exploring the impact of telecollaboration in initial teacher education: The EVALUATE project. *The EUROCALL Review*, 25(2), 38–41.
- O'Dowd, R., & Eberbach, K. (2004). Guides on the side? Task and challenges for teachers in telecollaborative projects. *ReCALL*, 16(1). <https://doi.org/10.1017/S0958344004000217>
- O'Dowd, R. (2018). From telecollaboration to virtual exchange: Stage-of-the-art and the role of UNCollaboration in moving forward. *Journal of Virtual Exchange*, 1(23). Research-publishing.net. <https://doi.org/10.14705/rpnet.2018.jve.1>
- O'Rourke, B. (2007). Models of telecollaboration (1): eTandem. In O'Dowd (Ed.), *Online intercultural exchange: An introduction for foreign language teachers* (pp. 41–61). Multilingual Matters.
- Pfingsthorn, J., Kramer, C., Czura, A., & Stefl, M. (2019). The formation of professional identity and motivation to engage in telecollaboration in foreign language education. *European Journal of Language Policy*, 11(2), 143–165. Liverpool University Press.
- Raźewska, A. (2022). *Developing group work and peer cooperation through telecollaboration*. Unpublished M.A. Thesis. Pedagogical University of Cracow.
- Richards, J., C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching* (3rd ed.). Cambridge University Press.

- Rockwood, H. S. (1995). Cooperative and collaborative learning. *National Learning Forum*, 4(6), 8–9. <https://doi.org/10.1002/ntlf.10024>
- Ruys, I., Van Keer, H., & Aelterman, A. (2012). Examining pre-service teacher competence in lesson planning pertaining to collaborative learning. *Journal of Curriculum Studies*, 44, 349–379.
- Schunk, D. H. (2012). *Learning theories. An educational perspective* (6th ed.). Pearson Education.
- Shea, D. P. (2017). Compelled to speak: Addressing student reticence in a university EFL classroom. *The Asian Journal of Applied Linguistics*, 4(2), 173–184.
- Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, 52(1), 119–158.
- Taskiran, A. (2019). Telecollaboration: Fostering foreign language learning at a distance. *European Journal of Open, Distance and e-Learning*, 22(2), 87–97.
- Tuckman, B. W. (1965). Developmental sequences in small groups. *Psychological Bulletin*, 63(6), 384–399. <https://doi.org/10.1037/h0022100>
- Tuckman, B. W., & Jensen, M. A. C. (1977). Stages of small-group development revised. *Group & Organization Management*, 2(4), 419–427. <https://doi.org/10.1177/105960117700200404>
- Turula, A., & Raith, T. (2015). Telecollaboration insights: Learning from exchanges that fail. *The EUROCALL Review*, 23(2), 19–37.
- Vygotsky, L. S., & Cole, M. (1978). *L.S. Vygotsky: Mind in society: The development of higher psychological processes*. Harvard University Press.
- Ware, P. (2018). Telecollaboration. In J. I. Lontos & M. DelliCarpini (Eds.), *The TESOL Encyclopedia of English Language Teaching*. John Wiley & Sons, Inc.
- Yang, S. J. (2020). Affordances and challenges of telecollaboration for pre-service teachers. *Language Learning & Technology*, 24(3), 30–41. <http://hdl.handle.net/10125/44733>
- Zarrinabadi, N., & Pawlak, M. (2021). *New perspectives on willingness to communicate in a second language*. Springer. <https://doi.org/10.1007/978-3-030-67634-6>

## **Informed consent for participation in the research**

*Dear Student,*

You are currently participating in an international collaboration between students at the Pedagogical University KEN in Krakow, Poland, and students from Beit Berl College in Israel. In the framework of this collaborative course students exchange vlogs, participate in a synchronous online meeting, and write reflections. We would like to ask for your permission to use the contents of the communications for research purposes. Within that framework, we will not share any identifying information about students. All references to students will remain anonymous. Participation is on a voluntary basis. You may withdraw your permission for your data to be used at any time.

*If you agree to participate, please fill in the following consent form.  
Thank you*

## **Consent of students to participate in the study**

Surname and first name of the student .....

Specialty / year of study .....

I agree to participate in an international project and research on teaching English under the title “Vlogs between students in Israel-Poland: A Reflective Collaboration” (VIP-RC).

I consent to the use of:

1. Recorded Vlogs.
2. Reflections written by me.
3. The answers to the questions (in the form of a questionnaire and/or interview).
4. Entries on the university MS Teams platform.
5. Online video recording of the meeting (MS Teams).

I certify that I have been informed about:

1. the purpose of the research of a purely scientific nature.
2. the course and form of the study.
3. keeping my personal data confidential and encoding all my works.
4. the possibility of withdrawing from participation in the research project at any time.

At the same time, I acknowledge that this material will be used only for scientific purposes, and any data that could allow the identification of the respondents will not be made available to third parties, nor will they appear in the final study (report).

Student's signature

.....

Date:

.....

*Thank you very much for participating in the study!*

Appendix 2

## **Sample of Telecollaboration Observation Field Notes**

Platform: ZOOM + Padlet

Meeting focus: *Similarities and differences*

Number of Students present:  $N = 6$  ( $n = 5$  female:  $n = 3$  Polish &  $n = 2$  Israeli;  $n = 1$  male Israeli)

Role of the researchers: Observers

Date: 25 January 2022

Length of the breakout room session: 49'13''

FLIPGRID		ZOOM/Padlet		
Coding categories	Themes	Descriptive Notes (Reconstructions of conversation)		Reflective Notes
		SIMILARITIES	DIFFERENCES	
University life	exams, enjoyment, hardship, English, Hebrew, language,	<p>"Our willingness to learn about another culture."</p> <p>"Although from different cultures, some similarities shine through like our expectations of the future and what we want from life."</p> <p>"We all want to learn."</p> <p>"Just like in your language, there are words in Hebrew that don't have a replacement in English."</p> <p>"We are all students."</p> <p>"We try to speak English."</p> <p>"Having a shared language allows us to communicate with people who are different from us."</p> <p>"We all are willing to cooperate."</p> <p>"We can understand each other even with our differences."</p> <p>"We can find a challenge in everything even the smallest activities."</p> <p>"The way we arrange our time."</p> <p>"We are not ready to deal with the world."</p>	<p>"Unlike you, I do not want to be a teacher."</p> <p>"Going to the army have an impact and it's a big difference between us."</p> <p>"Our educational systems are very different."</p> <p>"Your country/what you study is completely different."</p> <p>"We learn differently."</p> <p>"Israeli students are more confident than we are."</p> <p>"...but I will never be a Vlogger I guess."</p> <p>"We treat students in youth movement differently that in our traditional schools."</p>	<p>Levels of "failed communication" (O'Dowd &amp; Ritter, 2006)</p> <p><b>The individual level:</b></p> <ul style="list-style-type: none"> <li>- learners' intercultural knowledge, stereotypes,</li> <li>- expectations (What are the expectations of life/family? *no examples given).</li> </ul> <p><b>The classroom level:</b></p> <ul style="list-style-type: none"> <li>- teacher-to-teacher relationship,</li> <li>- group dynamics,</li> <li>- matching of the participants</li> </ul> <p>(*no justification as to how the educational systems are different).</p> <p><b>The socio-institutional level:</b></p> <ul style="list-style-type: none"> <li>- course organisation,</li> <li>- workload and time constraints (*Why are Vlogs so problematic?; How do students arrange their time?);</li> <li>- assessment.</li> </ul> <p><b>The interaction level</b></p> <ul style="list-style-type: none"> <li>- cultural differences (Where do the Israeli Ss confidence vs Polish Ss resentment, shyness come from?);</li> <li>- learners' engagement (Ss mentioned disengagement in ref. to university life).</li> </ul>
Personal life	movies, books, (extreme) sport, hobbies			
Professional development	teacher profession, translators, and interpreters, youth movement, military			
Participation	social life, school milieu, (youth) movement(s),			
Behaviour	conflict resolution, positive attention, disengagement,			
Family and friends	spouse, friends, pets (dogs, cats)			
Miscellaneous	shyness, trust, connection, distance, protest, weather			

## Questionnaire in Google Forms

### Telecollaboration in higher education

Hello! My name is Aleksandra. I am a second-year student of MA studies, majoring in English Studies. Thank you for your contribution to this project I am conducting research on developing collaborative collaboration through telecollaboration (connecting two distant classrooms with the help of Internet communication tools, projects, collaborative tasks and group work). Through this Google Form, I would like to learn about your experience with collaboration in this project I would like to ask you to fill in this form and answer according to your thoughts. The form is fully anonymous and any personal information will not be published or shared with a third party. Thank you.

Where are you from?

Poland      Israel

How old are you?

18–25      25–35      35–40      40+

What is your gender?

Male    Female      I prefer not to say

1. I can usually understand other group members' ideas.

Not at all true of me 1    2    3    4    5      Very true of me

2. It is best when each person helps each other within a group.

Not at all true of me 1    2    3    4    5      Very true of me

3. Sometimes I feel nervous when I have to give my ideas or communicate within a group.

Not at all true of me 1    2    3    4    5      Very true of me

*\*The following questions refer to your experience your experience of group work*

4. I often find it difficult to understand what the group task is.

Not at all true of me 1    2    3    4    5      Very true of me

5. I like to work alone even when placed in a group.

Not at all true of me 1    2    3    4    5      Very true of me

6. It is important that other group members take responsibility for the outcomes of my learning.

Not at all true of me 1 2 3 4 5 Very true of me  
7. I usually make a strong personal contribution to group work.

Not at all true of me 1 2 3 4 5 Very true of me  
8. I often think the work becomes too confusing when done in a group rather than individually.

Not at all true of me 1 2 3 4 5 Very true of me  
9. I am often afraid to ask for help within my group.

Not at all true of me 1 2 3 4 5 Very true of me  
10. I like group work more when we can make up our own groups.

Not at all true of me 1 2 3 4 5 Very true of me  
11. Even when groups are well organised, I believe there are more effective ways of using class time.

Not at all true of me 1 2 3 4 5 Very true of me  
12. I sometimes feel let down by other group members.

Not at all true of me 1 2 3 4 5 Very true of me  
*\*The following questions aim to understand your perception of this telecollaboration in your learning*

13. Based on your experience with your peer natives how would you describe the telecollaboration?

14. Based on your experience with non-native participants how would you describe the telecollaboration?

15. Taking part in a telecollaboration project was a positive experience for me.

Strongly disagree 1 2 3 4 5 Strongly agree

16. I enjoyed using the internet communication tool (Flipgrid).

Strongly disagree 1 2 3 4 5 Strongly agree

17. I enjoyed meeting online via Zoom.

Strongly disagree 1 2 3 4 5 Strongly agree

18. I learned how to communicate better with people of other cultures by taking part in online exchanges.

Strongly disagree 1 2 3 4 5 Strongly agree

19. What have you learnt about the other culture?

20. I improved my foreign language skills by taking part in the telecollaboration.

Strongly disagree 1 2 3 4 5 Strongly agree

21. In the telecollaboration I developed skills which will make me more employable.

Strongly disagree 1 2 3 4 5 Strongly agree

22. The most challenging in the telecollaboration has been... (choose 2 aspects).

- overcoming cultural differences,
- completing given assignments,
- lack of motivation in my group,
- arranging time to meet as a group,
- lack of time on my behalf to dedicate to the project,

- difference in language proficiency of groups,
  - overcoming my shyness.
23. The most beneficial in the telecollaboration has been... (choose 2 aspects).
- getting to know new culture,
  - communicating with new people,
  - enriching my language skills,
  - learning how to use online tools,
  - becoming more open to new challenges,
  - improving my organizational skills,
  - improving teamwork in our groups.
24. What aspects of your language learning have improved after the telecollaboration?
- speaking and communicating,
  - listening,
  - vocabulary,
  - self-monitoring (error correction skills),
  - no significant changes in any of the aspects.

Sabina Nowak, Aleksandra Rażewska

### **Herausfordernde, aber lohnende Aspekte der Telekollaboration: Das virtuelle israelisch-polnische Projekt (VIP)**

#### Zusammenfassung

Ausgehend von der Prämisse, dass die Telekollaboration ein anspruchsvolles Unterfangen sein kann, ist anzunehmen, dass ein solcher virtueller Austausch positive Ergebnisse bringen sollte. Trotz umfangreicher Forschungsarbeiten zur Telekollaboration haben bisher nur wenige Studien die Gruppendynamik im Kontext des Online-Austauschs untersucht. In der aktuellen Studie wurde der Fall einer israelisch-polnischen ( $N = 100$ ) interkulturellen Telekollaboration analysiert. Bei den israelischen Teilnehmern handelte es sich um Studierende eines BEd-Studiengangs im Rahmen informeller Bildung und bei den polnischen um Studierende eines BA-Studiengangs in Übersetzung. Das Hauptziel des Telekollaborationsprojekts bestand darin, die englischen Sprachkenntnisse und die interkulturelle kommunikative Kompetenz (Byram, 1997) der Studierenden zu entwickeln. In der vorliegenden Studie sollte allerdings recherchiert werden, wie die Gruppendynamik den Telekollaborationsprozess und die Projektergebnisse beeinflusste. Die Studierenden tauschten asynchrone Videoaufnahmen (Vlogs) und synchrone Sitzungen über Zoom und füllten einen Online-Fragebogen aus. In dem vorliegenden Beitrag werden einige der Herausforderungen im Zusammenhang mit der internationalen Telekollaboration dargestellt, darunter technologische und institutionelle Schwierigkeiten, kontextspezifische Anforderungen und einige individuelle Unterschiede zwischen den Teilnehmern. Mittels „thematischer Analyse“ (Braun & Clarke, 2012) untersuchten die Autorinnen Ursachen der „misslungenen Kommunikation“ (O'Dowd & Ritter, 2006) und wie sie sich auf den Verlauf der Zusammenarbeit auswirkten, sowohl technisch als auch in Bezug auf den Inhalt

und die Entwicklung der Interaktionsfähigkeiten, um bestimmte Themen aufzudecken, die mit Herausforderungen der Lernenden und Lehrenden verbunden sind. Die Studie schließt mit pädagogischen Implikationen für eine effektivere Umsetzung von Telekollaboration in der Hochschulbildung.

*Schlüsselwörter:* Telekollaboration, Gruppendynamik, Vlogging, Zurückhaltung





**Emerson Case**

California State University, Bakersfield, USA

 <https://orcid.org/0000-0002-7299-8637>

**Agnieszka Kaczmarek**

University of Applied Sciences in Nysa, Poland

 <https://orcid.org/0000-0002-5770-0329>

**Sebastian Zatylny**

University of Applied Sciences in Nysa, Poland

 <https://orcid.org/0009-0006-6145-100X>

## **Reading, Discussing, and Sharing: Creation of a Vibrant Transnational Online Learning Community through the International Book Club**

### **Abstract**

The present paper reports on the findings from the first four years of an ongoing, longitudinal study which examines the experiences and perceptions of Polish college participants in the International Book Club, an extension of a campus common read program at an American university. More specifically, the paper examines participants' experiences with the book, both inside and outside the classroom, and examines their opinions about the issues discussed in each year's book selection, their attitudes toward that issue and whether or not their attitude toward the issue was changed through this experience, their attitudes toward public discussions and whether or not their attitude toward public discussions was changed through this experience, and whether or not reading the book and participating in the club helped to improve their English-speaking abilities. Quantitative data, in the form of a nine-statement survey, and qualitative data, in the form of a nine-question open-ended questionnaire, were collected and analyzed. While a great deal of the literature on common read programs has explored the experiences of participants within the North American context, there is a dearth of research outside of that context, including the Polish context. The present study attempts to ameliorate this situation. The main findings from the study show that, overall, participants enjoyed the experience of reading the book and that participation in the program has been beneficial for them, including getting them to read more, to engage in discussions more, to think more deeply about more diverse topics and issues, and perhaps most importantly, to improve their English language abilities.

*Keywords:* common read programs, International Book Club

## Common Reads and the One Book, One City Movement

Over the last 25 years, common read programs have proliferated and expanded at public libraries in towns and cities across the United States. As Rogers (2002, p. 16) points out, localized book clubs have long been a part of the fabric of public libraries in the United States. In 1998, however, he writes, the concept of localized book clubs of “small groups of diehard readers” was “amplified to the tenth power” when librarian Nancy Pearl of the Washington Center for the Book at the Seattle [Washington] Public Library initiated a city-wide book club, dubbed “If All of Seattle Read the Same Book,” where library patrons throughout the entire library system were encouraged to read in common the book *The Sweet Hereafter* by author Russell Banks.

In her brief article entitled “‘One Book’ Programs Span the Nation,” *Der Spiegel* points out that these programs, often referred to as “One Book, One City” programs, or alternatively as “One City, One Book” programs, “can focus on a single community or span whole cities” (2005, p. 19) and the idea has even been extended to include an entire state when Arizona launched its “OneBookAZ” program in April 2002 (Rogers, 2002, p. 16).

Nowadays, each fall, with the commencement of the new school year, students at colleges and universities across the United States find themselves joining in on this phenomenon through campus common read programs. Published in 2015, under the auspices of the National Association of Scholars (NAS), Thorne’s research (2015) showed 341 American colleges and universities assigning common read texts, with a total of 231 different books assigned, during the 2013–2014 school year. Randall (2019), following up on Thorne’s work through the NAS-produced publication *Beach Books*, found just a few years later that that number had grown during the 2018–2019 school year to 518 college common read selections at 475 institutions, including 45 institutions in the state of California alone, and could be found in all but three American states.

## Literature Review

### Campus Common Read Programs

So what, exactly, do campus common read programs entail? Laufgraben (2006, p. ix), writing in the foundational text *Common Reading Programs Going Beyond the Book*, states that “[a] simple definition is that common reading

programs are educationally purposeful programs that engage students in a variety of in- and out-of-class academic and social experiences.” As she points out, campus common read programs can vary greatly in “type, expectations, scope, and size” (p. 11). She expands, explaining that some campus programs will require students who are just entering the college or university to read the book that has been selected during the summer immediately prior to their arrival on campus. They will then be expected to participate in discussions on the book with both teachers and fellow students once they have arrived on the campus. Other campuses will have students read the selected book during their period of orientation to the campus. Other programs, rather than selecting a single book, may provide students “with a reader consisting of multiple reading sections organized around a theme” (Laufgraben, 2006, p. 11).

While most campus common read programs focus solely on campus participants, other programs may invite members of the surrounding community to read a book in common with the members of the university community, while other campus read programs “have grown to include library exhibits, film series, theatrical performances, and grant-funded faculty development experiences” (Laufgraben, 2006, p. 11). “The ‘common’ component,” Laufgraben emphasizes, “involves a targeted audience asked to read the same text(s) in a defined period of time so that participants can engage in an academic discourse about those texts” (p. 11).

## Goals and Purposes of Common Read Programs

The goals and purposes of the common read programs that have proliferated across American colleges and universities in the quarter of a century since Nancy Pearl began the movement tend to fall into five categories, which Thorne (2015) has classified by examining key words found in the rationales that schools provide for their programs. The categories that she has identified include “(1) building community on campus, (2) setting academic expectations, (3) starting a conversation, (4), encouraging students to become involved in social activism, and (5) inspiring thoughtfulness or ‘critical thinking’” (136).

Other perspectives argue that common read programs can be used as an effective way to encourage students to explore their personal values and ethics; to help campus constituents become more aware of the cultural diversity that exists around them, both on the campus itself as well as in the communities that surround those campuses; to help students feel that they are valued members of that wider community; and to help integrate social and academic experiences into university life (Boff et al., 2007; Nadelson & Nadelson, 2012).

Much of the literature on common read programs also claims that reading books not directly associated with homework can serve to encourage students

to enjoy the reading process more, as well as to increase the overall amount of reading that students do, including reading that is done for pleasure as well as learning (Paulson, 2006; Twiton, 2007; Mallard et al., 2008; Nadelson & Nadelson, 2012). An additional study found that students who read the entire common read text were more likely to report a stronger connection to their campus community and even had higher academic achievement in college (Daugherty & Hayes, 2012).

Campuses have also begun to recognize that common read programs can serve as an effective way to at least partially fill the gap that has been created by the loss of the traditional core curricula that characterized colleges and universities for so many generations and that having students in the same year read at least one book in common is an effective method to ensure that students have a shared intellectual experience with their peers (Thorne, 2015).

### **Expanded Partnerships across the Campus**

While most campus common read programs have traditionally been the province of first-year experience programs, other entities from across the campus have also become deeply involved in campus common read programs over the last two decades.

One group that has been particularly active in the success of campus common read programs are first-year writing programs, which are widely seen as a convenient and logical place to integrate common reads since virtually all first-year college students must take at least one of these courses (Van Vaerenwyck, Clark, & Pasinella, 2002). And while the goals and purposes of common read programs can vary widely, as detailed above, nearly all overlap with the goals of first-year writing programs and the work carried out by those who administer writing programs (Benz et al., 2013).

The approaches that first-year writing programs use to achieve their desired outcomes can vary widely as well. Many writing programs take a traditional focus on the book, incorporating the common read text into their first-year writing courses and addressing the reading through a set of inquiry-based assignments (Van Vaerenwyck et al., 2002). At the other end of the spectrum, some programs, such as that described by Wranovix and Isbell (2020, p. 2) take a more modern digital approach by turning “to the concept of social annotation in a digital environment to reimagine the common read as an interactive experience, one in which students read a text together and can see in real time the thoughts and interpretations of their peers.”

Another entity that has become integral to campus common read program success is the academic library. Librarians are considered to have, as Angell (2019, p. 1) puts it, the ability to “collaborate with a wide variety of campus

departments to better incorporate a common read into the fabric of university life” and by playing “significant roles in selecting, promoting, and further supporting the common read.” Academic librarians are also seen as being able to contribute to the success of common read programs in ways that first-year experience or first-year writing programs cannot, such as through the creation of faculty teaching resources directly related to the book, which can include highly customized research guides, and by creating such extracurricular additions as book displays that help “to bring the book to life” (Chestnut, 2011, p. 55).

The inclusion of the academic library into the mix has been found to produce numerous other benefits as well. Within an academic environment, for example, libraries can bolster partnerships among students, librarians, faculty, and administrators. In addition, Rodney (2004, p. 1) claims, “students are often unaware of the many research resources that university libraries can provide and if faculty, such as first-year writing program faculty, don’t explicitly include library/research activities in their course curriculum, academic libraries can easily become storehouses rather than centers of learning.” By participating in campus common read programs, students are introduced to novel research topics and experience diverse opportunities to learn and practice important critical thinking, oral communication, and written communication skills, “while at the same time providing exciting and intellectually stimulating experiences that will encourage reading and open discussion and debate beyond the classroom” (p. 1).

Other entities have taken the campus common read program in a number of different and more specific directions, including using the campus read program as a focal point when creating a health sciences specific learning community and using the book to expand and improve the experiences of students participating in that learning community (Virtue et al., 2018); using the common read to create an off-shoot program whose focus is to explicitly foster empathy in college students (Graves & Reinke, 2021); using the read as a way to foster greater collaboration between student affairs and academic affairs (Delmas & Harrell, 2014); or focusing exclusively on memoirs set in New York City and written by authors who are available to visit campus (Moser, 2010).

And while most campus common read programs involve students only during the first academic term of the school year, Maloy et al. (2017) describe a year-long common read model that uses a first term book club experience to promote faculty development which allows for cross-disciplinary interactions, including the development of both events and assignments, which can then be integrated into courses during the following term by the instructors who were involved in the program. They argue that the structure of their program promotes college reading at their institution as it “builds an intellectual community of students and faculty across the campus. It posits college reading as a sustained collaborative, intellectual enterprise in which students and faculty

critically consider the context and implications of a text across disciplines” (p. 68).

As can readily be seen from the discussion above, there is a great deal of literature on the experiences of students participating in common read programs within the North American context, particularly those participating through colleges and universities in the United States. On the other hand, research that addresses the experiences of common read participants beyond those borders is greatly lacking. The present paper strives to partially remedy that situation through its examination of a transnational common read program, especially within the Polish context, thereby expanding both the research and the understanding of this interesting and important topic.

## **Background and Institutional Context**

The International Book Club (IBC) has its origins in the campus common read program of California State University, Bakersfield (CSUB), where it is widely known throughout the community as the One Book Project. CSUB is a comprehensive, four-year, regional university located roughly 180 kilometers to the north of Los Angeles, California. As a member of the 23-campus California State University system, CSUB serves over 11,000 students per school year, the majority of whom are Hispanic and the first in their family to go to college.

Begun in fall 2007, the One Book Project was originally created to support the university’s First-Year Experience (FYE) program, whose courses are designed to help students make the transition from high school to college. From its inception, the One Book Project has been an “in-semester” common read program where students read and study the selected book in first-year courses throughout the first academic term of the school year. At CSUB, the selected text is a required component of both First-Year Experience and Composition classes—both of which are required of all first-time freshmen. Over the last 15 years, it has also grown to be included in courses throughout the University’s General Education program.

The goals of the One Book Project, derived primarily from Laufgraben (2006), include:

- providing a common academic experience for all first year and transfer students;
- strengthening the academic atmosphere of the university;
- connecting students to their peers and instructors;
- getting students involved in campus activities through a shared learning experience;
- modeling academic behaviors;

- setting expectations for student success;
- fostering campus and community involvement;
- promoting more meaningful learning.

The fall 2021 semester marked the 15th year that CSUB sponsored the One Book Project. During the Fall 2018 and Fall 2019 editions, the selected authors were invited to campus to attend an afternoon reception, which included having the author help to present writing and art awards to both college and high school students, and were then asked to give a 30- to 40-minute keynote speech, followed by a half-hour question-and-answer session and a book signing. Unfortunately, due to the outbreak of the COVID pandemic, the 2020 and 2021 author keynote speeches and question-and-answer sessions were held virtually over Zoom, eliminating the possibility of book signings.

### **The International Book Club**

Begun during the fall 2018 semester and designed to be a transnational extension of the One Book Project, the IBC is a common read partnership program that initially brought students from the University of Applied Sciences in Nysa, Poland, together with students enrolled in CSUB's First-Year Experience and First-Year Composition courses, as well as with students studying at CSUB's intensive English program, the American Language Institute. During the four years that the IBC has been in existence, the program has evolved to also include participation, at different times and to different degrees, from students in Ensenada, Mexico, Coban, Guatemala, Buenos Aires, Argentina, Hradec Králové, Czech Republic, and Bratislava, Slovakia.

Over the last four years, the nature of the program has also evolved substantially, especially in response to the changes within both the Polish and American educational systems that have been necessitated by the COVID pandemic. Beginning initially as an asynchronous project where students would read the book and discuss a predetermined set of questions in their own individual classes, then interact with their international partners only through a fairly simple online blog, the IBC has now grown into a more sophisticated program with a more advanced website in which students can add vlogs, including those that introduce themselves to their international partners, and through which they can receive additional information and resources beyond what they get in their individual classes. More significantly, the program now also includes synchronous, virtual meetings between program participants in which they can directly interact with, and learn from, their international student counterparts. Because the IBC involves countries on three different continents, as well as universities with quite different course schedules and timetables, the creation of a workable plan for synchronous meetings has been challenging at best.

Based on evidence that this study has gathered thus far, however, the effort appears to be well worth it.

## The Study

The data examined in this article is derived from an ongoing, longitudinal study that examines the experiences and impressions of Polish college students who have participated in the IBC. Thus far, there have been four editions of the project, with data collected after each run. The analysis presented in this article analyzes data gathered between October 2018 and December 2021. In none of the project surveys was a division made between female and male members of the IBC. To date, the study has been conducted in Nysa, a small city in southwestern Poland, with a population of over 40 thousand residents. Over the course of the study, most of the college students who participated in the common read program lived in Nysa or in smaller towns and villages in the nearby region.

## Participants

In the first stage of the study, from October 2018 to January 2019, second-year students taking courses through the Department of Modern Languages at the University of Applied Sciences in Nysa, Poland, discussed *Incarceration Nations* (2016) by Baz Dreisinger. A total of 11 second-year college students were surveyed in the fall 2018 project's run as shown in Table 1 (see Appendix). In the second year of the study, from October 2019 to January 2020, project participants read *The Boy Who Harnessed the Wind* (2009) by William Kamkwamba and Bryan Mealer. During this program, the groups of readers got larger, including freshmen and sophomores this time and growing to 36 first-year and 23 second-year college participants as presented in Tables 2 and 3 (see Appendix). As part of the third and fourth editions of the study, second-year students at the University of Applied Sciences in Nysa discussed *Broad Band: The Untold Story of the Women Who Made the Internet* (2018) by Claire L. Evans during October 2020, and then *A Dream Called Home* (2018) by Reyna Grande in October 2021. Thirty-one second-year college students participated in the 2020 common read program (Table 4, see Appendix), while 21 college learners, who were enrolled in second-year courses, took part in the 2021 IBC project (Table 5, see Appendix).

## Instruments

For this study, both quantitative data and qualitative data were gathered in an attempt to give a full picture of the experiences and impressions of participants in the IBC during the four-year period of study. Quantitative data was gathered through a nine-statement survey administered to all participants after each of the four editions of the IBC. This survey attempted to gather data on the experiences of reading each year's selection, examining participants reading experiences both inside and outside of class, as well as ascertaining whether participants' viewpoints had been influenced by reading the book. Specifically, the questionnaire asked participants to respond with a simple "yes" or "no" to the following statements: (1) I enjoyed reading the book, (2) I would recommend the book to others, (3) I talked to my friends about the book outside of class, (4) I talked to my family about the book outside of class, (5) Reading the book encouraged me to read more, (6) Reading the book has made me want to read more about the subject, (7) Reading the book caused me to explore ideas and issues that I might not have thought about otherwise, (8) Reading the book caused me to have a change in attitude toward the subject, and (9) Reading the book has helped me understand English better.

Qualitative data was then gathered through a follow-up electronic questionnaire which attempted to delve deeper into the experiences of participants in the program through more open-ended questions. These questions included the following: (1) Do you think that the IBC project, in cooperation with California State University, Bakersfield, should be continued? Why? Why not?, (2) Have you read all the chapters of the book you were assigned to? If not, why not?, (3) After reading a chapter, did you prepare answers to the questions assigned as homework? If yes, how? If not, why was this the case?, (4) What have you gained by taking part in the project?, (5) Has [the book] changed your attitude to any of the problems touched upon in the book? Why or why not?, (6) In what ways has the book changed your attitude to the above-mentioned problems? If the book has not changed your attitude to any of the issues described in the book, why do you think this is the case?, (7) Do you think that public discussions on topics mentioned in the book are important? Why or why not?, (8) How important to you was the online discussion of the book?, (9) What have you gained by taking part in the online discussion?

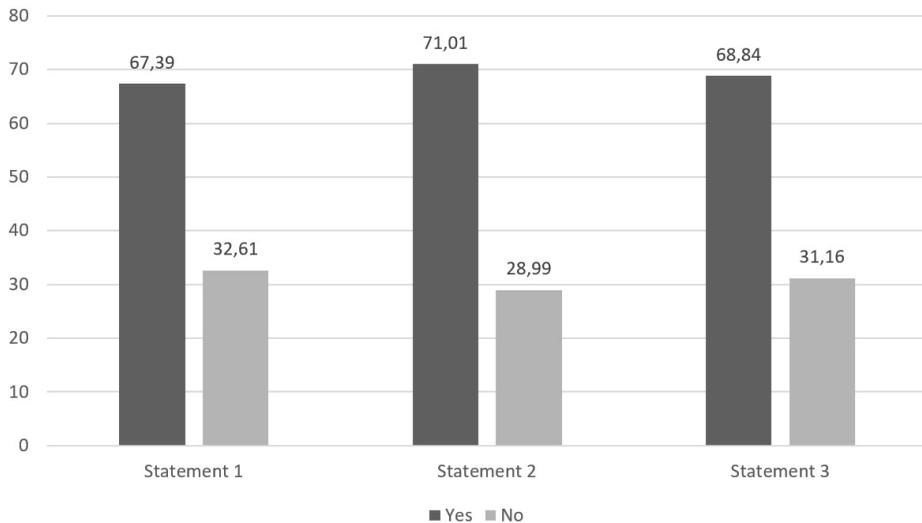
## Survey Results

Overall, the data gathered through the four editions of the survey administered at the end of each edition has proven to be quite promising, with some notable exceptions. (Please see the Appendix for a complete breakdown of the

responses for each of the separate editions and for each of the different sub-groups surveyed.)

**Figure 1**

*Collective Percentage Showing IBC Project Participants' Responses to Statements 1, 2, and 3 Gathered through the Four Editions of the Survey*



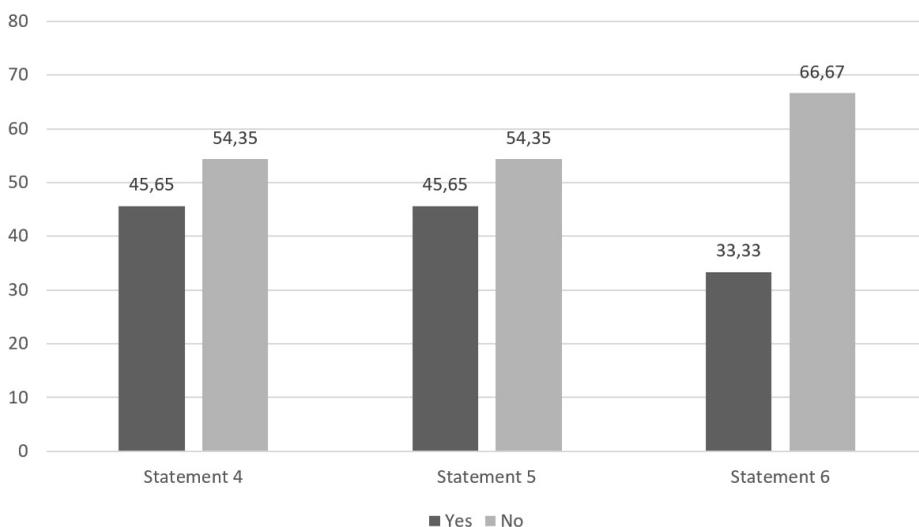
As can be seen in Statement 1 in Figure 1, over 67% of the respondents indicated that they enjoyed reading the book selection. Likewise, a very healthy 71% of the participants expressed that they would recommend the book to others (Statement 2). And as Statement 3 demonstrates, 68% of the students taking part in the IBC also asserted that they had talked to their friends about the book outside of class. That shows that the majority of the readers taking part in the following runs of the project appreciated the reading assignments, which makes the IBC a form of education approved by the students. Interestingly enough, the overall number of participants who were willing to recommend the book they had perused was even higher than the group of college students that found pleasure in turning the pages. The data also showcase that the selected readings made an impact on the majority of the students as they touched upon the content of the work when conversing with their friends outside the class as well, which with the high percentage proving enjoyment reinforces the educational value of the project and does not restrict it to a classroom.

Unfortunately, not all of the responses were quite so positive. For example, as Statement 4 in Figure 2 shows, only 45% of the respondents indicated that they had discussed the book with family members outside of class, with almost

55% expressing that they had not done so. Likewise, only 45% of the participants responded that reading the book had encouraged them to read more, while almost 55% declaring that it did not (Statement 5). On the most negative side, as can be seen in the data from Statement 6, only one-third of the students participating in the IBC replied that reading the book had made them want to read more about the issues discussed in the book, while a full two-thirds replied that it did not.

**Figure 2**

*Collective Percentage Showing IBC Project Participants' Responses to Statements 4, 5, and 6 Gathered through the Four Editions of the Survey*

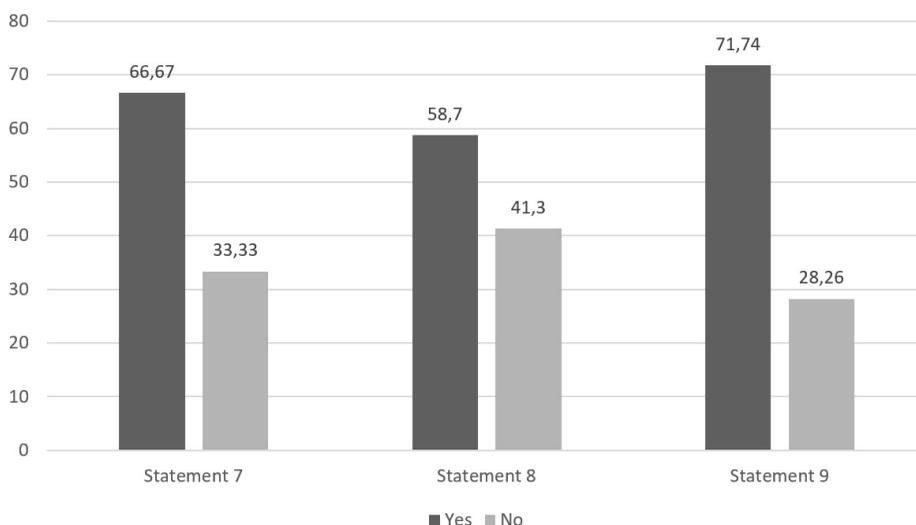


Analyzing the data, it is worthwhile to highlight that the students were more willing to share their impressions of the book with their friends (Statement 3) than with their family members (Statement 4). Nonetheless, the reasons explaining why less than half of the respondents discussed the book with their relatives go beyond the scope of this study as they pertain to the students' relationships with their families, an issue that was not examined in the survey. In addition, the low scores evidencing the participants' unwillingness to peruse are not stunning. According to the research conducted by the Polish National Library in March 2022, encompassing over 2000 respondents aged 15 and beyond, only 38% of the participants answered affirmatively to a question asking if they had read at least one book in a span of 12 months preceding the survey (Zasacka & Chymkowski). What is reassuring is that 45% of the Nysa students participating in the project responded that reading the book had encouraged them to read more, which surpassed the 2022 national readership average in Poland, proving the validity of the IBC project. Still, taking into account that only

one-third of the students replied that reading the book had made them want to read more about the issues discussed in the work clashes significantly with the numbers proving the students' general enjoyment (Statement 1) and thus raises further questions that should be encompassed in subsequent runs of the International Book Club.

**Figure 3**

*Collective Percentage Showing IBC Project Participants' Responses to Statements 7, 8, and 9 Gathered through the Four Editions of the Survey*



On the other hand, as the data in Figure 3 indicates, participating in the IBC can have a quite positive effect. As can be seen in Statement 7, a full two-thirds of participants felt that reading the selected text had caused them to explore ideas and issues that they might not have thought about otherwise. And while respondents may not have been motivated to read more about the subject, as indicated by Statement 6 above, members of the IBC did express that participation in the project had caused them to have a change of attitude toward the topic discussed in the book (Statement 8), which showcases a very satisfactory outcome of the project and evidences its validity. And finally, and perhaps most importantly, as Statement 9 denotes, almost 72% of the participants stated that participating in the club and reading the book helped them to understand the English language better, the claim proven through the questionnaire results interpreted in detail below.

## Questionnaire Results

In order to delve deeper into the experiences and impressions of participants, a follow-up online questionnaire was administered to all participants of the IBC at the end of each edition's run. A total of nine open-ended questions were included in the questionnaire and participants were encouraged to give as complete answers as possible. The following discussion highlights the most salient responses for each of the four editions, including both positive and negative responses.

### *The Fall 2018 Edition*

From October 2018 or January 2019, second-year students in the Department of Modern Languages at the University of Applied Sciences in Nysa participated in the first edition of the project by reading and discussing the book *Incarceration Nations* by Baz Dreisinger. At the end of the first edition, a questionnaire was administered to the 11 college participants regarding their experiences and impressions in reading *Incarceration Nations* and participating in the project.

All 11 respondents declared that the project should be continued due to several reasons. To begin with, some of the students felt that the project encouraged them to participate in discussion activities and having the opportunity to exchange opinions was important. Secondly, some of the participants enumerated the importance of the development of language skills, such as speaking, vocabulary, syntax, or reading comprehension. Finally, the students valued the fact that they discovered new issues and ideas which they were not familiar with.

When asked what the students had gained by taking part in the project, the responses were varied. To begin with, some respondents acknowledged that the project had broadened their viewpoints and enriched their knowledge. Moreover, there were students who admitted that they had read the first book in English in their lives. One student claimed that reading nowadays is not popular.

When asked to state whether public discussions on the issue were important, every student agreed that they were. One respondent explained that these issues have "a tremendous impact on the society" and thus people should focus on them. Another student stated: "it enables students to deliberate on the issues we usually don't think everyday." It was also claimed that nowadays "people read different books and that does not stimulate public discussion."

When asked how important for the students the online discussion based on the book was, every student stated it had been important. To justify their answers, the respondents indicated the opportunities to exchange the opinions with people from different parts of the world. One respondent summed it up nicely, stating: "The project enables the exchange of views with students from

a different part of the world, teaches us how to debate in a well-mannered way, and shapes our opinions on important issues that we usually neglect in everyday life.”<sup>1</sup>

### ***The Fall 2019 Edition***

In February 2020, at the finish of the second edition of the project, 22 of the 36 first-year students who had completed the initial survey at the University of Applied Sciences in Nysa were administered the questionnaire to learn about their experience in reading and discussing *The Boy Who Harnessed the Wind* by William Kamkwamba and Bryan Mealer and their experiences in participating in the IBC project.

When asked to comment on whether the International Book Club should be continued, all but one student answered affirmatively. Positive comments about the project included such statements as “the project integrates students through conversations,” “is an excellent way to exchange opinions,” and “is a nice opportunity to exchange insights with people of other nationalities about the book.” One participant claimed that the “[p]roject made me read and finished first book in my life.” Another student added: “It is an interesting and fresh experience. It makes us break away and learn something new at the same time. Such reading not only enriches our English dictionary, but also the very knowledge of the world. In the case of such interesting books as *The Boy Who Harnessed the Wind*, which was not too boring, it was also suspenseful and touching. It was not such a typical standard *reading* book. It brought something to the head and heart.” On the negative side, the student who did not agree that the project should be continued stated that “the project does not introduce anything important into the studies’ program.”

When asked what the students had gained by taking part in the project, the answers were varied. Some students stressed the improvement of linguistic skills “both through reading and later discussions,” and “in speech and writing.” What is more, some of the students claimed that the experience had changed their ways of thinking with one claiming: “I am less prejudiced than I was before,” and another stating that “William’s story has taught me that it’s not worth giving up.” Several of the answers focused on raising awareness of the starvation issue discussed in the book. Finally, three respondents asserted that they were encouraged to read more books of different genres they used to read. Once again, on the negative side, the student who did not agree that the project should be continued claimed that they gained “completely nothing.”

---

<sup>1</sup> The responses given by the students participating in the project are quoted in the entire article as they were given, with no additional correction.

When asked in what way the book has changed the students' attitude to the issues mentioned in the book the answers were again varied. One student summed it up: "The book, written in the first person, gave the opportunity to empathize with a person who lives in that world and struggles with problems that we do not even assume that we could have, such as hunger, extreme poverty, lack of access to medicine and education."

When asked to state whether public discussion on the issue was important, the project participants agreed that it was. One student claimed that "only by talking about problems in public we are able to make a difference." Another student added that "[t]he topics covered in the book are quite heavy, so they should be discussed and analyzed." When asked what they had gained by taking part in online discussions, the participants stressed the importance of exchanging opinions. As one student stated, "the discussion was important because it helped me to look at the topic from different perspectives." Furthermore, two students stated that it was an enjoyable experience. Another respondent pointed out: "The book changed my attitude to education."

### ***The Fall 2020 Edition***

At the end of the third run of the project, the open-ended questionnaire was once again administered to 31 second-year students at the University of Applied Sciences in Nysa regarding their experiences in reading and studying *Broad Band: The Untold Story of the Women Who Made the Internet* by Claire L. Evans.

In response to whether the International Book Club should be continued, all students once again agreed that it should. The students highlighted the importance of international cooperation, discussion opportunities including the exchange of ideas, the development of language skills, and the opportunity to discover another culture. One student called the project "illuminative." Another student, in explaining their opinion, stated that "it forces students to read books that they might not have read themselves. Books are also about very valuable issues, broadening horizons." A different participant added: "It seems to me that the books we read are not very popular among people, but they are very interesting and show problems and people who are heroes because they changed the world and who we had no idea about."

When asked what they had gained by taking part in the project, the answers were numerous and varied. The students indicated that the project had developed both their language skills and knowledge of the discussed topic. One participant explained: "I believe that by participating in the project, I gained new knowledge both in terms of vocabulary and knowledge about women important for the scientific world and the whole world, and I also learned about many new games, computer programs, portals that I had not previously heard

about.” Another student asserted: “I have gained knowledge about the history that was not told to me in the history class when the topic was about the internet and the development of it. Moreover, I have gained better use of the English language and improved my pronunciation while expressing my thoughts, and myself so it helped me think critically about something I had read about and gave me the confidence to speak about it.”

When asked in what ways reading the book had changed their attitudes about the topic, the interviewed students underlined the importance of raising awareness of the topic. One student expressed: “The book shows how important women are, how much they have achieved, what they have created, so if they succeeded, why should I fail?” Another participant stated: “It completely changed my view of gender, I always thought it was a man’s job. That’s why I respect the woman who paved the way for other women in the computer industry.” A different student added: “I started to believe more in my abilities.”

When asked to state whether public discussions on the issue were important, every interviewed student agreed. One participant explained: “I believe they are very important. It’s the 21st century and women are still treated worse than men. They are insulted, harassed, do not have full civil rights or even human rights, and in the end they are treated as stupider, as those who cannot achieve more than a man. When a woman discovers something the man hasn’t done, she is relegated to the shadows. It is only a woman who is told what they can or cannot, what is right for them and what is not, a man is not forbidden to do anything, any of their behavior is justified. Therefore, in every possible way we should talk about inequality and fight against it.” A different student claimed that “[c]hanging perspectives and gender roles is important. Often these very problems are very harmful and prevent a person from developing. Through these stories it is possible to show that women are also gifted and full of ideas, even in areas that are attributed to men.” Finally, one student asserted, “Public discussions are important, because they concern me directly as a woman.”

When asked to enumerate what they had gained by taking part in the online discussion, the students stressed the opportunity of exchanging opinions and spotting different points of views. As one student explained, “The very idea of the project is very good and it certainly allows you to understand a different culture as well, because each nationality may have a different perspective on a given problem.”

### ***The Fall 2021 Edition***

The online questionnaire was routinely again administered at the end of the fourth run of the project to students from the University of Applied Sciences in Nysa. For this edition, 21 second-year students were queried about their experiences in reading and discussing *A Dream Called Home* by Reyna Grande.

When asked to comment on whether the International Book Club should be continued, 19 of the 21 students answered affirmatively. One student supported their response by stating “It’s a good practice. It extends the capacity of vocabulary and skills of critical thinking.” Another student underlined the significance of international cooperation, asserting that “It is a nice opportunity to exchange insights with people of other nationalities about the book, and we do not have such opportunities every day.” Another participant added, “I think that these projects are important because thanks to them we have the opportunity to learn the language in a way that is much more practical than normal reading books.” A different student explained: “This project should be continued due to the fact that its participants conduct intensive discussions on important topics also outside the classroom.” Finally, one student admitted that the project had caused them to read and finish the first book in their life. On the other hand, one student who was not in favor of continuing the project acknowledged, “It was a nice project with good intentions. However, I think next time the students should be able to pick the book for reading to make the club more interesting.” Another student added that different types of books ought to be selected, ones “that focused on more fictional and general topics so that they would appeal to everyone.”

When asked what the participants gained by taking part in the project, the vast majority of them enumerated its advantages. The arguments were the following: motivation to complete further work, redefinition of someone’s priorities, gaining courage to embrace difficult topics, an enrichment of vocabulary, and knowledge about the world. Moreover, the students underlined the significance of the development of their speaking skills that prepared them to participate in the class discussions. One student stated: “I believe that thanks to the project, I overcame my fear of English-language literature and I will not limit myself only to literature in Polish any longer, but I will start reading books in English too.” The only opposite opinion was the following: “I’m sorry to say that I gained completely nothing. Only good thing was that I was reading something which I’ve been trying to do for a long time. Unfortunately, the book choice made me less encouraged to read books.”

When asked in what ways reading a book changes students’ attitude about the topic, students in general acknowledged that they understood the issue more deeply. Furthermore, one student stated: “The book changed my attitude to education.” Another student explained: “Better understanding of the problems reflected in the book and most of all a real joy to have open and friendly discussion with other people (teachers/students).” On the other hand, there were some statements that the students gained nothing. For instance, one participant stated, “The book did not change my approach in any way. I don’t care about problems that don’t concern me.”

When asked in what way the book had changed the students' attitude to the problems mentioned in the book, the participants enumerated plenty of varied examples. Some students underlined the importance of appreciating their life, comparing it to the hardship presented in the book. One student claimed, "it made me realize there are other people which are in the way more worse situation than me." Another student stated that they began to appreciate their mother more. On the contrary two students pointed out that the issues presented in the book were biased: "[T]he opinions loosely based on her (the author's) prejudice towards everyone." The second student added: "As someone who experienced discrimination in Germany because of my Polish origins I felt really mad at author at some point because being white doesn't give any privileges and book made it look like it does."

When asked to state whether public discussions on the issue were important, every student agreed that they were. In order to justify their statement, one student explained: "It was very useful, and reading answers to questions from people from other countries, from other universities was both useful and enjoyable." Another student added: "It allowed me to get acquainted not only with the comments of students I see every day, but also of people from other parts of the world, which was very developing and broadened my point of view." To sum up, one student claimed that by talking about particular issues in public, people can make a difference.

When asked how important for the students the online discussion based on the book was, 17 out of 22 students stated that it was important. The students explained that the discussion was "interesting," "the topic was important," and "it [was] a good opportunity to train speaking skills." Finally, when asked to enumerate what the students had gained by taking part in the online discussion, the answers were varied. One student claimed: "The online discussion was an important part of the project for me, as I was able to see how many people were getting involved and sharing their opinions, as well as engaging in a conversation on a given topic." Another one added: "I really like the online discussion because it's a good opportunity to train listening skills while talking to students from abroad."

## Discussion

Overall, based on the data gathered thus far, the International Book Club has proven to be a success, although some limitations remain. On the positive side, for the most part, participants in the program have enjoyed the experience of reading the book, with at least one student remarking that it caused them to

read a complete book for the first time in their life. More importantly, being a part of the International Book Club has caused participants to have more frequent, and perhaps more sustained, discussions with their classmates than they might otherwise have had. It has also caused them to think more deeply about the issues discussed in each year's selection than they normally would. The experience has also clearly caused many participants to have a change of attitude toward the subject through their reading and discussion of the book. Likewise, the experience of participating in the IBC has exposed them to diverse ways of thinking and learning, helping them to see the importance of the exchange of ideas and opinions. And clearly, based on answers from both the survey and the questionnaire, participation in the project has helped to improve the English language abilities of the members of the club. And since participation in the club exposed participants to books that they would not normally have access to, participation has also helped them become more familiar with English language literature.

Interestingly, although the International Book Club is indeed a "book" club, this improvement has not been limited to only an enhancement in reading comprehension ability, with various participants remarking that it has also helped them improve their speaking and listening in English, allowing them to engage more freely in conversations and discussions. Even more, as many participants remarked, it has helped them improve their grammatical competency and expand their vocabularies.

On the other hand, the limitations of the project must be recognized. While participation in the project clearly caused participants to have expanded and more in-depth discussions with fellow classmates, this effect did not seem to extend beyond the academic context, since respondents reported not having discussions outside of the classroom setting. Likewise, participation in the program does not seem to have caused a significant change in the reading habits of club members, since participants reported that they did not read more because of participation in the club. Both of these issues must be addressed in future editions of the project.

## Conclusion

The aim of the present study was to detail the findings from the first four years of an ongoing, longitudinal study that delved into the experiences and perceptions of Polish college students as they participated in the common read experience called the International Book Club, which brought them together virtually with students from the United States, Mexico, Guatemala, Argentina,

the Czech Republic, and Slovakia. These early results show that participants in the project did benefit from the IBC program and that it had a positive effect on their reading experiences, helping them to delve deeper into, and have deeper discussions on, topics and issues that they normally would not be exposed to. This experience also clearly caused participants to experience changes in their attitudes toward these topics, exposing them to diverse points of view and belief systems. And significantly, it also clearly caused participants to experience an improvement in their English language abilities.

In addition, this research, through its analysis of the Polish context, is meant to expand the contexts in which common read programs are researched. Looking forward, future research needs to be extended into ever-increasing contexts, countries, and cultures, if we hope to gain a fuller understanding of the potential of common read programs. For our part, we have been convinced enough by the feedback that we have so far received from participants in our project to continue the International Book Club into the future, with participants in the Fall 2023 edition reading *Sitting Pretty: The View From My Ordinary Resilient Disabled Body* by Rebekah Taussig.

## References

- Angell, K. (2019). The academic library and the common read: A multitude of possibilities for collaboration with campus programs and departments. *The Journal of Academic Librarianship*, 45(6), 1–3.
- Benz, B., Comer, D., Juergensmeyer, E., & Lowry, M. (2013). WPAs, writing programs, and the common reading experience. *WPA. Writing Program Administration*, 37(1), 11–32.
- Boff, C., Schroeder, R., Letson, C., & Gambill, J. (2005). Building uncommon community with a common book: The role of librarians as collaborators and contributors to campus reading programs. *Research Strategies*, 20(4), 271–283.
- Chesnut, M. T. (2011). Reading between the lines: Extending the reach of a university common reading program via the academic library. *Practical Academic Librarianship: The International Journal of the SLA Academic Division*, 1(1), 51–58.
- Daugherty, T. K., & Hayes, M. W. (2012). Social and Academic Correlates of Reading a Common Book. *The Learning Assistance Review (TLAR)*, 17(2), 33–41.
- Graves, N. A., & Reinke, J. S. (2021). Common read for the common good: A course project to foster college students' empathy. *Journal of Family & Consumer Sciences*, 113(4), 69–75.
- Laufgraben, J. L. (2006). *Common reading programs going beyond the book. The first-year experience monograph series no. 44*. National Resource Center for The First-Year Experience and Students in Transition.
- Mallard, J., Lowery-Hart, R., Andersen, A., Cuevas, M., & Campbell, K. (2008). We'll give you the story; you make it your own: Developing the common reader program at West Texas A&M University. *Enrollment Management Journal*, 2(2), 86–103.

- Maloy, J., Counihan, B., Dupre, J., Madera, S., & Beckford, I. (2017). The un-common read: Perspectives from faculty and administration at a diverse urban community college. *What Is College Reading*, 67–88.
- Moser, J. (2010). The Uncommon in Common Reading Programs: The Freshman Reading Program at Brooklyn College. *Currents in Teaching and Learning*, 2(2), 89–97.
- Nadelson, S. G., & Nadelson, L. S. (2012). In search of the right book: Considerations in common read book selection. *Journal of College Reading and Learning*, 43(1), 59–66.
- Paulson, E. J. (2006). Self-selected reading for enjoyment as a college developmental reading approach. *Journal of College Reading and Learning*, 36(2), 51–58.
- Randall, D. (2019). Beach books: 2018–2019. What do colleges and universities want students to read outside class?. *National Association of Scholars*. Retrieved from <https://www.nas.org/reports/beach-books-2018-2019/full-report>.
- Rodney, M. L. (2004). Partnerships and connections: Building community partnerships: The “One Book, One Community” experience. *College & Research Libraries News*, 65(3), 130–155.
- Rogers, M. (2002). Libraries offer chapter and verse on citywide book clubs: In Seattle, Chicago, Syracuse, and even some states the concepts and tactics vary, but collective reading grows. *Library Journal*, 127(6), 16–18.
- Spiegel, P. (2005). “One Book” programs span the nation. *American Libraries*, 36(5), 19.
- Thorne, A. (2015). Common reading programs: trends, traps, tips. *Academic Questions*, 28(2), 135–146.
- Twiton, A. (2007). Common reading programs in higher education. Folke Bernadottee Memorial Library, Gustavus Adolphus Library, St. Peter, MN.
- Van Vaerenwyck, L. M., Clark, S., & Pasinella, A. (2022). Inquiry-based learning in the context of a common read program encourages desired learning outcomes in first-year writing courses. *Journal of College Reading and Learning*, 52(1), 23–41.
- Virtue, E., Wells, G., MacKusick, C., Murphy-Nugen, A., Rose, A., & Snyder, M. M. (2018). The immortal life of Henrietta Lacks: Using a common read to transform a learning community. *Learning Communities: Research & Practice*, 6(1), 1–17. Retrieved from <https://washingtoncenter.evergreen.edu/lcrpjournals/vol6/iss1/4>.
- Wranovix, M., & Isbell, M. (2020). The digital common read: Creating a space for authentic engagement with social annotation. *Journal of the European Honors Council*, 4(1), 1–10.
- Zasacka, Z., & Chymkowski, R. (2022). *Stan czytelnictwa książek w Polsce*. Warszawa: Biblioteka Narodowa.

## Appendix

**Table 1***The Fall 2018 Edition: Second-year College Students' Responses to the Project*

Statements	YES	NO
I enjoyed reading <i>Incarceration Nations</i> .	6	5
I would recommend <i>Incarceration Nations</i> to others.	9	2
I talked to my friends about <i>Incarceration Nations</i> outside of class.	8	3
I talked to my family about <i>Incarceration Nations</i> outside of class.	4	7
Reading <i>Incarceration Nations</i> encouraged me to read more.	5	6
Reading <i>Incarceration Nations</i> has made me want to read more about this subject.	4	7
Reading <i>Incarceration Nations</i> caused me to explore ideas and issues that I might not have thought about otherwise.	10	1
Reading <i>Incarceration Nations</i> caused me to have a change in attitude toward the incarcerated and formerly incarcerated.	10	1
Reading <i>Incarceration Nations</i> has helped me understand English better.	6	5

**Table 2***The Fall 2019 Edition: First-year Students' Responses to the Project*

Statements	YES	NO
I enjoyed reading <i>The Boy Who Harnessed the Wind</i> .	24	12
I would recommend <i>The Boy Who Harnessed the Wind</i> to others.	28	8
I talked to my friends about <i>The Boy Who Harnessed the Wind</i> outside of class.	26	10
I talked to my family about <i>The Boy Who Harnessed the Wind</i> outside of class.	16	20
Reading <i>The Boy Who Harnessed the Wind</i> encouraged me to read more.	21	15
Reading <i>The Boy Who Harnessed the Wind</i> has made me want to read more about this subject.	14	22
Reading <i>The Boy Who Harnessed the Wind</i> caused me to explore ideas and issues that I might not have thought about otherwise.	26	10
Reading <i>The Boy Who Harnessed the Wind</i> caused me to have a change in attitude toward education.	24	12
Reading <i>The Boy Who Harnessed the Wind</i> has helped me understand English better.	31	5

**Table 3***The Fall 2019 Edition: Second-year Students' Responses to the Project*

Statements	YES	NO
I enjoyed reading <i>The Boy Who Harnessed the Wind</i> .	17	6
I would recommend <i>The Boy Who Harnessed the Wind</i> to others.	17	6
I talked to my friends about <i>The Boy Who Harnessed the Wind</i> outside of class.	15	8
I talked to my family about <i>The Boy Who Harnessed the Wind</i> outside of class.	12	11
Reading <i>The Boy Who Harnessed the Wind</i> encouraged me to read more.	9	14
Reading <i>The Boy Who Harnessed the Wind</i> has made me want to read more about this subject.	5	18
Reading <i>The Boy Who Harnessed the Wind</i> caused me to explore ideas and issues that I might not have thought about otherwise.	14	9
Reading <i>The Boy Who Harnessed the Wind</i> caused me to have a change in attitude toward education.	14	9
Reading <i>The Boy Who Harnessed the Wind</i> has helped me understand English better.	16	7

**Table 4***The Fall 2020 Edition: Second-year Students' Responses to the Project*

Statements	YES	NO
I enjoyed reading <i>Broad Band</i> .	20	11
I would recommend <i>Broad Band</i> to others.	24	7
I talked to my friends about <i>Broad Band</i> outside of class.	19	12
I talked to my family about <i>Broad Band</i> outside of class.	13	18
Reading <i>Broad Band</i> encouraged me to read more.	17	14
Reading <i>Broad Band</i> has made me want to read more about this subject.	14	17
Reading <i>Broad Band</i> caused me to explore ideas and issues that I might not have thought about otherwise.	25	6
Reading <i>Broad Band</i> caused me to have a change in attitude toward education.	16	15
Reading <i>Broad Band</i> has helped me understand English better.	23	8

**Table 5***The Fall 2021 Edition: Second-year Students' Responses to the Project*

Statements	YES	NO
I enjoyed reading <i>A Dream Called Home</i> .	11	10
I would recommend <i>A Dream Called Home</i> to others.	9	12
I talked to my friends about <i>A Dream Called Home</i> outside of class.	14	7
I talked to my family about <i>A Dream Called Home</i> outside of class.	7	14
Reading <i>A Dream Called Home</i> encouraged me to read more.	1	20
Reading <i>A Dream Called Home</i> has made me want to read more about this subject.	2	19
Reading <i>A Dream Called Home</i> caused me to explore ideas and issues that I might not have thought about otherwise.	7	14
Reading <i>A Dream Called Home</i> caused me to have a change in attitude toward education.	7	14
Reading <i>A Dream Called Home</i> has helped me understand English better.	11	10

Agnieszka Irena Kaczmarek, Emerson Case, Sebastian Zatylny

### **Lesen, Diskutieren und Teilen: Gründung einer lebhaften übernationalen Online-Lerngemeinschaft durch den internationalen Buchclub**

#### **Zusammenfassung**

Die vorliegende Arbeit berichtet über die Ergebnisse der ersten vier Jahre einer fortlaufenden Längsschnittstudie, in der Erfahrungen und Wahrnehmungen polnischer Teilnehmer an dem internationalen Buchclub – einer Erweiterung eines gemeinsamen Leseprogramms an einer amerikanischen Universität, unter die Lupe genommen wurden. Im Einzelnen wurde die Leseerfahrung der Teilnehmer untersucht, sowohl innerhalb als auch außerhalb des Klassenzimmers, ihre Meinung und Einstellung zu den im Rahmen der jährlich gewählten Bücher behandelten Themen und ob sich diese Einstellung durch ihre Erfahrung geändert hat oder nicht, bzw. ihre Einstellung zur öffentlichen Diskussion und ob sich diese Einstellung durch ihre Erfahrung geändert hat oder nicht, sowie die Frage, ob die Buchlektüre und die Teilnahme am Club zur Verbesserung ihrer Englischkenntnisse beigetragen haben oder nicht. Es wurden quantitative Daten in Form eines Fragebogens mit neun Stellungnahmen und qualitative Daten in Form eines Fragebogens mit neun offenen Fragen gesammelt und ausgewertet. Während ein großer Teil der Literatur über gemeinsame Leseprogramme die Erfahrungen der Teilnehmer im nordamerikanischen Kontext behandelt hat, gibt es immer noch einen Mangel an Forschung außerhalb dieses Kontextes, einschließlich des polnischen Kontextes. Die vorliegende Studie versucht, diese Lücke zu schließen. Die wichtigsten Ergebnisse der Studie zeigen, dass die Teilnehmer ihre Leseerfahrung insgesamt genossen haben und dass die Teilnahme an

dem Programm für sie von Vorteil war. Sie wurden unter anderem dazu gebracht, mehr zu lesen, sich mehr an Diskussionen zu beteiligen, tiefer über verschiedene Themen und Fragen nachzudenken und, was vielleicht am wichtigsten ist, ihre Englischkenntnisse zu verbessern.

*Schlüsselwörter:* gemeinsame Leseprogramme, internationaler Buchclub





**Katarzyna Budzińska**

Lodz University of Technology, Poland

 <https://orcid.org/0000-0003-1412-2140>

## Foreign Language Classroom Anxiety in Emergency Remote Instruction

### Abstract

Although language anxiety is the most widely studied emotion in second language acquisition, the extent to which emergency remote teaching (ERT) provokes anxiety is still insufficiently understood. The present research fills this gap by analyzing variables affecting anxiety in the new mode of instruction, while also exploring apprehension-relieving pedagogical approaches. This qualitative study was conducted among 218 university students and data were obtained by an online questionnaire and a focus group interview. Additionally, an online questionnaire was used to find out how teachers adapted their approaches to ensure student wellbeing in ERT. The findings reveal several factors affecting anxiety levels as well as pedagogical approaches that students find effective for anxiety reduction. On the whole, the study extends language anxiety research by looking closely at the variables affecting FLCA increase and decrease in a new, insufficiently explored mode of instruction.

*Keywords:* COVID-19, foreign language classroom anxiety, language learning emotions, emergency remote teaching, tertiary education, synchronous online instruction

Foreign Language Classroom Anxiety (FLCA), defined as “the worry and negative emotional reaction aroused when learning or using a second language” (MacIntyre, 1999, p. 27), was recognized by Horwitz et al. in 1986 and has been acknowledged ever since. Previously, anxious learners had not always been understood by instructors. Skipping class or seeking “refuge in the last row in an effort to avoid the humiliation or embarrassment of being called on to speak” (Horwitz et al., 1986, p. 130), apprehensive learners were frequently regarded as being unmotivated or as having a bad attitude. Thanks to the recognition of FLCA, instructors now understand that language learning can be accompanied by negative emotions, and that affected learners should be identified, supported, and treated with sensitivity (Horwitz et al., 1986). The

substitution of online learning for physical classroom learning, as a result of the COVID-19 pandemic, thus raises the following questions: To what extent do students experience anxiety in emergency remote teaching (ERT)? What are the FLCA sources in ERT? How do instructors reduce FLCA in ERT?

## Literature Review

The concept of FLCA first entered foreign and second language acquisition research with the seminal study by Horwitz et al. (1986). Extremely anxious language learners displayed psycho-physiological manifestations (Horwitz et al., 1986), while Young (1991) had also mentioned “avoiding activities in class, coming unprepared to class, acting indifferent, cutting class, putting off taking the foreign language until the last year, crouching in the last row, and avoiding having to speak in the foreign language in class” (p. 430).

Horwitz et al. (1986) identified three performance anxieties related to FLCA: communication apprehension, fear of negative evaluation, and test anxiety. They observed that learners who were challenged by speaking to an audience found it harder still when using the limited FL resources at their disposal, especially when they were being evaluated. As Horwitz (2017) clarified, while the three performance anxieties go some way to explain the nature of language anxiety, this specific anxiety represents more than the sum of its parts. Research indicates a variety of FLCA triggers, such as relationships between learners, the relationship between learners and their teacher, competition, and adverse comparison with peers (Horwitz, 2017). Anxiety can also be caused by the method of instruction or teaching style (Budzińska, 2015). While language anxiety can turn learning into an unpleasant experience, the crucial point about language anxiety is that it leads to suboptimal achievement (MacIntyre, 2017). Therefore, as Horwitz (2017) highlights, the goal of language anxiety research is “to make language learning more comfortable for vulnerable learners” (p. 44). Recent studies of emotions in SLA have emphasized the beneficial role of enjoyment in SLA (Botes et al., 2021; Dewaele et al., 2018; Jin & Zhang, 2021). As Dewaele et al. (2018) suggest, “teachers should strive to boost FLE rather than worry too much about students’ FLCA” (p. 678).

Although anxiety in foreign and second language acquisition has been researched extensively, relatively few studies have examined anxiety in the planned online environment (Coryell & Clark, 2009; Lee & Hsieh, 2019; Pichette, 2009). It was the recent outbreak of the COVID-19 pandemic and the global transfer to online education that urged a number of scholars (e.g., Alemany-Arrebola et al., 2020; Frascini & Tao, 2021; Kaiser & Chowdhury,

2020; Maican & Cocoradă, 2020; Resnik & Dewaele, 2021; Resnik et al., 2022) to look into learner emotions experienced during their new, suddenly introduced method of instruction, labeled as ERT (Hodges et al., 2020).

Maican and Cocoradă (2021), who investigated enjoyment and FLCA among Romanian FL university students during the pandemic, observed mixed feelings towards the new learning environment, which is consistent with Gregersen's (2020) comment that "language tasks are not unanimously enjoyed or universally anxiety provoking" (p. 67). Russell (2020) pointed out that online students could experience anxiety both as a result of learning a language and as a result of using new instructional technologies or platforms, which Kaiser and Chowdhury (2020) referred to as "technophobia" (p. 135). According to the recent Resnik et al.'s research (2022), exploring differences in online and face-to-face learners' FLCA before and during the pandemic, the use of technology was the most frequently reported apprehension source in ERT. The scholars also found that the main ERT anxiety sources differed from those inherent in face-to-face instruction. Other reported features of anxiety during ERT arose from participants' physical isolation from their peers (Sun & Zhang, 2021) and intrusion by other members of the household, including domestic pets (Adedoyin & Soykan, 2020; Maican & Cocoradă, 2021).

Instructors, who were already suffering from increased anxiety attributable to the pandemic crisis (Mercer, 2020), suddenly faced the challenge of teaching in a new medium without experience or adequate training (Adedoyin & Soykan, 2020; Bao, 2020; Gao & Zhang, 2020; MacIntyre et al., 2020). This may have led to lower-quality instruction (Derakhshan et al., 2021) and, consequently, may have aggravated students' anxiety about their progress or the effectiveness of their studies (Maican & Cocoradă, 2021; Resnik et al., 2022). Interestingly, according to Jelińska and Paradowski's findings (2021a), instructors who taught synchronously in the higher education sector coped well with adopting digital technology.

The hardships and uncertainties of the prevailing pandemic induced an increased level of general anxiety on teachers as well as students (Alemany-Arrebola et al., 2020). Additionally, emergency procedures introduced in order to facilitate remote learning, which increase students' workload, may also contribute to the anxiety and depression observed in the online learning environment (Commodari & La Rosa, 2020; Fawaz & Samaha, 2020; Jelińska & Paradowski, 2021b; Resnik et al., 2022).

Despite all these potential anxiety sources, some aspects of ERT may alleviate student apprehension. In Kaiser and Chowdhury's (2020) study, most respondents felt more secure in the virtual environment, as they were not exposed to the risk of losing face with their peers. Furthermore, not having to worry about appearance or classroom etiquette allowed others to feel physically relaxed. According to Resnik and Dewaele's (2021) findings, anonymity

and protection from public criticism in the online environment also contribute to lower anxiety levels. Moreover, the scholars demonstrated the prevalence of enjoyment over anxiety as well as a decrease in both anxiety and enjoyment in ERT classes. A drop in FLCA was linked, however, to students speaking the foreign language less in ERT classes than in face-to-face education. This decreased interaction was caused both by more teacher-centered instruction (Pawlak et al., 2022) and increased task evasion due to limited class management (Gao & Zhang, 2020).

In order to support learners struggling with the novel, imposed upon them (Russell, 2020) instruction mode, teachers had to adjust their instruction (Jelińska & Paradowski, 2021b). Nevertheless, pedagogical approaches employed in order to help anxious learners to cope with the suddenly introduced online instruction mode remain unexplored.

So far, research focusing on FLCA in ERT has been insufficient. Furthermore, as Resnik and Dewaele (2021) pointed out, ERT research would benefit from including interviews “to capture the complexity and nuances of learners’ unprecedented experiences” (p. 23). The present study uses narratives as well as focus group interviews to capture such nuances and thus extends language anxiety research. Furthermore, the present study investigates pedagogical approaches aiming to enhance student emotional comfort during ERT, which has not yet been considered by existing research.

## **Research Questions**

Based on the above-mentioned literature, the following research questions were posed:

1. To what extent do students experience anxiety in ERT?
2. What are the FLCA sources in ERT?
3. How do instructors reduce FLCA in ERT?

## **Material and Methods**

### ***Participants and Context***

The study was conducted at a technical university language center in Poland, where students are required to take an FL course in addition to their major

subject. The sample consisted of 218 students and 11 teachers. The vast majority of students ( $N = 212$ ) were aged between 20 and 23 years. The participants were mostly Polish ( $N = 206$ ), with five Ukrainians and two Belarusians. There were 148 males and 70 females. One hundred and sixty-nine of the respondents were learners of English, while 40 studied German, eight Italian, and one French. Advancement levels ranged from Pre-Intermediate to Advanced.

At the onset of the pandemic, face-to-face classes at the institution were replaced with synchronous online instruction on Microsoft Teams. This study was conducted in June 2020, three months after the introduction of ERT.

### *Instruments and Procedure*

Two anonymous Google Forms online questionnaires (one for students and one for instructors) and a focus group interview were used. All teachers working for the institution were contacted by the author by email and asked to participate in the research. They were informed about the study's purpose and assured anonymity. Links to the questionnaires were included in the email. Online consent was obtained individually at the beginning of each questionnaire.

The student questionnaire (Q) was administered during class by teachers who agreed to take part in the research. It consisted of eight questions (see Appendix A). The questions were in English and Polish, and participants could answer them in either language. The author translated the answers given in Polish into English before including them in the result section. The questionnaire was followed up with a focus-group meeting (FG) with volunteers organized as a video conference on Microsoft Teams. At the end of the questionnaire, there was an invitation to take part in a subsequent focus group meeting. Thirteen students accepted the invitation by providing their email addresses. They were then contacted by the author and a meeting was organized as a video conference on Microsoft Teams. The participants agreed that the conversation could be recorded and used for the purpose of this study. They were assured that the interviews would remain confidential and that their real names would be concealed. The meeting generated more details on the data gathered by the questionnaire. It lasted 76 minutes. Participants had the choice of the meeting to be conducted in Polish or English, and they preferred the latter. The questions posed are listed in Appendix B.

The teacher questionnaire consisted of two questions (listed in Appendix C) investigating whether instructors believed ERT induced greater anxiety than face-to-face education and whether they took any measures to reduce FLCA. The total data corpus is 41,774 words. Participants had the choice of the meeting to be conducted in Polish or in English and they preferred English.

## *Analysis*

Data analysis involved transcribing the interviews using Otter.ai, compiling them with the narratives from the questionnaire, manual coding and inductive data analysis. A grounded theory approach was applied (Charmaz, 2006). According to grounded theory, data collected should be analyzed for repeated themes, which are then tagged with codes. Codes are subsequently grouped into categories that can give rise to new theories. The transcripts were scrutinized by the author and an external researcher with a view of pinpointing anxiety sources and factors responsible for anxiety reduction self-reported by the respondents, as well as anxiety-reducing pedagogical approaches. At this stage, annotations were made to record any salient thoughts. The frequency of each reported anxiety source was calculated using Excel. Subsequently, categories were identified, and the data were analyzed again, assigning units of analysis to categories, that is, coding. Each data sample was analyzed in depth, subjecting it to multiple waves of coding and categorization, until the point of saturation was achieved. The data were analyzed by the two researchers separately. Any disagreements regarding interpretation and categorization were discussed and resolved. The inter-coder agreement reached 95%, which is considered acceptable (Miles & Huberman, 1994).

## **Results**

Findings are presented below, in conjunction with the research questions they relate to.

### ***Results for Research Question 1: To What Extent Do Students Experience Anxiety in ERT?***

The answer that ERT learning induced greater anxiety than face-to-face learning was mentioned as frequently as the answer that this type of language class was more relaxing. Many participants felt the same degree of anxiety in ERT as they did in face-to-face settings.

The most frequent explanation for feeling less anxious in ERT was anonymity. One student commented that physical signs of anxiety, such as blushing, are not visible when the camera is switched off:

*I feel less stressed during my online language class because others can't see me when I am blushing. (Q)*

Some respondents reported that it was less stressful to make a mistake in the company of other people who were heard but not seen:

*During online courses there's much less stress for the students because you are not directly responsible if you say something not right because during stationary courses if you say something not right, everybody is going to look at you or you may get bad reputation for, you know, saying something wrong and on online classes it's much less impactful to you because we don't see each other. (FG)*

*During my online language class, I am less anxious than in the classroom because there are just voices, and I don't feel afraid to make a mistake. (Q)*

The familiar, home environment provided comfort to some learners.

*Online classes make me a little less anxious because I am in the place where I feel safe. (Q)*

Additionally, some participants felt more relaxed at home as they could consult a dictionary discreetly with their camera switched off.

*I feel more comfortable sitting in my own room, where I know, if needed, I can, get a dictionary or open the book on certain pages (usually I don't, but my mental comfort increases and I'm not stressed so much, which also helps me come up with ideas and think faster during classes). (FG)*

Several participants commented that the absence of eye contact helped them concentrate and put them at ease.

*It's definitely less stressful for me, partially because an eye contact is distracting. (Q)*

*Online classes are less stressful because one does not have to look directly at the teacher. (Q)*

*I'm more stressful in class because the other students are looking at me. (Q)*

Some respondents found it easier to concentrate in the online setting, since there was only one speaker at a time, with other voices silenced. This additionally facilitated listening comprehension.

*Online learning is less stressful because classroom environment is louder. (Q)*

*Online classes stress me less because I can focus better on listening. (Q)*

There were also some comments that studying without the company of others was less anxiety-provoking than face-to-face learning. Some students felt more confident speaking to the computer than to a physical person:

*Well, I actually like much more learning in my home, in this environment. Speaking—I find it much easier through the camera, I'm slightly bit of an introvert, so this is also very helpful for me because I feel slightly bit uncomfortable when there is a lot of people around, yeah I think this is the main reason, I am a lonely learner, I am a slightly bit a loner, but not in a bad way, I like to be one in my room, for me it was good, ERT actually encouraged me to speak more than during physical classes, it helped me actually. (FG)*

Several respondents found ERT less anxiety-provoking because avoidance is easier to effect in an online environment than it is in a physical classroom.

*If you want to avoid participation, all you have to do is pretend your microphone doesn't work. (Q)*

*Online classes give me opportunity to put myself on mute and actually I do not have to pay attention all the time. (Q)*

### **Results for Research Question 2: What Are the FLCA Sources in ERT?**

The following causes of FLCA in ERT are derived from the data. They have been grouped under similar themes and ranked (in descending order) from most to least frequently mentioned.

Speaking online:

- Feeling that everybody is focusing on the speaker;
- Having to use a microphone;
- Not seeing other students.

Technological challenges:

- Experiencing unreliable Internet connections;
- Suffering from inadequate sound quality;
- Experiencing difficulty taking a test in an online format.

Online test anxiety.

Unnatural setting:

- Not seeing other students, lacking eye contact or non-verbal communication;

- Feeling that the classes seem unnatural;
- Lacking live contact or normal human interaction;
- Not being able to see other students' reactions;
- Having to use a camera.

Unsatisfactory instruction:

- Experiencing greater difficulty understanding teachers' explanations;
- Having inadequate contact with the instructor;
- Feeling a lack of learning, that classes are being used only for checking homework;
- Sensing a teacher's inability to see and appreciate effort;
- Finding less predictability in online classes;
- Getting few contributions from other students;

Fear of negative evaluation.

Lack of peer support.

The home environment as a classroom:

- Feeling embarrassed to speak a foreign language in front of one's family;
- Worrying about distracting household noises;
- Being unsettled by the lack of boundaries between home and university;

Fatigue and other issues related to student life during the pandemic.

**Speaking Online.** Speaking was shown to be the primary cause of FLCA in ERT, mainly because participants felt they were the center of attention:

*I'm more anxious during online classes because then I know everybody is listening to me when I speak. During normal classes, I can feel less watched. (Q)*

The background noise typical of a physical classroom was absent, so the fear of being judged was more intense. Participants were self-conscious when communicating through a microphone because they had the impression of being scrutinized by others:

*I never volunteer to speak. I feel stressed because of other students' muted microphones and the fact that everybody hears clearly what I am saying. (Q)*

Respondents clearly indicated that being called to speak without preparation caused the most apprehension. They also reported difficulties experienced when divided into pairs or groups, and the fear that their partner would be unavailable or unwilling to respond.

**Technological Challenges.** Technology-related apprehension was mainly attributable to the quality of the Internet connection. Several participants were concerned about the sound quality since it led to misunderstandings:

*I was most stressed when the professor couldn't understand me and when I couldn't understand the professor. (I can barely understand him so good sound quality is necessary). (Q)*

They also worried that they would not adequately be heard, or that they would not hear properly themselves:

*I feel stressed because I don't know if my voice is loud enough. (Q)*

Moreover, learners reported interruptions in connectivity and the associated anxiety experienced when trying to understand a task the rest of the group had begun. One participant mentioned the embarrassment he felt when he realized he had been speaking into a muted microphone. Participants also reported that they had been stressed when their cameras or microphones switched on unexpectedly and broadcast unwanted sounds or images. Stress related to the unreliability of technology was also mentioned.

**Online Test Anxiety.** Testing was a frequently mentioned anxiety cause. Many learners worried more about taking their final test online than they would about taking it in a physical classroom environment. Primarily, they were anxious because they had not experienced online testing before:

*I am more worried about the online form of the final test; so many things can go wrong. I don't know how it works and that makes me feel stressed. (Q)*

Some respondents were apprehensive about losing their Internet connection during the final test and failing as a result. Learners were also worried about the time limit (the test was automatically deactivated after a set amount of time) and concerned that after ERT they could “not be prepared as well as after regular studies.” Moreover, participants felt uneasy about the requirement to speak to a camera in oral tests.

**Unnatural Setting.** Another significant reason many respondents found ERT more anxiety-provoking than face-to-face learning was that it felt unnatural. They referred to the new mode of instruction as “strange,” devoid of “normal human interaction,” physical contact, and body language. They found it stressful to study without seeing other students and their reactions:

*I get confused when I'm supposed to talk to other people I can't see. This makes me stressed. In the classroom I feel more confident. (Q)*

*I am more anxious during my online language class because it is something different and I am not used to this. (Q)*

Additionally, participants remarked that not being able to see other students' reactions made online speaking more anxiety provoking:

*The fact that you don't see another [sic] people means that you can't really know how they reacted to what you've said. (Q)*

*I often choose to stay silent during my language class because I can't judge whether other students are laughing at me when they have their microphones muted and cameras switched off. (FG)*

While some teachers did require students to use the camera on their devices, results unfortunately demonstrate that respondents found this to be a major cause of anxiety. Several learners regarded the requirement to appear on camera as the most stressful aspect of online education. Some respondents admitted their unwillingness to reveal their private homes. Several learners mentioned they would not like to be seen wearing casual clothes that would not be appropriate in the school environment. Feeling self-conscious about one's appearance was also reported:

*In my opinion the most stressful thing about online learning is using the camera to show my face. I am not always looking good especially now. (Q)*

*I'm always worried that my teacher will want me to have the camera turned on, and I will have messy hair :) but jokes aside this really is stressful (like the camera thing, not hair). (FG)*

However, one participant did comment that making cameras obligatory, while stressful and initially difficult, would benefit learners in the long run, since being able to see other speakers' facial expressions and body language would facilitate communication in a foreign language.

**Unsatisfactory Instruction.** Some participants felt anxious because of the way Internet-based language learning was delivered. They thought online instruction was generally substandard and did not engage students sufficiently.

*I am more anxious because we only check homework in classes. It feels like we don't learn anything new. I would change our classes so that we*

*learn or explain new topics and then do homework. Our teacher doesn't present us with new topics. (FG)*

Some respondents expressed dissatisfaction with the lesson pace.

*I am an ambitious student as for English courses and I like having challenge, having regular tasks to do and during online classes the problem is that the courses are at a much slower pace than on stationary courses and so that we do much less in more time, and this is inefficient. (FG)*

Participants also complained that the lack of physical contact with the educator made classes difficult to follow:

*We had to work in groups, and I didn't know what I was supposed to talk about. (FG)*

Consequently, participants were apprehensive that they might lose the progress they had made so far.

**Fear of Negative Evaluation.** Many learners worried more about making mistakes and about being negatively evaluated online than in the physical classroom. The main reason for this was the lack of the background noise, which allowed others to hear their speech more clearly and focus on it more critically (see section Speaking Online).

*I feel slightly more worried because other classmates need to be completely silent during online lessons and a quiet environment seems to make your mistakes more noticeable. (FG)*

Participants admitted to feeling greater embarrassment when the instructor corrected their mistakes online, as other students seemed to pay more attention to their comments. One respondent noted that online correction felt more personal, as it seemed that the teacher was making an example of one particular student.

*In the classroom we don't feel evaluated. When we say something wrong, the teacher can say it correctly and the whole class repeats it, but during online classes the teacher seems to speak to this one person. (FG)*

There was a comment that errors made online caused more anxiety since they were "saved," although this only referred to mistakes made in the chatbox or on the interactive board. One participant pointed out that mistakes seem more

pronounced when the body language is absent, and one is unable to display their physical reactions to mistakes they make.

*When we are in the classroom, my embarrassment and my reaction to mistakes is visible, which reduces their scale. But when I am heard only, mistakes are more visible. (FG)*

**Lack of Peer Support.** Several participants felt more secure in a physical classroom, where they could ask their peers for help when they were finding it difficult to express themselves, for example when they were searching for words or struggling to complete a task. Additionally, they relied on peer support when they did not understand the teacher's instructions, or when they got distracted and needed someone to explain the task at hand. In the digital learning environment students found themselves alone and unable to ask for help discreetly, without losing face. Thus, the need to be wholly self-reliant added up to their anxiety.

*Online classes can be more stressful. I sometime lose the plot and there is no one nearby to tell me what to do. (Q)*

*I felt more comfortable in the classroom. I think it's because in a normal classroom someone can help me when I answer. (Q)*

**The Home Environment as a Classroom.** When educational institutions switched to online learning, students' private homes suddenly became "classrooms," which increased learner apprehension. One respondent found it challenging to sustain his motivation when the boundary between home and university was blurred.

*Over my entire life, I've got used to the fact that home is for home stuff and school is for learning, so now the two worlds have collided and it's extra difficult for me to force myself to do something. (FG)*

Several students felt uneasy speaking an FL in front of family members.

*I find it embarrassing to speak English to my computer in the presence of my family members, even if I am alone in my room and the door is closed. (Q)*

Household noise and distractions from cohabitants were also reported as sources of FLCA.

*I am more stressed when studying online because when I am speaking, my family members behaving in a noisy way may be heard. (Q)*

*I live alone but sometimes my parents or grandparents are visiting so I don't really know when they are visiting, so they can be there during the class and they are quite noisy I would say, so there is a lot of distraction that makes us not to focus on our lessons. (FG)*

### **Fatigue and Other Issues Related to Student Life during the Pandemic.**

In addition to becoming familiar with a new form of study, participants were under other pressures. One finding was that online learning caused greater fatigue than classroom-based instruction.

*Most of the times I am tired but stress doesn't play a huge role. (Q)*

*I'm more tired of online class. (Q)*

*I wouldn't say I had a much different emotion than I would have with normal face to face classes, but I sometimes would feel a little bit tired, and the classes would be so long and long. (FG)*

Another source of tension was the challenge of studying technical subjects online, which added to their workload and thus increased their overall anxiety.

*I can't fully concentrate on language bearing in mind that I already have 5 hours of painful remote laboratories. (FG)*

Participants repeatedly underscored the amount of coursework required by his major subject and pointed out that foreign language assignments contributed to his heavy workload.

*Sometimes I had to finish my projects for the same day that English class was happening and I was up until like 5 or 6 am and I woke up at 12 because we had lessons at 12 and I thought: "No, I cannot go through this, it's too difficult!" (FG)*

Furthermore, respondents suffered from stress resulting from having to do additional household duties that arose during the pandemic.

*Online learning is more anxiety provoking because I am looking after my younger brother, which is a major distraction. (Q)*

Others mentioned screen fatigue.

*It's pretty hard to concentrate on the computer while sitting on the computer like for 10 hours. (Q)*

### ***Results for Research Question 4: How Do Instructors Reduce FLCA in ERT?***

Most teacher participants claimed they tried to reduce students' apprehension during online classes, showing an awareness that learners were adapting to a new medium of instruction. Pedagogical approaches aimed at alleviating oral FLCA that appeared in teacher narratives include:

- Allowing extra time for thinking;
- Using a flipped classroom approach;
- Dividing learners into pairs or groups for speaking;
- Using the think-pair-share strategy;
- Arranging speaking assessment individually instead of in front of the group;
- Enabling screen-shared presentations;
- Using picture dictionaries and other visual aids;
- Engaging in small talk;
- Creating a friendly atmosphere;
- Giving encouragement;
- Allowing students to leave their cameras off.

For the most part instructors implemented these techniques intuitively, but one respondent reported that their choice of pedagogical approach was based on asking learners about their preferences.

Student narratives were mostly consistent with those of the teachers. Interestingly, learners mentioned several stress-alleviating techniques that were not reported by teachers. This may suggest that teachers used those techniques in a natural way, without a conscious effort to reduce oral FLCA. Anxiety-alleviating pedagogical approaches that appeared in student narratives include:

- Teacher manner;
- No pressure to use cameras;
- Similarity of instruction;
- Teacher's camera on;
- Soft error correction;
- Small talk;
- Personalisation;
- Individual approach;
- Appealing instruction.

**Teacher Manner.** Students valued teachers who were kind, supportive, warm, and cheerful, but also patient, understanding, and non-judgmental.

*My teacher has a magical aura to feel calm [sic] and not judging us. (Q)*

*My teacher is very calm and friendly, and it helps a lot. She is always happy to explain the problem, she is still smiling and is very positive. (Q)*

Since learners were under a lot of pressure from other subjects, they additionally praised educators who did not criticize them for not doing homework. Additionally, learners appreciated using humor and laughter.

*My teacher laughs with us, tells stories and creates a really good atmosphere. I think that helps everyone feel comfortable. (Q)*

*My teacher is very funny and outgoing, she makes us all laugh. (Q)*

Students also mentioned not being forced to answer questions, particularly without preparation.

*Everybody can contribute when they want. Nobody is called on to answer the question. (Q)*

Praise and encouragement to speak, albeit with mistakes, were also considered helpful.

**No Pressure to Use Cameras.** Participants reported feeling less anxious when allowed to work without cameras. (see section Unnatural Setting).

*I switch my camera off because it helps me to be more confident when I speak. (Q)*

*We don't need to turn our cameras on, and I think it helps people to speak more. (Q)*

**Similarity of Instruction.** Making online classes synchronous and, consequently, similar to face-to-face learning was a critical step that helped many learners adapt to ERT. Consequently, the switch was smooth and learners were reassured that their classes would continue in the similar manner. Break-out rooms helped to preserve social interactions among learners.

*Lessons run as in our previous classes, except that they are now online. This makes me feel less anxious. (Q)*

*Online classes were similar to face-to-face ones, we also had group speaking in break-out rooms, it was slightly easier because we didn't have to spend time getting into the other desks, the teacher would just make rooms and we would just chat with each other, so that's fine. The fact that classes were similar helped a lot. It did encourage some people who didn't speak a lot or sometimes they wouldn't speak ever and I think that addressed social needs in online classroom. (FG)*

**Teacher's Camera On.** Instructors preserved the natural character of online classes by keeping their cameras switched on.

*My teacher switches her camera on and it makes me reassured because I can see her facial expressions and thanks to that I can feel like she is in the same room as us. (FG)*

Additionally, teachers' cameras switched on facilitated understanding and communication thanks to the educator's facial expressions and body language. Learners found it comforting and easier to communicate when they could see the instructor react to what they were saying.

*Thanks to the camera I can see the teacher's emotions, for example a smile when we give correct answers. (Q)*

**Soft Error Correction.** A sensitive approach to mistakes seemed to alleviate fear of negative evaluation. Participants welcomed lack of criticism, guidance toward self-correction, explanations, and repetition of rules. This was particularly relevant in the ERT context, which made error correction seem more personal (see section Online Test Anxiety).

*There is no pressure if anybody makes a mistake. What is more, the teacher tries to explain the rules once again. (Q)*

*My teacher's tone of voice is so nice when she corrects my pronunciation that it does not feel like pointing out a mistake, but genuinely wanting to help. (Q)*

**Small Talk.** Small talk featured in both student and teacher narratives as a pedagogical tool to diminishing online language anxiety. Learners found it particularly beneficial at the beginning of a class. They felt comforted when

the teacher asked them how they were feeling, discussed the weather, or told a personal anecdote.

*We always do some small talk to feel more comfortable at the start of the lesson. (Q)*

*We chat in a natural way, so it does not feel like a lesson but pleasure. (Q)*

**Personalization.** Personalization seemed particularly effective in reducing FLCA. Students felt relaxed when asked about their lives during lockdown. They happily shared information about their weekend plans, their university assignments, or personal matters.

*Yes, our teacher tries to make us more relaxed, for example we talk about how our life looks during the pandemic. (Q)*

They welcomed the opportunity to show household pets to the class, using their cameras. This sharing of personal space made the virtual classroom more human.

*She is convincing other students to show their pets and it's great. She is the most cheerful teacher at this sad, sad university. (Q)*

**Individual Approach.** Participants valued teachers who addressed individual needs, for example, by giving more time to finish an exercise, offering help, or adjusting tasks to suit learners' abilities.

*The teacher decreases our stress level by offering help when someone is struggling with an exercise. (Q)*

Respondents were also appreciative when an instructor who observed signs of heightened anxiety allowed learners to relax, instead of exerting pressure. They acknowledged that recording a presentation for the instructor was less stressful than making it live in front of other students. Additionally, thoughtful selection of students of mixed abilities for break-out rooms was reported as a stress-reducing technique.

*My teacher knew that online learning is something new for us and he understood that there are differences between online and classroom learning, and he made adjustments to make those classes better. For example, when there were students who didn't want to talk about certain exercises, he made smaller groups and those groups I think he made with knowledge*

*about our differences because there was always a group in which there were students who couldn't speak fluently or cannot communicate fluently and students who were better and they encouraged him or her to speak, they engaged other people, so those groups were pretty thought through I'd say. Those groups were made specially to improve skills of some students and to make other students comfortable, when someone wasn't comfortable speaking. (FG)*

**Appealing Instruction.** Participants reported reduced anxiety when they were absorbed by varied lessons, visual aids, stimulating topics, exercises, or supplementary materials such as films, pictures or games.

*The teacher makes me feel comfortable because lessons are varied. There are language games, films, reading. (Q)*

*My teacher tries to make the atmosphere more relaxing by preparing different exercises about interesting topics. (Q)*

Clear explanations were also identified as a factor in helping learners feel at ease.

*I don't feel anxious because everything is always well explained, and we get satisfying answers to our questions. (Q)*

Some respondents spoke in favor of the “flipped classroom,” admitting they felt more relaxed when they received the topic of the next lesson in advance. It allowed them to review material and to prepare mentally for lessons.

## Discussion

The purpose of the first research question was to find out to what extent students experience anxiety during ERT. The main finding of the study was that participants reported increased and decreased levels of FLCA when faced with ERT. This is consistent with Maican and Cocoradă's (2021) observation that ERT provoked mixed emotions towards the new learning environment. Crucially, there was a group of learners that felt anonymous and thus more comfortable speaking once their camera was switched off, which corroborates Resnik and Dewaele's (2021) observations. Additionally, online education appeared to be more relaxing owing to the familiarity of the home environment

(Kaisar & Chowdhury, 2020), the removal of classroom background noise, the lack of physical presence, and the ease of avoiding tasks.

The second research question focused on the FLCA sources in ERT. There seems to be some overlap between FLCA and FLCA in ERT. The three physical classroom anxieties described by Horwitz et al. (1986) (i.e., communication apprehension, fear of negative evaluation, and test anxiety) apply additionally to ERT, but they seem to exhibit a modified profile when they are observed online.

Frustration with the method of instruction (Budzińska, 2015; Maican & Cocoradă, 2021) is another source of anxiety common to both online learning and face-to-face instruction. Emergency remote teaching, a product of the sudden transition to online learning, made some learners anxious about their progress (Maican & Cocoradă, 2021). However, only a small number of participants reported anxiety attributable to the quality of teaching, confirming Jelińska and Paradowski's (2021a) finding that instructors who taught synchronously in the higher education sector coped well with the adoption of digital technology.

The remaining FLCA triggers identified by this study, which involve technological challenges (Resnik et al., 2022), a lack of physical presence (Resnik & Dewaele, 2021), isolation (Sun & Zhang, 2021), and the home environment as a place of instruction (Adedoyin & Soykan, 2020; Maican & Cocoradă, 2021), are exclusive to the ERT. This concurs with Resnik et al.'s (2022) findings that the nature of anxiety arising from ERT is specific to that environment. It seems that these anxiety sources should also be relevant in computer assisted language learning.

It should be noted that the technological challenges reported were a result of weak Internet connections and unreliable equipment (Maican & Cocoradă, 2020; Resnik et al., 2022), rather than “technophobia” (Kaisar & Chowdhury, 2020) or using the instructional platform itself (Russell, 2020). This is probably because the participants—students at a technical university—were already familiar with a range of digital technologies.

In the present study, loneliness and social isolation were manifested in the loss of peer support and the need to be fully self-reliant. As expected, respondents were further stressed by challenges related to the lockdown, such as screen fatigue, heavier workload (Commodari & La Rosa, 2020), and distraction by co-habitants who were also forced to spend time at home (Maican & Cocoradă, 2021). These anxiety sources are characteristic to the COVID-19 period and are not likely to occur during planned online courses.

The third research question focused on pedagogical approaches reducing language anxiety. Jelińska and Paradowski (2021b) observed that “a humanistic approach and a pedagogy of compassion [...], while always crucial, gains particular importance in difficult times like these” (p. 9). As these results testify, instructors who took part in the study also did a lot to alleviate their learners' apprehension. This is quite remarkable given that educators had to adjust to

online delivery of courses overnight, putting an incredible strain on them, on top of increased levels of general anxiety attributable to the pandemic crisis (Mercer, 2020). Instructors' efforts and dedication were clearly recognized by learners participating in the study (Resnik & Dewaele, 2021).

Respondents cited a wide range of anxiety-reducing pedagogical approaches employed by their teachers. Teacher manner seemed to be one of the most frequently mentioned stress-alleviating aspects. The reason may be the perception of reduced distance between learner and teacher when communicating on a digital platform. In online education, the teacher is right in front of students on their screen. Personalization was also highlighted as a pedagogical approach that could alleviate student apprehension in ERT. Learners appreciated this approach as it showed the instructor's sincere, human interest in them, in contrast with a detached, impersonal style where the teacher functions as a knowledge-transmitting robot. This approach is particularly relevant in the light of Resnik and Dewaele's (2021) finding that "disembodied classes have less emotional resonance" (p. 1). Furthermore, respondents found that highly engaging instruction alleviated language anxiety, which is entirely consistent with the claim by Dewaele et al. (2018) that teachers should focus on increasing student FLE rather than reduce their FLCA.

Since language anxiety affects online learners as well as face-to-face learners, online instructors should take measures to lessen their apprehension. Accordingly, anxious online students will benefit from pedagogical approaches directed at them, such as those described in the present study. Here it is important to recall that the starting point when dealing with FLCA in the online learning environment is for the instructor to become aware that a problem exists, just as it is with face-to-face FLCA. It is then necessary to identify and assist anxious learners, who may mistakenly appear to be arrogant, lazy, unmotivated, or uninterested (Horwitz et al., 1986). In the virtual environment there will be learners who fail to log on and there will be those who use problems with their connection or equipment as excuses to avoid speaking.

ERT has altered the way languages are taught. Many language teaching institutions have started running synchronous online courses alongside traditional face-to-face ones as the demand for them has grown. This study was undertaken during the early stages of the COVID-19 pandemic. Future research could demonstrate how far this study was influenced by the rushed transition to online teaching in 2020, and whether FLCA sources during the post-COVID online language education remain the same.

The present study has some limitations. Firstly, the fact that students' comments on anxiety increase or decrease were self-reported involves a possible lack of objectivity. Secondly, the study did not find the exact proportion of learners whose anxiety dropped or stayed the same during the transfer to ERT. Future research could further explore the change in anxiety levels as a re-

sult of the shift to online language education. Thirdly, the effect of particular pedagogical approaches has not been empirically explored, as this was not the purpose of this study. Future research could investigate the stress-alleviating effect of particular approaches identified by the present exploration.

## Conclusion

The present study has examined the anxiety experienced by synchronous online language learners forced to study remotely due to the global COVID-19 pandemic. This type of anxiety seems to be specific to the online learning environment, and not just the combination of language anxiety with the apprehension inherent in using educational technologies. The findings demonstrate that while ERT contributed to lessening anxiety in some learners, other respondents reported increased apprehension compared with face-to-face learning. The study extends language anxiety research by closely inspecting the variables affecting FLCA increase and decrease in a new, insufficiently explored mode of instruction. The findings capture some “nuances of learners’ unprecedented experiences” (Resnik & Dewaele, 2021, p. 23) that have not been found before. Furthermore, the anxiety-reducing pedagogical approaches perceived by respondents as effective are relevant to online education in general and can be employed globally by others to promote wellbeing of language learners.

## References

- Adedoyin, O. B., & Soykan, E. (2020). Covid-19 pandemic and online learning: The challenges and opportunities. *Interactive Learning Environments*, 31(2), 863–875. Retrieved March 31, 2021. <https://doi.org/10.1080/10494820.2020.1813180>
- Alemany-Arrebola, I., Rojas-Ruiz, G., Granda-Vera, J., & Custodio Mingorance-Estrada, Á. (2020). Influence of COVID-19 on the perception of academic self-efficacy, state anxiety, and trait anxiety in college students. *Frontiers in Psychology*, 11. Retrieved March 25, 2021. <https://doi.org/10.3389/fpsyg.2020.570017>
- Bao, W. (2020). COVID-19 and online teaching in higher education: A case study of Peking University. *Human Behavior and Emerging Technologies*, 2(2), 113–115. <https://doi.org/10.1002/hbe2.191>
- Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, 12(2), 205–232. <https://doi.org/10.14746/ssl2022.12.2.3>

- Budzińska, K. (2015). *Foreign language classroom anxiety in a further education context*. Unpublished PhD thesis, University of Łódź.
- Charmaz, K. (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Sage.
- Commodari, E., & La Rosa, V. L. (2020). Adolescents in quarantine during COVID-19 pandemic in Italy: Perceived health risk, beliefs, psychological experiences and expectations for the future. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2020.559951>
- Coryell, J. E., & Clark, C. (2009). One right way, intercultural participation, and language learning anxiety: A qualitative analysis of adult online heritage and nonheritage language learners. *Foreign Language Annals*, 42(3), 483–504. <https://doi.org/10.1111/j.1944-9720.2009.01037.x>
- Derakhshan, A., Kruk, M., Mehdizadeh, M., & Pawlak, M. (2021). Boredom in online classes in the Iranian EFL context: Sources and solutions. *System*, 101, 102556. <https://doi.org/10.1016/j.system.2021.102556>.
- Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety in the FL classroom: The effect of teacher and learner variables. *Language Teaching Research*, 22(6), 676–697. <https://doi.org/10.1177%2F1362168817692161>
- Fawaz, M., & Samaha, A. (2020). E-learning: Depression, anxiety, and stress symptomatology among Lebanese university students during COVID-19 quarantine. *Nursing Forum*, 56(1), 52–57. <https://doi.org/10.1111/nuf.12521>
- Fraschini, N., & Tao, Y. (2021). Emotions in online language learning: Exploratory findings from an *ab initio* Korean course. *Journal of Multilingual and Multicultural Development*, 42, 1–19 <https://doi.org/10.1080/01434632.2021.1968875>
- Gao, L. X., & Zhang, L. J. (2020). Teacher learning in difficult times: examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.549653>
- Gregersen, T. (2020). Dynamic properties of language anxiety. *Studies in Second Language Learning and Teaching*, 10(1), 67–87. <https://doi.org/10.14746/sslft.2020.10.1.4>
- Hodges, C., Moore, S., Locke, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. *Educause Review*. <https://er.educause.edu/%20articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning> (accessed 27 April 2022).
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70(2), 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 31–50). Multilingual Matters.
- Jelińska, M., & Paradowski, M. B. (2021a). Teachers' engagement in and coping with emergency remote instruction during COVID-19-induced school closures: A multinational contextual perspective. *Online Learning Journal*, 25(1), 303–328. <https://doi.org/10.24059/olj.v25i1.2492>
- Jelińska, M., & Paradowski, M. B. (2021b). Teachers' perception of student coping with emergency remote instruction during the covid-19 pandemic: The relative impact of educator demographics and professional adaptation and adjustment. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2021.648443>
- Jin, Y., & Zhang, L. J. (2021). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. *International Journal of Bilingual Education and Bilingualism*, 24(7), 948–962. <https://doi.org/10.1080/13670050.2018.1526253>

- Kaisar, M. T., & Chowdhury, S. Y. (2020). Foreign language virtual class room: Anxiety creator or healer? *English Language Teaching*, 13(11), 130–139. <https://doi.org/10.5539/elt.v13n11p130>
- Lee, J. S., & Hsieh, J. C. (2019). Affective variables and willingness to communicate of EFL learners in in-class, out-of-class, and digital contexts. *System*, 82, 63–73. <https://doi.org/10.1016/j.system.2019.03.002>
- MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 24–45). McGraw-Hill.
- MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11–30). Multilingual Matters.
- MacIntyre, P. D., 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(4), 1–13. <https://doi.org/10.1177%2F1362168820965897>
- Maican, M.-A., & Cocoradă, E. (2021). Online foreign language learning in higher education and its correlates during the COVID-19 pandemic. *Sustainability*, 13(2), 781. <https://doi.org/10.3390/sul3020781>
- Mercer, S. (2020, June 8). *Teacher wellbeing in a time of crisis: Lesson for now and the future*. [Video]. YouTube. <https://www.youtube.com/watch?v=3TL3foXoFII>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage.
- Pawlak, M., Derakhshan, A., Mehdizadeh, M., & Kruk, M. (2022). Boredom in online English language classes: Mediating variables and coping strategies. *Language Teaching Research*, 26(4), 598–616 <https://doi.org/10.1177/13621688211064944>
- Pichette, F. (2009). Second language anxiety and distance language learning. *Foreign Language Annals*, 42(1), 77–93. <https://doi.org/10.1111/j.1944!9720.2009.01009>
- Resnik, P., & Dewaele, J.-M., (2021). Learner emotions, autonomy and trait emotional intelligence in 'in-person' versus emergency remote English foreign language teaching in Europe. *Applied Linguistics Review*, 4(3), 473–501. <https://doi.org/10.1515/appl-rev-2020-0096>
- Resnik, P., Dewaele, J.-M., & Knechtelsdorfer, E. (2022). Differences in the intensity and the nature of foreign language anxiety in in-person and online EFL classes during the pandemic: A mixed-methods study. *TESOL Quarterly*, 57(2), 618–642. <https://doi.org/10.1002/tesq.3177>
- Russell, V. (2020). Language anxiety and the online learner. *Foreign Language Annals*, 53(2), 338–352. <https://doi.org/10.1111/flan.12461>
- Sun, J., & Zhang, X. (2021). Exploring Chinese college students' emotions as they engage in online learning during a pandemic. *Asia Pacific Journal of Education*, 43(4), 984–955. <https://doi.org/10.1080/02188791.2021.1965541>
- Young, D. J. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *Modern Language Journal*, 75(4), 426–437. <https://doi.org/10.1111/j.1540-4781.1991.tb05378.x>

## Appendix A

**Student Questionnaire**

1. Are you more anxious during your online language class than your regular class in the classroom? Why?
2. Are you more worried about making a mistake during your online foreign language (FL) class than during your regular class in the classroom? Why?
3. Are you more worried about your pronunciation during your online class than during your regular class in the classroom? Why?
4. Are you more anxious about your final test after studying online? Why?
5. Which aspects of online learning did you find anxiety-provoking?
6. Which online activities make you feel anxious? Why? Would you prefer to avoid them?
7. Does the teacher do anything to make you feel comfortable during your online classes?
8. Have you ever felt stressed because of technical issues?

## Appendix B

**Focus Group Questions**

1. Are you anxious during your online language class? What are the symptoms?
2. Do you feel the same during your classes in the classroom?
3. Would you benefit more from your classes if you were not anxious?
4. Do you participate in an active way despite your anxiety? Why?
5. If you were to continue studying a foreign language online, how would you change your classes to make you feel better?
6. What influences your FLCA in ERT?

## Teacher Questionnaire

1. Do you think that an online foreign language class is more anxiety-provoking than face-to-face learning?

2. Do you do anything to reduce your students' anxiety during your online class? If so, what do you do?

Katarzyna Budzińska

### **Angst im Fremdsprachenunterricht in Emergency Remote Instruction**

#### Zusammenfassung

Obwohl Sprechangst die am gründlichsten untersuchte Emotion bei dem Zweitspracherwerb ist, ist noch immer unzureichend bekannt, inwieweit Emergency Remote Teaching (ERT) zur Auslösung von Angst führt. Die vorliegende Studie füllt diese Lücke, indem sie Variablen analysiert, die einen Einfluss auf Angsterfahrung in der neuen Unterrichtsform haben, und dabei auch angstlösende pädagogische Ansätze erforscht. Die qualitative Studie wurde unter 218 Universitätsstudenten durchgeführt. Die Daten wurden mittels eines Online-Fragebogens bzw. eines Fokusgruppeninterviews erhoben. Zusätzlich wurde ein Online-Fragebogen verwendet, um herauszufinden, wie die Lehrkräfte ihre Lehrmethoden anpassen, um das Wohlbefinden der Studierenden in ERT zu gewährleisten. Die Ergebnisse weisen auf mehrere Faktoren hin, die das Angstniveau beeinflussen, sowie auf pädagogische Ansätze, die sich nach Meinung der Studierenden als wirksam für Abbau von Angst erwiesen. Insgesamt erweitert die Studie die Sprechangstforschung, indem sie die Variablen, welche die Zunahme und Abnahme von Angst (FLCA) in einer neuen, unzureichend erforschten Unterrichtsform beeinflussen, genauer betrachtet.

*Schlüsselwörter:* COVID-19, Angst im Fremdsprachenunterricht, Emotionen beim Sprachenlernen, Emergency Remote Teaching, tertiäre Bildung, synchroner Online-Unterricht



**Nettie N. Ndou**

University of Pretoria, South Africa

 <https://orcid.org/0000-0003-0700-9756>

**Segun Emmanuel Adewoye**

University of South Africa, South Africa

 <https://orcid.org/0000-0001-8323-0933>

## **Challenges of Learning in Second Language among South African School Learners with Developmental Language Disorder**

### **Abstract**

Language in education plays a critical role in effective teaching and learning worldwide. This study aimed to explore the challenges of learning in a second language among secondary school learners with developmental language disorder (DLD). The study also unveils strategies used by professionals to support learners and learners' attitudes towards support. The study participants were learners ( $n = 12$ ), teachers ( $n = 5$ ), a speech Language therapist, and an educational psychologist. A qualitative research approach was utilised employing a case study as the research design. Data generation sources included non-participatory observations, interviews, and focus group discussions. Data was analysed using thematic analysis. Findings indicated that learning in a second language as the medium of instruction posed challenges for learners with DLD. They struggled with comprehension, reading, and word recognition of learning content. Teachers indicated that they used strategies such as remedial lessons and giving learners more reading material to enhance their vocabulary and narrative skills. They also referred learners to other professionals for further assistance. Parental involvement is also encouraged in support of learners. However, one of the factors limiting learners' support is learners' attitudes. Learners' attitudes towards support may be attributed to limited awareness of DLD. Hence, there is a need to raise awareness of DLD among the learners. An in-depth course on special education in secondary teacher training programmes is necessary to further equip teachers with strategies to enhance inclusive classrooms. The promotion of local languages as a medium of instruction must be highly prioritised, even at a secondary level of education.

*Keywords:* developmental language disorder, second language learning, medium of instruction, learning challenges, home language

Success in school learning depends heavily on learners' ability to understand language use as a medium of instruction. Language in education plays a key role in effective teaching and learning worldwide (Khudsen, 2012). In post-apartheid South Africa, English has become the preferred medium of instruction in schools (Manyike & Lemmer, 2014). After 1994, the South African government embarked on a series of incisive efforts to shape language policy in education.

All eleven South African languages were given equal status in the South African constitution in which South African learners have the choice of being taught in their first language. However, due to limited resources and minimal oversight by the government, South African school teachers are not enthusiastic about tutoring in African languages (Hazeltine, 2013). Moreover, there are cases where teachers and learners code-switch and code-mix English with a local language during teaching and learning process. For instance, a study conducted in Limpopo showed that teachers use code-switching and a trans-language process, alternating and merging languages to assist learners to understand concepts (Kretzer, 2019).

In most South African schools, learners are required to learn and use vocabulary in a second language (Ossai & Uzoegwu, 2019). In such circumstances, learners with DLD may have difficulty developing cognitive academic language proficiency (CALP). DLD is a neuro-developmental condition that impairs spoken or written language and is not associated with any known causal condition (McGregor et al., 2020). Learners with this condition struggle to learn, comprehend, and use spoken and written language (McGregory, 2020; Gillespie, 2015). Hence, using a second language as a medium of instruction may worsen communication challenges (Starling et al., 2011).

Communication difficulties substantially impact learners' academic progress and educational attainment (Tribushinina et al., 2020). Ngulube (2015) noted that such learners acquire language structures of English but find it hard to put them into understandable texts. In the same vein, Cakiroglu (2019) also highlighted that it is difficult for learners to acquire English as a second language because they lack comprehensible input to facilitate understanding.

Reading as a receptive skill assists learners in accumulating vocabulary; however, lack of it makes learners fail to infer meaning from texts (Cakiroglu, 2019). Manyike and Lemmer (2014) also argued that due to the use of English as the medium of instruction in South Africa, low academic achievement among black learners is revealed in the annual national school-leaving tests and annual literacy and numeracy assessments. Consequently, learners may develop negative attitudes towards using a second language as the medium of instruction and may not be able to participate effectively in classroom activities (Blair et al., 2018).

Krashen (1988), in his second language development theory, highlighted two independent systems of second language development, that is, acquired and learned systems. The acquired system is the result of meaningful interac-

tion using the target language. The process is similar to when children acquire their first language. The learned system is the product of formal instruction. It results in a child's knowledge of the language (e.g., grammar rules). The difference between the two systems is that the acquired system is the product of a subconscious process, while the learned system results from a conscious process (Krashen, 1988). Krashen (1988) added that Second Language Learning (SLL) does not facilitate Second Language Acquisition (SLA).

The learned system is generally used today in most schools where attention is focused on using the English language and its written form. The main objective for the learner is to understand the structure and rules of the language. Schütz (2007) argues that such a task requires intellectual effort and deductive reasoning, which a learner with DLD may struggle to attain.

Due to less knowledge and paucity of literature on challenges faced by South African school learners with DLD in learning a second language as an MOI, the individual specialised needs of South African learners with such a condition are not being adequately catered for in the mainstream classrooms (Manyike & Lemmer, 2014). Although evidence of support strategies for addressing barriers to learning exists (Adewoye, 2022), there are limited empirical findings on the challenges that a second language as a medium of instruction poses for effective teaching and learning among South African learners with DLD.

Therefore, the purpose of this study is to explore the challenges of learning in second language among South African school learners with DLD to propose effective instructional communication strategies suitable for implementation within the South African context.

Specifically, this study sought to answer the following research question:

- What are the challenges of learning in second language among secondary school learners with DLD?

## **Methodological Designs**

### **Research Approach and Design**

This study was conducted using a qualitative research approach. According to Basit (2010), qualitative research methods are suited for studies that seek to derive findings from participants' views. We used case studies to investigate real-life situations over time by collecting detailed, in-depth data from differ-

ent sources (Creswell, 2014). Two cases were purposively selected: Zimbabwe (Case 1) and South Africa (Case 2). However, this study reports the findings in South Africa (Case 2). Multiple case studies provided rich descriptions and interpretations on the phenomena under study and the replication of data collection across multiple sites helped to understand the challenges faced by learners with DLD in learning through a second language in different contexts.

### **Participants and Sampling Strategies**

The inclusion criterion was based on the aim of gathering evidence from participants. Three schools in Gauteng were purposively chosen to participate in the study. This decision was made because of our familiarity with the area. Twenty participants, including 12 learners (six boys and six girls), six teachers, a speech-language therapist and an educational psychologist, were purposefully selected to participate in the study. The learners participating in this study were in Grades 7 and 8. They were between the ages of 13 and 15.

English language teachers were asked to select learners with mild DLD in their respective schools. The teachers selected the participants based on their familiarity and understanding of the language challenges of the participants. The research design dictated the number and type of participants; as a qualitative study, the goal was to elicit in-depth perspectives of the participants on the challenges of learning in second language among learners with DLD.

### **Data Collection Techniques**

In this study, semi-structured interviews, classroom observations and focus group discussions were used to collect data. Classroom observations were conducted in both schools. Focus group discussions were conducted with learners who were purposively selected by their respective teachers. Semi-structured interviews were also conducted with the teachers, speech-language therapists, and an educational psychologist. Telephone interviews were used for member checking during data analysis.

Driscoll (2011) purports that the core advantage of mixing different research techniques is that since each method has strengths and weaknesses, one technique can cover the weaknesses of the other. Therefore, the concept of triangulation was highly utilised in this study. Triangulation is used to determine how different methods check, validate, or collaborate one another. This enables an understanding of a social phenomenon from different viewpoints (Rich-Mahadkar, 2015).

## **Ethical Considerations**

Ethics clearance from the author's institution was obtained. Permission to conduct research in Gauteng, South Africa, was granted by the Gauteng Department of Education. Throughout the study, informed consent, confidentiality, and voluntary involvement were all incorporated.

## **Data Analysis and Interpretation**

The qualitative data was analysed using thematic analysis. The procedure involved searching across data sets to identify, describe, and interpret data, and analyse and report repeated patterns (Kiger & Varpio, 2020). The audio recordings of the interviews were transcribed. According to Braun and Clarke (2006), inductive thematic analysis is a common method of qualitative research analysis to identify, investigate, and record patterns in data. We coded the data to identify and build significant patterns. This includes: (a) getting familiar with the data, (b) creating initial codes, (c) looking for themes, (d) reviewing the themes, (e) defining and labelling the themes, and (f) writing up the findings.

## **Reliability of the Study**

To ensure the credibility of the research findings, we prolonged the engagement with participants and triangulated data sources. Space triangulation was also used to overcome parochialism (Cohen et al., 2011). Also, transcribed data was emailed to the participants for verification (Gunawan, 2015).

## **Findings**

The outcomes of the data analysis were recorded and categorised into three themes. These themes were sub-divided into eight sub-themes. In this article, we discuss sub-themes on the challenges faced by learners with DLD in using a second language as the medium of instruction. We also investigated learner attitudes towards support strategies as one of the factors affecting the effective implementation of support by professionals.

## Difficulty in Comprehension

Learners highlighted the challenges they faced learning with a second language, English. They mentioned that understanding instructions given in English was a challenge; for instance, one of the participants stated:

Most times I don't understand instructions when the teacher is instructing me using English language. (P4)

Teachers indicated that learners with DLD demonstrate little comprehension of topics taught because it is difficult for them to interpret and respond to questions appropriately. According to the teachers, such learners frequently concentrate on important words while ignoring the rest of the text, which affects their comprehension of the questions. A teacher stated:

I realised when I use English throughout the lesson without the use of any vernacular language, learners with DLD are left behind. They show little comprehension of the lesson taught. (P3)

Other teachers shared a similar view that using a second language as the medium of instruction negatively affected learners with DLD academically. The use of a second language as the medium of instruction affects learners' performance in classroom activities. A teacher shared the following thoughts:

Yes, it does affect their academic performance because they find it difficult to comprehend what they are being taught when using [a] second language. Eventually, it affects their performances in tests and exams. (P7)

The Speech Language therapist mentioned that using English as the medium of instruction affects learners with DLD. This is because they struggled to acquire language at an appropriate level of proficiency. She further explained:

Well, we are looking at children with DLD, but the language of instruction is English, right? We are already burdening a child who has a burden, right? We are bringing in this foreign language which to them is difficult, yet they are struggling with their own local language. Hence, we have a very serious problem. (P1)

As a result of their difficulties in comprehension, learners with developmental language disorders perform poorly in all learning areas. One of the learners declared:

I struggle to read and interpret word sums in Mathematics; I cannot comprehend and interpret what the questions are, especially in Social Sciences and Natural Sciences. (P6)

### **Difficulty in Reading**

During the classroom observation in schools, we noticed that learners with DLD experienced difficulties reading texts. Some of these learners were passive. They could not contribute to anything during the class activities. Most learners feel too embarrassed to participate during lessons because of the fear of being laughed at. To substantiate this observation, a teacher stated thus:

The use of [a] second language as the medium of instruction negatively affects these learners in reading, and they become very passive during lessons to avoid embarrassment. Most [of the] time, they fail to participate in class activities because they can't read in English. They withdraw from learning. They just keep quiet in the classroom because of fear of embarrassment (P18)

When asked why learners with DLD find reading English text difficult, an educational psychologist explained that:

The use of a second language as the medium of instruction kind of put a learner with a language disorder at further disadvantage because they are already disadvantaged by their condition. The use of English as the medium of instruction makes matters worse for such learners because they already struggle acquiring their local language at [an] appropriate level of proficiency. (P5)

Learners must be able to answer questions requiring them to read instructions, explain, suggest, solve, discuss, analyse, comprehend, and justify aspects of the work to succeed in school. However, learners with DLD experience difficulties with their reading tasks. Their grammatical and textual competency is limited, contributing to their anxiety in class. As a result of limited vocabulary, they are unable to construct logical phrases and paragraphs to compose meaningful text.

The same question was posed to a teacher who answered thus:

Learners with developmental language disorder do not like reading tasks because they sometimes read without understanding and fail to answer the comprehension test questions. As a result, they often perform poorly in their assessment tasks, which make[s] them feel inferior to their peers. (P8)

## Difficulty in Word Recognition

Almost all the teachers indicated that most learners with DLD have trouble with word recognition. Their inability to recognise or use words correctly leads to poor performance. The excerpts that follow illustrate the challenges that difficulty in word recognition posed for such learners.

A lack of vocabulary forces them to produce work of poor quality as they cannot do as well as their peers. They do not understand idioms, proverbs, and expressions. (P3)

Difficulty in recognising words while reading confuses the learners. The learners face various challenges, such as word problems, comprehension problems, as well as problems in creative writing activities. They scored poorly in tests and examinations owing to poor writing quality and an inability to interpret questions appropriately. One teacher buttressed this point by saying:

They lack the skills to recognise and pronounce important words in a sentence. If they cannot pronounce such words, the learners would just omit them without making any further attempts to complete the task. In addition, the learners struggle to pronounce words correctly. They mix vowel sounds such as *a*, *e*, and *i*. As a result, they end up guessing where to place the stress and subsequently read words incorrectly. (P4)

Learners were also asked to narrate their experiences regarding being taught in school with a second language to shed more light on this. One learner stated:

I don't like being taught with English language because the words are confusing, I don't always understand what is required of me. I find it very difficult to interpret sentence[s], when I try, others laugh at me. (P2)

Using the English language for practical purposes is a significant challenge for learners with a developmental language disorder. They failed to ask questions in class and write paragraphs in English because they have difficulty understanding, reading, and recognising words, as well as spelling and sentence construction. On some occasions, they knew and understood tasks, but their grammatical deficiencies made it difficult to write sentences appropriately, affecting their academic performance in all learning areas. They reportedly struggled to write down their ideas properly.

## Learners' Attitudes towards Support

Despite challenges highlighted by the learners, the teachers argued that there are support systems to assist such learners (see Table 1).

**Table 1**

*Support Systems for Learners with DLD*

Participants	Support system for Learners with DLD
P18	Giving learners more books to read, Remedial lessons, More parental involvement, Referrals to the District-Based Support Team (DBST).
P19	Remedial lessons, Giving learners more books to read.
P20	Remedial lessons, Availability of more books for learners to read, Referring learners to DBST.

However, one of the factors impeding the effective implementation of these strategies is the learners' attitude towards support. One of the teachers highlighted that such learners do not cooperate; most of them are aggressive and uncontrollable. She also added that it was difficult to monitor them as the classes were too big:

They are many learners who are not supposed to be here, honestly. They are supposed to be doing vocational subjects. Because of their language disorder, they become so embarrassed. Some become uncontrollable, rude, and aggressive sometimes. They hide out because our classes are just too big. (P4)

Furthermore, one learner confirmed that he had challenges in understanding English as the medium of instruction, but he could not disclose it to his teacher.

I honestly lose focus in class [laughs]. I do have challenges in understanding lessons taught in English, I am not open, so my teachers will be thinking I am okay while not okay. [I] am not that open, you know. (P9)

Some learners also highlighted that they avoided contributing in class because the other learners would laugh at them:

I cannot participate in class. Some students always laugh when I miss the correct answer. If you spell the word wrong, they laugh in class, so I will be ashamed, you know. So, it is best to keep your cool in class, you know. (P10)

Learners' attitudes also affect the referral process; teachers cannot refer learners for further support to the speech-language therapist as they try to hide their condition. This explains why very few learners are referred to speech-language therapists at the secondary level of education. The speech-language therapist argued that she received referrals mostly from the primary level compared to the secondary level. This means that many unidentified learners with DLD in the mainstream classrooms are not receiving support from professionals.

## Discussion

Ability to communicate is a fundamental skill required for effective participation in school activities. Learners with DLD are facing significant language difficulties (Ziegenfusz et al., 2022), especially in education contexts where the second language is used as the medium of instruction. This study's findings revealed that using a second language as the medium of instruction posed challenges for learners with DLD. Learners experienced difficulties in comprehension, reading, and word recognition. Teachers reported that learners with DLD mispronounced words and did not fully comprehend text. Hence, they hardly understand the lessons being taught.

Existing literature is replete with the challenges of using a second language as the general medium of instruction for learners. Studies by Khan and Khan (2016), Krugel and Fourie (2014), Nawaz et al. (2015), Ngulube (2015), as well as Cakiroglu (2019) have concluded that the use of a second language as the medium of instruction also contributes to poor academic achievement of learners. In Australia, O'Connor et al. (2015) concluded that learners who entered school were not yet proficient in English, were at risk of experiencing low education outcomes. In Tanzania, Lupogo (2014) concluded that using a second language in vocational education training was a disadvantage because most students failed to understand the language of instruction. The findings of this study also resonate with those of Vuzo (2018), who argued that using a second language in Tanzania contributed to school dropout because using unfamiliar language as a medium of instruction contributes to learners' lack of interest in learning.

The presence of language difficulties has a significant impact on the academic success and educational attainment of learners with DLD. Such learners often function at the Basic Interpersonal Communication Skills (BICS) level and do not develop Cognitive Academic Language Proficiency (CALP) (Lillywhite, 2011). Academic language skills include inferential language, that is, communicating ideas across contexts and using a narrative language where learners are expected to describe a series of events and understand the range of academic vocabulary and grammatical structures (Wissel, 2016). It is, therefore, not surprising that such learners are in dire situations in terms of poor academic achievement (Tuite, 2019). Thus, considering the nature and characteristics of developmental language disorders, support is necessary to include and accommodate such learners in school.

Inclusive education refers to a wide range of strategies and processes that seek to make the universal right to quality, relevant, and appropriate education a reality. "It is about changing the system to fit the student, not changing the student to fit the system. It locates the 'problem' of exclusion firmly within the system, not the person or their characteristics" (Stubbs, 2008, p. 8). Therefore, the education system must have supportive structures and services to support the needs of all learners.

Teachers reported that they do support learners with DLD. They gave them more books to read and conducted remedial lessons. These strategies enhanced learners' vocabulary and narrative skills (Joffe et al., 2019; Murphy et al., 2017; Lowe & Joffe, 2017; Spencer et al., 2017). The teachers also explained that they encouraged more parental involvement during support. In some cases, they referred such learners to the District-Based Support Team for further assistance and support. However, learners' attitudes towards support affect accurate identification and effective support.

Although the findings of this study indicated that learners struggled with reading, comprehension, and word recognition, they had developed compensatory strategies to hide their challenges. They resorted to compensatory strategies such as absenteeism, not responding to questions, avoiding oral presentations, and not attending lessons. This made it very difficult for their teachers to identify and support them. The National Behaviour Support Service [NBSS] (2011) highlighted that many learners with DLD may have an undetected or hidden difficulty with language acquisition because they often develop compensatory strategies, for example, always agreeing or disagreeing with the conversational partner, remaining silent or responding using learned phrases or being absent.

There is a stigma attached to being given support among learners with DLD in mainstream secondary school. Support is viewed as a punishment and an embarrassment. Such learners have low self-esteem and lack confidence. A lack of confidence emanates from the labels attached to the learners (Miller, 2019). Such learners experience a variety of difficulties, socially and emotionally.

Social withdrawal has been a particular concern for such learners, particularly reticent withdrawal, or shyness in a school setting (Miller, 2019). Miller (2019) added that learners' social withdrawal patterns co-occur with poor social outcomes, such as difficulties in establishing friendships, social isolation, and victimisation. Raines et al. (2012) argue that learners with special needs such as DLD may suffer social isolation, lower self-esteem, substandard education, and are twice as likely to drop out of school.

There is limited awareness of the importance of speech-language communication needs by teachers, general school staff, and the community. This shortcoming is because there is an assumption that language development and early intervention only happen early in the primary years and not in secondary school. There is a narrow understanding of the role of language in the secondary school curriculum. Hence, fewer professionals have the skills, knowledge, resources, and assessment tools to identify or support learners with DLD on time. So, learners are labelled "dull" with "deviant behaviour" and just fall through the system's cracks.

Furthermore, education training for teachers in mainstream classrooms rarely prepares teachers for working in diverse classrooms, and it does not equip teachers with the knowledge and skills to support learners with DLD effectively (Krystal, 2015; Zwane & Malale, 2018). Handling learners with special needs requires extensive expertise. Also, parenting children with challenges such as DLD requires a high level of knowledge and access to resources, information and services. In developing countries such as South Africa, such services are not always available to every learner with DLD challenges (Taderera & Hall, 2017).

## **Conclusion and Recommendations**

This study contributes to the literature on the challenges that learners with DLD face in secondary schools. Tribushinina et al. (2020) noted that there is limited literature on the challenges of using second language as a medium of instruction for such learners. Therefore, this study explored the challenges such learners face in secondary school where second language is used as a medium of instruction. These challenges affect learners in all learning areas and overall academic performance. By using a qualitative research approach, professionals and learners could provide detailed descriptions of their experiences on the topic under study, which would have been limited if a quantitative approach had been utilised.

In-depth courses on special education in secondary teacher training programs are necessary to equip teachers with knowledge of inclusive education.

Interprofessional collaboration is recommended, especially between teachers and speech-language therapists. This will ensure effective support of learners with DLD. Professionals will share their knowledge and expertise, and this will enhance the support of the learners in mainstream schools. Assessment tests used to identify and support learners with DLD must be relevant to learners in multilingual contexts.

There is a need for a paradigm shift in the way the second language is used and taught in schools. Krashen's language acquisition and learning systems must be used to cater to all learners' needs. The medium of instruction must be inclusive to accommodate the needs of learners with language acquisition challenges, such as learners with developmental language disorders. Raju and Joshith (2018) argued that learners need to get a wide variety of language inputs for better acquisition through reading books, having conversations, and listening to the target language. The teachers are responsible for ensuring that the materials and tasks they use in their classrooms are organised and implemented to meet individual learner's needs (Ivančević-Otanjac, 2016).

The promotion of local languages must be highly prioritised, even in secondary school. Literacy in local languages must be developed beyond the decoding of narrative texts. Learners with developmental language disorder should therefore be given a chance to develop their first language fully, then they should gradually be introduced to the second language. The learner should acquire BICS and CALP in their local languages before acquiring and learning a second language. It must be highlighted that high levels of cognitive language proficiency are required for successful learning, and as learners move from one level of education to the next, learning becomes increasingly more complex and demanding. Therefore, it might be unfair for such learners to be expected to go through such a learning process using a second language.

The school can hire more teachers who can teach in local languages to assist learners with DLD. Such teachers could assist learners who require additional clarification on lessons or topics discussed in class, as well as those who do not possess a sufficient understanding of the English language. Support teachers can assist in supervising learners' homework in the afternoon for one hour after school before the extracurricular activities and extra classes begin. It might be easier for learners to access learning if they could overcome the language barrier.

## Acknowledgments

We appreciate the principals of the schools where this study took place and all the participants for cooperating with us before, during, and after data collection.

## Disclosure

There is no conflict of interest.

## Author Contributions

Both authors co-conceptualised the idea for the research, the design and methodology adopted, the analysis of the data and the finalisation of the study results. Ndou N. N. was the lead author.

## References

- Adewoye, S. E. (2022). Factors that enhance flexible teaching of learners with special needs in South Africa in the COVID-19 era: Implications for post-COVID-19 classrooms. *E-Journal of Humanities, Arts and Social Sciences*, 3(11) 213–225. <https://doi.org/10.38159/ehass.2022sp31117>
- Basit, T. N. (2010). *Conducting research in educational contexts*. Continuum International Publishing Group.
- Blair, A., Haneda, M., & Bose, F. N. (2018). Reimagining English-medium instructional settings as sites of multilingual and multimodal meaning-making. *TESOL Quarterly*, 52(3), 516–539. <https://doi.org/10.1002/tesq.449>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in psychology*, 3, 77–101.
- Cakiroglu, A. (2019). The language acquisition approaches and the development of literacy skills in children. *International Electronic Journal of Elementary Education*, 11(2), 201–206. <https://doi.org/10.26822/iejee.2019248600>
- Cohen, L., Manion, L., & Morrison, K. (2011). *Research methods in education*. Routledge.

- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, mixed methods approaches*. Sage.
- Driscoll, D. L. (2011). Introduction to primary research: Observations, surveys and interviews. *Writing Spaces: Readings and Writing*, 2, 153–174.
- Gillespie, T. (2015). Language differences versus language disorder. *Journal of Educational Psychology*, 112(5), 1042–1060. <https://doi.org/10.1037/edu0000394>
- Gunawan, J. (2015). Ensuring trustworthiness in qualitative research. *Belituns Nursing Journal I(1)*, 10–11. <https://doi.org/10.33546/bnj.4>
- Hazeltine, R. (2013). *Language policy and education in multilingual South Africa*. University of Hawai'i at Hilo Hawai'i Community College.
- Joffe, V. L., Rixon, L., & Hulme, C. (2019). Improving storytelling and vocabulary in secondary school students with language disorder: A randomised controlled trial. *International Journal of Language & Communication Disorders*, 54(4), 656–672. <https://doi.org/10.1111/1460-6984.12471>
- Ivančević-Otanjac, M. (2016). Students with language learning disabilities and difficulties in a foreign language classroom. *Specijalna Educativa i Rehabilitacija*, 15(4), 461–474. <https://doi.org/10.5937/specedreh15-12071>
- Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data. *Medical Teacher*, 42(8), 846–854. <https://doi.org/10.1080/0142159X.2020.1755030>
- Khan, T. J., & Khan, N. (2016). Obstacles in learning English as a second language among intermediate students of districts Mianwali and Bhakkar. *Pakistan. Open Journal of Social Sciences*, 4, 154–162. <http://dx.doi.org/10.4236/jss.2016.42021>
- Khudsen, B. S. (2012). Do students' backgrounds in the language of instruction influence secondary school academic performance? *Journal of Education and Practice*, 6(30), 78–95.
- Krashen, S. (1988). *Second language acquisition and second language learning*. Prentice-Hall.
- Kretzer, M. M. (2019). South African teachers switch languages in class: Why policy should follow. [Blog Post]. <https://theconversation.com/south-african-teachers-switch-languages-in-class-why-policy-should-follow-122087>. Retrieved September 8, 2019.
- Krugel, R., & Fourie, E. (2014). Concerns for the language skills of South Africa learners and their teachers. *International Journal of Educational Sciences*, 7(1), 219–228. <https://doi.org/10.1080/09751122.2014.11890184>
- Krystal, A. (2015). *Teaching strategies for students with developmental speech delays in the mainstream classrooms*. Unpublished Thesis. Ontario Institute for Studies in Education of the University of Toronto.
- Lillywhite, K. M. (2011). Developing cognitive academic language proficiency (CALP) in diverse classrooms. *Journal of Interprofessional Care*, 35, 821–831.
- Lupogo, I. (2014). Language of instruction: A challenge for secondary schools and tertiary Institutions in implementing VET in Tanzania. *Journal of Educational Policy and Entrepreneurial Research*, 1(3), 26–30. <https://www.iiste.org/Journals/index.php/JEPER/index>
- Manyike, T. V., & Lemmer, E. M. (2014). Research in language education in South Africa: problems & prospects. *Mediterranean Journal of Social Sciences*, 5(8), 251–251. <https://doi.org/10.5901/MJSS.2014.v5n8p251>
- McGregor, K. K., Goffman, L., Horne, A. O. V., Hogan, T. P., & Finestack, L. H. (2020). *Developmental Language Disorder: Applications for advocacy, research and clinical service. Perspectives of the ASHA special interest groups*. Vol. 5. American Speech Language-Hearing Association.
- McGregor, K. K. (2020). How we fail children with Developmental Language Disorder. *Language. Speech and Hearing*, 5(7), 25–42.

- Miller, M. E. (2019). Sociability in children with Developmental Language Disorder. Unpublished Dissertation, Brigham Young University. <https://scholarsarchive.byu.edu/etd/7493>
- Ngulube, I. E. (2015). *Learn English son: A practical approach to the English language*. Ulamba Publishers.
- Nawaz, S., Umer, A., Tabasum, M., Zaman, M., Batool, A., & Aslam, S. (2015). Difficulties facing students of L1 in adopting L2. *European Journal of English Language, Linguistics and Literature*, 2(2), 1–6.
- O'Connor, J., & Gieger, M. (2009). Challenges facing primary school educators of English Second (or other) language learners in the Western Cape. *South African Journal of Education*, 29(2). <https://doi.org/10.15700/saje.v29n2a167>
- Ossai, C., & Uzoegwu, P. (2019). Improving the productive language skills of senior secondary school English language learners with expressive language disorders through inclusive education introduction. In *Conference proceedings: Inclusive education for national Developmental Faculty of Education*. University of Nigeria, Nsukka.
- Raju, N., & Joshith, V. P. (2018). Krashen's theory of second language acquisition: A practical approach for English language classrooms. *International Journal of Innovative Knowledge Concepts*, 6(12), 179–184.
- Rich-Mahadkar, S. (2015). *Strategic asset management for improved healthcare infrastructure planning in English NHS Trusts*. Doctoral dissertation, Loughborough University.
- Schütz, R. (2007). Stephen Krashen's theory of second language acquisition. *English Made in Brazil*, 2(2), 54–78.
- Starling, J., Munro, N., Togher, L., & Arciuli, J. (2011). Supporting secondary school students with language impairment. *Speech, Language and Hearing*, 13, 26–30
- Stubbs, S. (2008). *Inclusive education: Where there are few resources*. Atlas Alliance.
- Taderera, C., & Hall, H. (2017). Challenges faced by parents of children with learning disabilities *Africa Journal of disability*, 6(1), 83–98.
- Tribushinina, E., Dubinkina-Elgart, E. D., & Rabkina, N. (2020). Can children with DLD acquire a second language in a foreign-language classroom? Effects of age and cross language relationships. *Autism & Developmental Language Impairments*, 4, 1–12.
- Tuite, F. (2019). The impact of developmental language disorder on overall child development: *Speech and Language Therapists' Perspectives*, 6(32), 56–78.
- Vuzo, M. (2018). Towards achieving the sustainable development goals: Revisiting language of instruction in Tanzanian secondary schools. *International Journal of Science and Research*, 6(4), 1230–1241.
- Wissel, S. (2016). *Foundational skills to support reading for understanding in kindergarten through 3rd grade (NCEE 2016–4008)*. Washington, DC: USD Department of Education, Institute of Educational Sciences, National Center for Education: Evaluation and Regional Assistance.
- Ziegenfusz, S., Paynter J., Fluckiger, T., & Westerveld M. F. (2022). A systematic review of the academic achievement of primary and secondary school-aged students with developmental language disorder. *Autism and Developmental Language Impairments*, 7(1), 1–33.
- Zwane, S., & Malale, M. (2018). Investigating barriers teachers face in the implementation of inclusive education in high schools in Gege branch. *South African Journal of Disability*, 7, 1–12. <https://doi.org/104102/9jod.v7io.391>

Nettie N. Ndou, Segun Emanuel Adewoye

## **Herausforderungen bei dem Lernen in einer Zweitsprache unter südafrikanischen Schülern mit Sprachentwicklungsstörung**

### **Zusammenfassung**

Sprache spielt in der Bildung weltweit eine entscheidende Rolle für effektives Lehren und Lernen. Ziel der vorliegenden Studie war es, die Herausforderungen des Lernens in einer Zweitsprache bei Schülern der Sekundarstufe mit Sprachentwicklungsstörungen zu untersuchen. Die Studie zeigt auch Strategien, die von Experten zur Unterstützung der Lernenden eingesetzt werden, sowie die Einstellung der Lernenden zur erhaltenen Unterstützung. Die Studienteilnehmer waren Lernende ( $N = 12$ ), Lehrer ( $N = 5$ ), eine Sprachtherapeutin und ein Schulpsychologe. Es wurde ein qualitativer Forschungsansatz verwendet, bei dem eine Fallstudie als Forschungsdesign eingesetzt wurde. Zu den Quellen der Datengenerierung gehörten nicht-teilnehmende Beobachtungen, Interviews und Fokusgruppendifkussionen. Die Daten wurden mittels thematischer Analyse ausgewertet. Die Ergebnisse deuten darauf hin, dass das Lernen in einer Zweitsprache als Unterrichtsmedium für Lernende mit SES eine Herausforderung darstellt. Sie hatten Schwierigkeiten mit Verstehen, Lesen und Worterkennung von Lerninhalten. Die Lehrkräfte gaben an, dass sie Strategien wie Förderunterricht und Bereitstellung von mehr Lesestoff einsetzen, um den Wortschatz und die narrativen Fähigkeiten der Lernenden zu verbessern. Außerdem verwiesen sie die Lernenden an andere Experten zur weiteren Unterstützung. Auch die Beteiligung der Eltern an der Unterstützung von Lernenden wird als fördernd angesehen. Einer der Faktoren, der die Unterstützung von Lernenden einschränkt, ist allerdings ihre Einstellung zur erhaltenen Unterstützung. Diese kann auf das unzureichende Bewusstsein für SES zurückgeführt werden. Daher ist es notwendig, das Bewusstsein der Lernenden für SES zu schärfen. Ein ausführlicher Kurs in Sonderpädagogik in der Ausbildung von Sekundarschullehrern ist erforderlich, um die Lehrkräfte mit Strategien zur Verbesserung des integrativen Unterrichts vertraut zu machen. Die Förderung der lokalen Sprachen als Unterrichtsmedium muss auch in der Sekundarstufe einen hohen Stellenwert haben.

*Schlüsselwörter:* Sprachentwicklungsstörung, Zweitspracherwerb, Unterrichtsmedium, Lernherausforderungen, Herkunftssprache





**Zofia Chłopek**

University of Białystok, Poland

 <https://orcid.org/0000-0001-5324-8710>

**Jacek Pradela**

Silesian University of Technology, Poland

 <https://orcid.org/0000-0002-2079-4253>

## Parents' Perceptions of Children's Bi-/Multilingualism

### Abstract

Nowadays, many children all over the world are raised with more than one native language. The aim of the present study was to investigate how parents/caregivers evaluate the language competences of children growing up with two or more languages from birth or from very early in life. The results indicate that the language skills of young bi-/multilinguals are generally perceived to be comparable to those of their monolingual peers, though they are typically asymmetrical. Cross-linguistic influence is a common phenomenon which does not raise parents' concern. The outcomes also point to certain factors which may shape children's language competences.

*Keywords:* early bilingualism and multilingualism, native language acquisition, cross-linguistic influence

The number of transnational and multilingual families is rising continuously in many parts of the world (Li & Hua, 2015). Consequently, more and more children are growing up in contact with more than one native language (L1) from birth or from an early age. Nowadays, additive bi-/multilingualism is often considered an advantageous situation which enriches the linguistic repertoire of a child (e.g., De Houwer, 2005; Genesee & Nicoladis, 2009; Legacy et al., 2018; Meisel, 1990; see the next section). Nevertheless, some scholars indicate bi-/multilingual children's delay in their language development (e.g., Genesee & Nicoladis, 2009; Gósy, 2007; Hoff et al., 2012; Oller et al., 2007; Ordóñez, 2004; see the next section). The present study aims to gain further insight

into this issue by investigating parents' opinions on the language competences of children raised with more than one L1.

## Early Bi-/Multilingualism

Early bilingualism (bilingual first language acquisition) results from an early contact with two native languages (or dialects). Similarly, early multilingualism (multilingual first language acquisition) means that three or more mother tongues (or dialects) develop in early childhood (Chłopek, 2016, pp. 51–52). Scholars are not unanimous as to when precisely contact with each language should begin for a child to be considered an early (and not late) bi-/multilingual (Chłopek, 2016, pp. 52). In the present paper, we adopt a neurolinguistic perspective on bi-/multilingualism based on the distinction between implicit linguistic competence which is subserved by procedural memory and explicit metalinguistic knowledge which is subserved by declarative memory (Paradis, 2004, 2009; Ullman, 2001). Procedural memory is a subconscious type of memory, which is available from birth, but already around the age of five begins to gradually lose its plasticity. Native languages develop predominantly by means of this memory. The period from around the age of two to around the age of five is extremely important, since this is when procedural/implicit language competences must develop. Declarative memory, conversely, which is responsible for conscious learning processes, starts to develop around the age of two, but it is not until a child's seventh birthday that it functions relatively well. After this age it keeps developing, partly in response to school instruction in the native language, growing literacy skills, and rising metalinguistic awareness (Paradis, 2004, 2009). After puberty, declarative memory is enhanced and grows until early adulthood, while procedural memory becomes less available, even though performance in procedural memory may increase with practice (Ullman, 2001). Since mother tongues are acquired mainly by means of procedural memory, we assume that early bi-/multilinguals are people who have reached a fairly communicative proficiency level in two or more languages by the age of five. Any languages learned later in life are non-native languages.

Several research studies indicate that, as Meisel (1990, p. 17) puts it, “bilingual first language acquisition does not differ in substantial ways from monolingual development.” It has been claimed that children acquiring one L1 and those acquiring two L1s go through the same developmental stages and at a similar speed (e.g., De Houwer, 2005; Genesee & Nicoladis, 2009; Legacy et al., 2018). Researchers also believe that in the case of multilingual children,

stages of language development are in agreement with monolingual norms (Barnes, 2006; Gatto, 2004; Navracscics, 1998).

It has also been postulated that bi-/multilinguals are characterised by high levels of metalinguistic and metapragmatic awareness. Metalinguistic awareness comprises sensitivity to language as a system of signs, the ability to concentrate on linguistic forms independently of their meanings and to analyse and manipulate them consciously, as well as the capacity to notice cross-linguistic similarities (Jessner, 2006, p. 42), and is believed to positively affect language development (D'Angelo & Sorace, 2022; Herdina & Jessner, 2002; Jessner, 2006; Thomas, 1988). Studies conducted with children show young bilinguals' advantage over young monolinguals as regards metalinguistic awareness (Marinova-Todd, Zhao, & Bernhardt, 2010; Vásquez Carranza, 2009), as well as exceptionally high metalinguistic skills of young multilinguals (Hoffmann & Stavans, 2007; Navracscics, 1998). Metapragmatic awareness, in turn, manifests itself as sensitivity to one's interlocutors (their communicative preferences, needs, and characteristics) and the situational context in which communication takes place, as well as the ability to collaborate with others in order to achieve communicative success (Tomasello, 1999; Verschueren, 2000). High metapragmatic awareness may translate into high effectiveness of language acquisition (Safont Jordà, 2003). Researchers particularly emphasise young bi-/multilinguals' awareness of the fact that different people should be addressed in different languages (Genesee & Nicoladis, 2009; Montanari, 2009; Navracscics, 1998).

Nevertheless, children who grow up with more than one language may initially develop their language competences more slowly in comparison with their monolingual peers. In the first years of life, the mental lexicon of a bi-/multilingual child in each of their languages<sup>1</sup> is often narrower than that of a monolingual child of the same age (Genesee & Nicoladis, 2009; Hoff et al., 2012; Oller et al., 2007), and also other competences, such as grammar (Hoff et al., 2012), narrative skills (Ordóñez, 2004), or phonological processing (Gósy, 2007) may develop more slowly. This is a natural consequence of the fact that, compared to a monolingual child, a child raised bi-/multilingually uses each of their languages to a lesser extent—for fewer waking hours, in fewer communicative situations, and for fewer purposes (De Houwer, 2009; Grosjean, 1989; Hoff et al., 2012).

A common aspect of bi-/multilingualism is cross-linguistic influence (CLI), which is a psycholinguistic phenomenon consisting in one language present in the bi-/multilingual mind being affected by another one (Grosjean, 1989; Herdina & Jessner, 2002; van Dijk et al., 2022) and resulting from constant co-activation of language systems stored in the brain (Thierry & Wu, 2007). CLI

<sup>1</sup> Nevertheless, De Houwer (2022, p. 5) believes that normally developing, healthy bilingual children “learn to understand and speak at least one language to levels similar to monolingual peers at similar ages.”

leads to different kinds of language mixing, such as, for example, interlingual transfer (i.e., a “takeover” of the grammatical, phonetic, semantic etc. properties of a non-target language for the purposes of target language production), code switching (i.e., the use of non-target lexical items, phrases, or clauses), or borrowing (i.e., the adaptation of non-target lexical items in accordance with target language principles) (Chłopek, 2016, pp. 34–35; De Angelis, 2007; Sharwood Smith & Kellerman, 1986, p. 1).<sup>2</sup> CLI may, at times, hamper communication, particularly with monolinguals; however, with sufficient exposure to each language, it disappears over time (Döpke, 1998; Müller & Hulk, 2001); moreover, cross-language communicative practices such as translanguaging are even believed to speed up the language learning process (Li, 2018).

## The Present Study

### Aim of the Study and Research Questions

The present study aimed to gain more insight into the language abilities of young bi-/multilinguals, based on parents/caregivers’ perceptions of them. Thus, the following research questions were addressed: (1) How do parents/caregivers evaluate the language competences of children raised in early bi-/multilingualism? (2) How do they perceive cross-linguistic influence in their children’s speech? It was expected that the outcomes would depend on the children’s age, a younger age predicting lower linguistic competences in comparison with monolingual children, as well as more intensive CLI. As mentioned above, even though young bi-/multilinguals’ linguistic abilities are believed to be comparable to those of their monolingual peers, some language areas (such as vocabulary) may be less well-developed (Genesee & Nicoladis, 2009; Gósy, 2007; Hoff et al., 2012; Oller et al., 2007; Ordóñez, 2004). Apart from the main goal of finding answers to the research questions, it was hoped that the study would provide additional information about young bi-/multilinguals’ characteristics, for example, their metalinguistic awareness, as well as some external factors influencing their language competences, such as the attitudes of the environment towards bi-/multilingualism.

---

<sup>2</sup>It must be noted that the terminology used in this paper is by no means applied universally in different psycholinguistic publications.

## Method and Procedure

Data was collected by means of a questionnaire directed at parents and caregivers of bi-/multilingual children. On the one hand, a questionnaire study may be considered as biased by subjectivity, since parents may tend to evaluate their own children too favourably. On the other hand, such a study provides researchers with information based on data gathered in natural communicative settings by people who have permanent and intimate contact with their children and thus are able to recognise their verbal and nonverbal intentions much better than a researcher who usually concentrates on selected aspects of language in an artificial setting of a research study. For this reason, questionnaire studies may be treated as complementary to those conducted in controlled conditions.

The questionnaire consisted of closed- and open-ended questions. These aimed at obtaining information about the children's language biography and their environment (five questions), their language competences (four questions), and any other aspects of their language development (three questions and a space for further comments). The participants were asked to evaluate the receptive and productive competences in their children's languages by comparing them with those of monolingual children of the same age they had contact with. In particular, they were asked to place each of the children's languages on a scale from 1 (worse than monolingual peers) to 5 (better than monolingual peers). The simple questionnaire format, making use of a 5-point Likert scale, was chosen because it was obvious that people who are not engaged in linguistic research are able to make only general evaluations and comparisons of language competences. The respondents were also encouraged to leave comments on various aspects of their children's language abilities. In order to minimise the effect of the participants' subjectivity and to increase the reliability of their responses, clues as to which criteria should be applied, as well as additional explanations (e.g., of interlingual transfer), were provided.

The questionnaire was designed in three languages (English, German, and Polish), initially in Microsoft Word format and later also by means of Google Forms. The Word file and the link to the online questionnaire were sent out to parents of bi-/multilingual children by e-mail. Contact with the respondents was sought mainly via Facebook, primarily on FB groups for parents raising their children with two or more languages (*Uczę swoje dziecko angielskiego* and *Dwujęzyczność dziecięca*), and partly among the present authors' family members and acquaintances. The data was collected from July 2016 to March 2021. All fully completed questionnaires were accepted, provided that the family's economic status was not below average and the child had no language disorders or other communicative difficulties.

## Description of the Children

Fifty-one completed questionnaires were subjected to analysis, with each questionnaire providing information about one child. Most respondents (36) referred to one child only, there were five respondents with two children, and one respondent with five children. Even though the questionnaire was addressed at parents or caregivers, all the respondents were actually parents.

At the time of the questionnaire completion, the children were between 1 year and 1 month and 20 years and 6 months old (mean = 7.9 years, SD = 5.5); there were 31 females and 20 males in the described group. The economic status of the children's families was either average (32 children) or higher than average (19 children).<sup>3</sup> None of the children had a health deficit which might have influenced the results.

The children had from 2 to 5 native languages or dialects<sup>4</sup>; 38 of them had two L1s, 8—three L1s, 4—four L1s, and 1—five L1s. These were as follows: Polish ( $N = 50$ ), English ( $N = 22$ ), French ( $N = 14$ ), German ( $N = 10$ ), Spanish ( $N = 5$ ), Catalan ( $N = 3$ ), Turkish ( $N = 3$ ), Czech ( $N = 2$ ), Italian ( $N = 2$ ), Hungarian ( $N = 2$ ), Danish ( $N = 1$ ), Arabic ( $N = 1$ ), Mandarin Chinese ( $N = 1$ ), Kunming (a Mandarin Chinese dialect) ( $N = 1$ ), Castiglione (an Italian dialect) ( $N = 1$ ), Finnish ( $N = 1$ ), and Polish Sign Language ( $N = 1$ ). All the children have (had) regular contact with all their L1s in natural communicative situations before the age of five. Many were unbalanced bi-/multilinguals, that is, some of their languages were dominant; only 13 children were equally proficient in both/all of their L1s. In addition, 17 children had some knowledge of non-native languages: 11—one, 5—two, and 1—three non-L1s.

For the purposes of the quantitative analysis, the children were divided into four groups, in accordance with the age of each child at the moment of the questionnaire completion. The current knowledge about the functioning of the procedural and declarative memory, as well as the literacy skills development, were taken into consideration. Obviously, we realise that any such divisions are largely artificial, particularly as the age when certain abilities emerge varies significantly from child to child. The following groups were created:

- Group 1: children aged 1–5 (a period of intensive language development with the engagement of procedural memory),  $N = 20$ ;
- Group 2: children aged 5;1–7 (procedural memory is still efficient, but declarative memory develops; literacy skills gradually develop),  $N = 8$ ;
- Group 3: children aged 7;1–12 (procedural memory is still efficient, but declarative memory begins to function effectively; children acquire literacy skills at school);  $N = 11$ ;

<sup>3</sup> As indicated by the respondents on a 5-point Likert scale.

<sup>4</sup> For convenience purposes, we will use the term “language” in the remaining part of the paper in order to denote both a language and a dialect.

- Group 4: children aged 12;1 and older (around puberty declarative memory becomes stronger and procedural memory weaker),  $N = 12$ .

Obviously, the responses for the last group were partially retrospective.<sup>5</sup> Due to the small numbers of the children in each of the four groups, no statistical analyses were conducted. In the next section we describe the results and seek answers to the earlier posed research questions.

## Results and Discussion

Table 1 shows all the children's languages, in numbers (no data for reading and writing is included for the youngest groups, since most of these children have not developed literacy skills yet). As can be seen, the children's receptive skills (listening comprehension, reading comprehension) and productive skills (speaking, writing) in each L1 are believed to be mostly comparable to the respective skills of their peers growing up with one language; this trend can be observed in all age groups. Moreover, some L1s are believed to be at a higher level of proficiency than the one represented by monolingual children, though, as a closer analysis of the data reveals, very few children have mastered both/all of their L1s better than their monolingual peers as regards the four language skills. More importantly, not one child represents all skills in both/all of their L1s at a lower level than monolingual children of the same age. Some language deficiencies reported by the parents are as follows: firstly, both the listening and speaking skills in some languages of the youngest children seem to be delayed; secondly, the speaking abilities of the children aged 5;1–7 and older seem to be unstable in some of their languages; thirdly, the older children's writing ability is largely not fully mastered in all of their languages.

Thus, the data is in line with earlier research indicating an asymmetry of young bi-/multilinguals' language competences (Genesee & Nicoladis, 2009; Gósy, 2007; Hoff et al., 2012; Oller et al., 2007; Ordóñez, 2004). As the present study suggests, a delay in the development of some L1s may even occur after early childhood; in fact, literacy skills in some L1s may remain at a low level even in adolescence or early adulthood. Such differences in language competences are often caused by varying input and output in each language (De Houwer, 2009; Grosjean, 1989; Hoff et al., 2012). Indeed, a few respondents<sup>6</sup> indicate in their comments that certain factors, such as the recency of use of a given language and the possibility to communicate in it with different

---

<sup>5</sup> In spite of the retrospective character of some respondents' comments, the present tense is retained in the description of the data, in order to avoid complex structures such as, for example, "some of the children's L1s are reported to be or have been at a higher level of proficiency."

<sup>6</sup> If a similar comment was provided by 2–5 participants, the number of the respondents was not provided for clarity purposes. Comments made by only one person have not been mentioned.

**Table 1**

*The Parents' Evaluation of Their Children's Language Competences (the Numbers Refer to the Children's L1s, with Reading and Writing Skills Not Taken into Consideration for Groups 1 and 2)*

	<b>Group 1</b> <b>1–5</b> <i>N</i> = 20	<b>Group 2</b> <b>5;1–7</b> <i>N</i> = 8	<b>Group 3</b> <b>7;1–12</b> <i>N</i> = 11	<b>Group 4</b> <b>12;1 and older</b> <i>N</i> = 12	<b>Total</b> <i>N</i> = 51
<b>N of L1s</b>	<b>52 (100%)</b>	<b>18 (100%)</b>	<b>23 (100%)</b>	<b>28 (100%)</b>	<b>121</b>
<b>Reception—listening</b>					
worse than monolingual peers	6 (11.5%)	1 (5.6%)	3 (13.0%)	1 (3.6%)	<b>11</b>
a bit worse than monolingual peers	6 (11.5%)	3 (16.7%)	2 (8.7%)	4 (14.3%)	<b>15</b>
as good as monolingual peers	33 (63.5%)	12 (66.7%)	12 (52.2%)	20 (71.4%)	<b>77</b>
a bit better than monolingual peers	5 (9.6%)	0 (0%)	3 (13.0%)	0 (0.0%)	<b>8</b>
better than monolingual peers	2 (3.8%)	2 (11.1%)	3 (13.0%)	3 (10.7%)	<b>10</b>
<b>Reception—reading</b>					
worse than monolingual peers	--	--	2 (8.7%)	2 (7.1%)	<b>4</b>
a bit worse than monolingual peers	--	--	4 (17.4%)	4 (14.3%)	<b>10</b>
as good as monolingual peers	--	--	12 (52.2%)	19 (67.9%)	<b>31</b>
a bit better than monolingual peers	--	--	3 (13.0%)	0 (0.0%)	<b>3</b>
better than monolingual peers	--	--	2 (8.7%)	3 (10.7%)	<b>5</b>
<b>Production—speaking*</b>					
worse than monolingual peers	10 (19.2%)	2 (11.1%)	3 (13.0%)	2 (7.1%)	<b>17</b>
a bit worse than monolingual peers	7 (13.5%)	6 (33.3%)	3 (13.0%)	5 (17.9%)	<b>21</b>
as good as monolingual peers	20 (38.5%)	7 (38.9%)	10 (43.5%)	16 (57.1%)	<b>53</b>
a bit better than monolingual peers	4 (7.7)	1 (5.6%)	3 (13.0%)	2 (7.1%)	<b>10</b>
better than monolingual peers	5 (9.6%)	2 (11.1%)	4 (17.4%)	3 (10.7%)	<b>14</b>
<b>Production—writing**</b>					
worse than monolingual peers	--	--	6 (26.1%)	4 (14.3%)	<b>10</b>
a bit worse than monolingual peers	--	--	4 (17.4%)	8 (28.6%)	<b>12</b>
as good as monolingual peers	--	--	8 (34.8%)	10 (35.7)	<b>18</b>
a bit better than monolingual peers	--	--	2 (8.7%)	4 (14.3%)	<b>6</b>
better than monolingual peers	--	--	2 (8.7%)	2 (7.1%)	<b>4</b>

\* Two children from group one do not speak yet.

\*\* One child from group three does not write in one of her (three) languages.

people and in a variety of contexts, have a significant impact on its development (often at the expense of other languages), especially in early childhood, for example:

For the first five years [of our child's life] we were able to notice that after a 2–3 weeks' stay in a given country, when contact with the other language was very weak, her speaking skills [in this language] deteriorated, she tended to forget words, she stuttered, and wasn't able to formulate her thoughts. After moving to another place a reverse process occurred. From the age of six both languages have been so strong that forgetting hasn't taken place any more. [34/13;2/3]<sup>7</sup>

The teachers at the English school equip the child with knowledge of English in different areas (e.g., mathematics, geography, English, etc.). The child's home, the Polish school, and church fill in slightly different lexical fields, even though there are some common domains. [16/6;9/2]

Nineteen questionnaires include comments about various positive aspects of early bi-/multilingual development. Most of them concentrate on the ease with which children acquire their languages, even those which were introduced later than at birth, for example:

We were really impressed by the speed of the language acquisition of our son. He is learning very fast and he consciously uses many new words in both languages. He likes to repeat new words and phrases after us. [...] [10/1;3/2]

From the earliest years he has communicated with ease, he has a rich vocabulary, he is very talkative. [47/7;3/2]

[Her teachers] say that she uses much wider vocabulary than her peers and applies "adult" expressions. [7/8;10/3]

Such opinions are not surprising—research studies show that with each new language, acquisition becomes easier. This is because the already gained language competences become a source of positive transfer. Furthermore, meta-

---

<sup>7</sup>The respondents' comments have been edited and checked for spelling and grammar. The comments in Polish and German have been translated into English. All the information which might reveal the identity of the respondents or their children has been removed. At times, translations and explanations have been added in square brackets. At the end of each comment, the number of the questionnaire/child, the child's age at the moment of questionnaire completion, and the number of their L1s, have been provided.

linguistic and metapragmatic sensitivity, which grow as a consequence of multiple language acquisition, may facilitate communication and boost conscious language learning (for an overview, see Cenoz, 2003; Chłopek, 2011: ch. 3.10).

Only six questionnaires include comments on some negative aspects of an early acquisition of two or more languages. One of the participants describes a disruption in the development of all her daughter's languages after intensive contact with her fourth L1: after starting a Catalan-speaking nursery school at the age of two, the child's language competences became unstable and language mixing appeared, even though she had been able to separate her languages (Polish, Spanish, and English); this lasted until the child was about 3.5 years old, after which time she began to separate her languages better [23/3;8/4]. The other five children are reported to present some delay in one or more L1s, for example:

Our daughter began to speak single words [in her three languages] a bit later than her peers (she was about 3 months "late" compared to other children born in the same month or one month younger than her), but within a few weeks she made up for this delay; at the age of 1.5 she could speak Italian using full sentences, even though she always placed the pronoun "I" at the end of the sentence. Simple sentences in Polish appeared about 8–10 months later. [7/8;10/3]

A separate questionnaire item asked about cross-linguistic influence. In order to ensure clarity, no distinction between different cross-linguistic phenomena was made, and the terms "interlingual transfer" and "language mixing" were used interchangeably. As the responses show, more or less intensive CLI was observed in as many as 45 children's speech. Nevertheless, most participants admit that this is a rather rare phenomenon. It is also worth noting that not one respondent expresses any concern about it. Research studies indeed show that language mixing is not very intensive in the case of bi-/multilingual children, compared to foreign language learners (Barnes, 2006; De Houwer, 2005; Hoffmann & Stavans, 2007; Navracscics, 1998). The information provided by the participants also suggests that moderate interlingual transfer within the domains of lexis, morphology, and syntax may be the most common type of CLI in early childhood. Code switching may also appear early in life, though the tendency to switch codes seems to become stronger after early childhood. For example:

Yes, she's mixing her languages a lot, she inserts single Polish words into English sentences (*I've got mleko* ["milk"]). At times she inflects Polish words according to English rules (building plural noun forms: *komputerys* ["computers"], *butelkas* ["bottles"]); she adds Polish verb endings to English words (e.g., *It's mine, not to taczyć* ["don't touch"]). [...] Sometimes there

is a Polish word in an English sentence (*Where is my szczurek?* ["rat"]), other times an English word in a Polish sentence (*H. a story book chciałam* [H. – the child's name, "I wanted a storybook"]). [11/2;8/2]

Since her second birthday, when in a monolingual environment [...], she's been speaking both languages but with the grammar of the language everybody speaks at the moment. For example, when we are on holiday at her Turkish grandparents', she says in French *small a car* (the Turkish way) instead of *a small car* (the French word order); she places verbs at the end of the sentence and uses no articles, like in Turkish. [2/3;0/2]

Yes [he mixed languages], but only sporadically and during the first (three?) years of life. For example, he used to say *csuzdálnia*, combining the [Hungarian] word *csuszda* [...] and the [Polish] word *zjeżdżalnia* ["slide"]. At the very beginning he used to say *niem*, combining the Polish word *nie* and the Hungarian word *nem* ["no"]. [...] [9/8;3/2]

As some respondents' comments indicate, language mixing serves specific purposes. At times, it allows young bi-/multilinguals to compensate for gaps in knowledge and avoid communication problems. A typical situation is when a given lexical item is not yet present in the child's mental lexicon, but its equivalent in another language has already been acquired. Such asymmetry of language skills often stems from the fact that bi-/multilinguals typically use their languages in different situations, in different settings, and with different people (e.g., De Houwer, 2009; Grosjean, 1989). Consequently, some concepts are easier to express in one language, others in another. As some participants point out, young bi-/multilinguals may find it difficult to verbalise some ideas and describe certain situations in some of their languages, for example:

The child mixes his languages when he doesn't know the Polish equivalent [of an English word]. [35/10;5/2]

Language mixing takes place only in the case of a lack of specific vocabulary (e.g., while describing scout customs). [31/12;11/2]

He has always told me very little about what happened at kindergarten or school. He's talked a lot more when asked by his dad or his Hungarian grandma. I guess this has been easier for him, since these events took/take place in Hungarian. [9/8;3/2]

However, switching codes is not necessarily a sign of communication problems. Children may mix their languages in some situations and inhibit non-target languages in others—probably for various reasons, for example:

[...] Recently I've been wondering why he mixes his languages; he inserts a lot of French words [into his Polish utterances] when talking to me, even though I'm using only Polish. I thought that was the only way he could speak, but when my parents came to stay with us for two weeks, he spoke to them in full sentences using only Polish. [5/6;3/2]

It seems that a common reason for code switching is that this strategy makes communication more efficient. Thus, it may occur when a non-target language ensures economy or ease of expression, for example:

[Language] mixing occurs when the need arises to express some information very quickly; in such cases the child switches to Polish [i.e., her stronger language]. [29/11;3/2]

At the very beginning of speech development, the child selected from both languages such words which were easier to pronounce. E.g., he used the Danish word for “car,” because this word is easier to pronounce [than the Polish equivalent], but he used the Polish word for “dog.” [3/5;7/2]

Code switching after early childhood may also be explained by the fact that the rising metalinguistic awareness allows young bi-/multilinguals to consciously choose elements of their languages in order to achieve specific communicative needs. Purposeful, creative language mixing for jocular purposes can be treated as a sign of well-developed metalinguistic awareness (Johnsen, 2022). Such language use has been observed by a few participants, for example:

Mixing [was] frequent and subconscious until primary school, later [it became] rather conscious, for fun. [44/19;2/2]

Moreover, as children grow older, language mixing may become part of communication practices within a bi-/multilingual family. Whereas very young children do not question the practice of separating languages, which sometimes requires repeating the same utterance in two or more languages (e.g., the mother's L1 and the father's L1), older ones may find it unnatural, especially after discovering that there is actually no need to translate. As one respondent writes:

We try to avoid language mixing, but since the child matured and in the case of topics related to the Polish reality language mixing has been more frequent. [40/20;6/2]

Some participants' remarks suggest various factors which shape the frequency and intensity of CLI. One of them is recency of language use. Namely, a recently activated language may override other languages present in the mind, for example:

Subconscious mixing of Mandarin and Polish, most easily noticed after a change in language environment—i.e., soon after coming to Poland or right after returning to China. [27/2;2/4]

Another significant factor that shapes CLI seems to be language proficiency. A more fluent language usually dominates weaker ones, for example:

Polish remained [...] the dominant language and transfer from this language occurred, e.g., Polish prepositions were taken over into German (example: *Was gibt es auf Nachtisch?* ["What's for dessert?," translation of the Polish preposition *na* as German *auf*, instead of *zu*; the correct form is *zum Nachtisch*]), more infrequently: incorrect articles (gender), later also idioms. [44/19;2/2]

Recency of use and language proficiency are factors which often shape CLI (for an overview, see Chłopek, 2011, ch. 4.3). Moreover, fatigue, stress, or illness are indicated by some respondents as potential triggers of language mixing phenomena. Yet another factor mentioned by a few participants is the child's environment, that is, the language mixing habits of family members, for example:

[...] From about the age of three, vocabulary mixing has become infrequent. At present it does occur, but it is used in a conscious and purposeful way, as an accepted way of communicating between family members. [34/13;2/3]

In general, however, the respondents' children are either hardly ever ( $N = 21$ ) or never ( $N = 26$ ) confronted with mixed speech. This may be one reason for the moderate level of code mixing observed in their speech.

Several respondents' comments suggest that bi-/multilingual children are characterised by high metalinguistic awareness, that is, they are able to think about their languages, analyse them, and compare them. Such remarks can be found in 18 questionnaires, for example:

Sometimes he translates words from Polish into Spanish or from Spanish into Polish; at times he uses words in both languages, as if he knew that one object has two names (*nóż* and *cuchillo* [“knife”]). [10/1;3/2]

“Daddy says like this, but mommy says like that.” [2/3;0/2]

At times she compares the Polish and Spanish pronunciation. [17/13;3/4]

Moreover, eight parents of older children (from groups 3 and 4) indicate that they switch codes consciously/purposefully, in order to achieve specific communicative goals (all younger ones are reported to mix their languages subconsciously/automatically). Purposeful language mixing is considered to be a sign of high metalinguistic awareness (Jessner, 2006, pp. 84–119). The present results find support in the respective literature, which indicates particularly high levels of metalinguistic sensitivity of bi-/multilingual children (Hoffmann & Stavans, 2007; Marinova-Todd, Zhao, & Bernhardt, 2010; Navracscics, 1998; Vásquez Carranza, 2009). Metalinguistic skills are a useful tool in the acquisition of both native and foreign languages (Herdina & Jessner, 2002; Jessner, 2006)—which explains another finding, that is, the parents’ high opinions with regard to their children’s language learning aptitude.

A few participants ( $N = 8$ ) also mention high metapragmatic awareness of their children. In particular, they underline their understanding of the fact that people speak different languages, which must be attended to in communication, for example:

She immediately adjusts the language [of communication] to her interlocutor, she is never wrong in this respect. [27/2;2/4]

Our daughter remonstrates when someone addresses her in a language different than their mother tongue. She is used to such communication only. [...] She is aware of the fact that she can speak three languages and that not everyone can understand them. When she wants to communicate something to both her mum and her dad, she repeats the same sentence—in Polish and in French. [19/4;1/3]

Moreover, as the comments suggest, the younger children usually and the older ones almost always mix their languages only if their interlocutor knows both the target and the source language. This is another proof of bi-/multilingual children’s well-developed awareness of other people’s communicative needs, which coincides with previous study results (Genesee & Nicoladis, 2009; Montanari, 2009; Navracscics, 1998). High metapragmatic awareness undoubt-

edly makes it easier for a child to communicate with other people, which in turn ensures better social skills and facilitates language acquisition.

It turns out, however, that the choice—or avoidance—of a given language depends not only on the presence of particular people, but also on the situational context or specific objects or events this language is associated with. A few participants indicate that a certain place, occasion, or topic may trigger the use of a particular language or cause unwillingness to use another one, for example:

I've noticed that when we are in the kitchen he speaks Polish more often, I suppose he associates this language with food. [5/6;3/2]

She doesn't like to speak English outside the home. She won't speak to her kindergarten teacher, who is a native speaker of English [...]. During Skype conversations with her [Polish speaking] grandparents she answers their questions using single words. When playing on her own [...], she talks to her toys switching between both languages [i.e., Polish and English], sentence after sentence. [18/2;10/3]

As some respondents' comments suggest, the intensity and length of exposure to input, along with the proficiency level in each language, are crucial factors influencing language choice or avoidance. For example, a girl who had good receptive knowledge of three languages (Polish, Spanish, Catalan) at the age of two, but was later immersed in a Polish-speaking environment, refused to use the other languages and did not even respond in her father's L1 (Spanish) when addressed by him. However, after intensive contact with her Spanish family at the age of 2;6, the situation changed and at the moment of questionnaire completion she was using four languages (the fourth one was English), with Spanish as the dominant one [21/5;8/4]. Two other cases are a boy [30/4;7/2] and a girl [33/4;7/2] (siblings), whose French (their mother's L1) is weak compared to their Polish (the language of the environment): both of them use Polish words when speaking French even if their interlocutors do not understand Polish.

Additionally, the importance of the attitudes of the close environment for the development of children's bi-/multilingual competences and their language choices must be underlined. One of the questions included in the survey concerned the attitudes towards bi-/multilingualism in the child's immediate environment (i.e., family, school, neighbourhood). As many as 42 respondents state that the knowledge of more than one language is generally perceived as highly desirable and a reason to be proud. A large majority of the participants (except for four who did not leave a comment on this issue) believe that rais-

ing their children bi-/multilingually was the right choice. They treat early bi-/multilingualism as a natural solution, especially if different languages are spoken within the family and in the environment. Since parents' attitudes and engagement are crucial for the development of children's bi-/multilingualism (Paradowski & Michałowska, 2016), it may be argued that this was an important factor in this study as well.

## Answers to Research Questions

The first research question was: How do parents/caregivers of children raised in early bi-/multilingualism evaluate their language competences? As the obtained data suggests, a characteristic feature of children raised with two or more L1s is an asymmetry of competences. Young bi-/multilinguals typically have at least one language at a level which is comparable to or at times even higher than the level of young monolinguals' competences; nevertheless, the skills attained in other L1s may be much lower. Very young children's oral skills seem to be at a particularly low level, but even older bi-/multilinguals may encounter problems when speaking in some of their L1s. Also, the ability to write in more than one language may be a serious challenge even for a teenager or a young adult. Thus, as expected (section "Aim of the Study and Research Questions"), the child's age turns out to be a good predictor of their language competences.

The second research question was: How do they perceive cross-linguistic influence in their children's speech? As it turns out, cross-linguistic influence is frequently observed in young bi-/multilingual children's oral utterances; older bi-/multilinguals are usually able to separate their languages very well, though most probably this depends on the language practices of their closest environment. Moderate interlingual transfer within the domains of lexis, morphology, and syntax may be expected in early childhood. Code switching may also appear in the first years of life, but it seems to be more intensive after early childhood. CLI may result from the need to compensate for gaps in knowledge or to ensure economy of expression, or from growing metalinguistic awareness.

Both L1 competences and cross-linguistic influence are reported to depend on certain factors, such as recency of language use or input quality and quantity. The latter factor, along with the proficiency in each language, also has an impact on language choice or avoidance. The attitudes towards bi-/multilingualism and towards language mixing, represented by the child's closest environment, probably belong to crucial variables influencing their bi-/multilingualism.

In spite of low competences in some languages, as well as CLI, parents are impressed by their children's language learning aptitude, not only for their L1s, but also for foreign languages learned at school. Their attitudes towards early bi-/multilingualism are very positive—they consider it to be something desirable or even indispensable in the contemporary world. The results also point to young bi-/multilinguals' high levels of metalinguistic and metapragmatic awareness, abilities which can boost successful communication and language acquisition.

## Conclusion

In conclusion, the results of the present study indicate that parents of children raised with more than one language perceive their bi-/multilingualism in a positive way. They notice an asymmetry of young bi-/multilinguals' competences, which in the case of some language skills may persist until early adulthood, they also indicate moderate cross-linguistic influence in their language production, but they accept them as natural phenomena. They are impressed by their children's language learning aptitude, metalinguistic awareness, and metapragmatic awareness.

Finally, it should be stressed that the respondents were mostly parents sensitive to and interested in the issues related to early bi-/multilingualism, representing positive attitudes towards knowing more than one language, and willing to support their children's language development. Therefore, future research should concentrate on children who acquire more than one L1 with less assistance on the part of their parents.

## Acknowledgments

We would like to warmly thank all the parents who participated in the study. Most of them preferred to remain anonymous, but some agreed to have their names included in the acknowledgments; these are: Agnieszka Basta, Ewelina Brito-Cysarz, Agata Cox, Sylwia Cytrynska, Marzena Decowska, Małgorzata Denoue, Karolina Gaca, Kamila Kubiciel-Kotarska and her daughter Hanna Kotarska, Joanna Lustanski, Katarzyna Siegel-Hertz, Dorota Tomaszewska, Urszula Twarowska (in alphabetical order). We would also like to thank Susan Ross-Żółkiewicz for proofreading the article.

## References

- Barnes, J. D. (2006). *Early trilingualism: A focus on questions*. Multilingual Matters.
- Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, 7(1), 71–87.
- Chłopek, Z. (2011). *Nabywanie języków trzecich i kolejnych oraz wielojęzyczność: Aspekty psycholingwistyczne (i inne)*. Wydawnictwo Uniwersytetu Wrocławskiego.
- Chłopek, Z. (2016). *Early bilingualism and multilingualism: Parents' and caregivers' attitudes and observations*. Peter Lang. <https://doi.org/10.3726/978-3-653-06204-5>
- D'Angelo, F., & Sorace, A. (2022). The additive effect of metalinguistic awareness in third or additional language acquisition. *International Journal of Bilingual Education and Bilingualism*, 25(10), 3551–3567. <https://doi.org/10.1080/13670050.2022.2064710>
- De Houwer, A. (2005). Early bilingual acquisition: Focus on morphosyntax and the separate development hypothesis. In J. F. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 30–48). Oxford University Press.
- De Houwer, A. (2009). *Bilingual first language acquisition*. Multilingual Matters.
- De Houwer, A. (2022). The danger of bilingual-monolingual comparisons in applied linguistic research. *Applied Psycholinguistics, First View*, 1–15. <https://doi.org/10.1017/S014271642200042X>
- Döpke, S. (1998). Competing language structures: The acquisition of verb placement by bilingual German-English children. *Journal of Child Language*, 25(3), 555–584. <https://doi.org/10.1017/S0305000998003584>
- Gatto, D. (2004). Ein Kind–drei Sprachen: Bericht einer longitudinalen Fallstudie. In B. Hufeisen & N. Marx (Eds.), *Beim Schwedischlernen sind Englisch und Deutsch ganz hilfreich. Untersuchungen zum multiplen Sprachenlernen* (pp. 163–176). Peter Lang.
- Genesee, F., & Nicoladis, E. (2009). Bilingual first language acquisition. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 324–342). Wiley-Blackwell.
- Gósy, M. (2007). Speech perception processing in first and second language in bilinguals and L2 learners. In Z. Lengyel & J. Navracsics (Eds.), *Second language lexical processes: Applied linguistic and psycholinguistic perspectives* (pp. 39–59). Multilingual Matters.
- Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36 (1), 3–15.
- Herdina, P., & Jessner, U. (2002). *A dynamic model of multilingualism: Perspectives of change in psycholinguistics*. Multilingual Matters.
- Hoff, E., Core, C., Place, S., Rumiche, R., Senior, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, 39, 1–27. <https://doi.org/10.1017/S0305000910000759>
- Hoffmann, C., & Stavans, A. (2007). The evolution of trilingual codeswitching from infancy to school age: The shaping of trilingual competence through dynamic language dominance. *International Journal of Bilingualism*, 11(1), 55–72. <https://doi.org/10.1177/13670069070110010401>
- Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh University Press.
- Johnsen, R. V. (2022). “Quiero juksar en la julaftenito”—Playfulness and metalinguistic awareness in translingual family interactions. *Multilingua*, 41(1), 57–83. <https://doi.org/10.1515/multi-2019-0120>

- Legacy, J., Zesiger, P., Friend, M., & Poulin-Dubois, D. (2018). Vocabulary size and speed of word recognition in very young French-English bilinguals: A longitudinal study. *Bilingualism*, 21(1), 137–149. <https://doi.org/10.1017/S1366728916000833>
- Li, W. (2018). Translanguaging as a practical theory of language. *Applied Linguistics*, 39(1): 9–30. <https://doi.org/10.1093/applin/amx039>
- Li, W., & Hua, Z. (2015). Challenges of multilingualism to the family. In U. Jessner-Schmid & C. J. Kramsch (Eds.), *The multilingual challenge: Cross-disciplinary perspectives* (pp. 21–38). Walter de Gruyter Mouton. <https://doi.org/10.1515/9781614512165-002>
- Montanari, S. (2009). Pragmatic differentiation in early trilingual development. *Journal of Child Language*, 36, 597–627. <https://doi.org/10.1017/S0305000908009112>
- Müller, N., & Hulk, A. (2001). Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. *Bilingualism: Language and Cognition*, 4(1), 1–21. <https://doi.org/10.1017/S1366728901000116>
- Navracsics, J. (1998). *The acquisition of Hungarian by trilingual children*. Unpublished doctoral thesis, Janus Pannonius University of Pécs, Hungary.
- Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28 (2), 191–230. <https://doi.org/10.1017/S0142716407070117>
- Ordóñez, C. L. (2004). EFL and native Spanish in elite bilingual schools in Colombia: A first look at bilingual adolescent frog stories. *Bilingual Education and Bilingualism*, 7(5), 449–474. <https://doi.org/10.1080/13670050408667825>
- Paradis, M. (2004). *A neurolinguistic theory of bilingualism*. John Benjamins.
- Paradis, M. (2009). *Declarative and procedural determinants of second languages*. John Benjamins.
- Paradowski, M., & Michałowska, M. (2016). Establishing a bilingual home: Parents' perspective on the effectiveness of the adopted communication strategies. *Lingwistyka Stosowana*, 17(2), 43–65.
- Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, 76(27), 1–23.
- Safont Jordà, M. P. (2003). Metapragmatic awareness and pragmatic production of third language learners of English: A focus on request acts realizations. *International Journal of Bilingualism*, 7(1), 43–69.
- Sharwood Smith, M., & Kellerman, E. (1986). Crosslinguistic influence in second language acquisition: an introduction. In E. Kellerman & M. Sharwood Smith (Eds.), *Cross-linguistic influence in second language acquisition* (pp. 1–9). Pergamon.
- Thomas, J. (1988). The role played by metalinguistic awareness in second and third language learning. *Journal of Multilingual and Multicultural Development*, 9(3), 235–246.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
- Ullman, M. T. (2001). The neural basis of lexicon and grammar in first and second language: the declarative/procedural model. *Bilingualism: Language and Cognition*, 4(1), 105–122. <https://doi.org/10.1017/S1366728901000220>
- van Dijk, C., van Wonderen, E., Koutamanis, E., Kootstra, G. J., Dijkstra, T., & Unsworth, S. (2022). Cross-linguistic influence in simultaneous and early sequential bilingual children: A meta-analysis. *Journal of Child Language*, 49(5), 897–929. <https://doi.org/10.1017/S0305000921000337>
- Vásquez Carranza, L. M. (2009). Cognitive advantages of balanced bilingualism. *Revista Pensamiento Actual*, 9(12–13), 69–78.
- Verschueren, J. (2000). Notes on the role of metapragmatic awareness in language use. *Pragmatics*, 10(4), 439–456.

Zofia Chłopek, Jacek Pradela

## **Die Wahrnehmung der Eltern/Erzieher von der Zwei-/Mehrsprachigkeit der Kinder**

### Zusammenfassung

Viele Kinder werden heutzutage mit mehr als einer Muttersprache erzogen. Das Ziel dieser Studie war es zu erkennen, wie Eltern/Erzieher die sprachlichen Kompetenzen der Kinder einschätzen, die mit zwei oder mehr Sprachen von Geburt oder sehr früher Kindheit an aufwachsen. Die erhaltenen Resultate zeigen, dass die Sprachkompetenzen zwei-/mehrsprachiger Kinder ähnlich wie diese ihrer einsprachigen Gleichaltrigen bewertet werden, obwohl sie oft asymmetrisch sind. Zwischensprachliche Einflüsse sind ein verbreitetes Phänomen, welches keine Sorgen der Eltern/Erzieher hervorruft. Die Ergebnisse weisen auch auf bestimmte Faktoren hin, welche die Sprachkompetenzen der Kinder beeinflussen können.

*Schlüsselwörter:* frühe Zwei- und Mehrsprachigkeit, Erstspracherwerb, zwischensprachliche Einflüsse



**Fatih Ünal Bozdağ**

Osmaniye Korkut Ata University, Turkey

 <https://orcid.org/0000-0002-9959-4704>

## **Probabilistic Analysis of English Dative Constructions in Academic Writings of English EFL Learners**

### **Abstract**

Grammatical patterns in learner writings are among the most investigated topics in second/foreign language acquisition, gaining particular momentum thanks to corpus studies. English dative constructions are among those for which comprehensive literature is also available, consisting of different perspectives to explain the linguistic phenomenon on various theoretical grounds. However, except for rare instances, the foci of interest have constantly been on comparing learner data to native speaker data, particularly in terms of frequency of use in the second/foreign language learning environment. Different from other studies, the current study investigated English dative alternation in learner data with a probabilistic point of view, scrutinizing 27 learner corpora of learners with different L1s. Results showed that differences in learners' native languages had little added value to variations among learners. Moreover, a tendency similar to priming the verb 'give' in dative constructions was observed for the other variables in the construction.

*Keywords:* English dative constructions, learner corpora, interlanguage, Bayesian regression, first language influence

English ditransitive constructions have been extensively studied as grammatical patterns in the syntax-semantic interface. Initially, generative or conventional approaches proposed prepositional object constructions (henceforth PD) as the base form, over which transformation rules are applied to form double object constructions (henceforth DO). In this view, both variants refer to the same semantic space, and shifts in patterns are due to different argument realizations required by the verbal object. On the other hand, contemporary theories propose that two distinct forms address two separate yet related meanings. In constructional accounts, ditransitive constructions are formed around the

canonical sense, that is, “agent argument acts to cause a transfer of an object to recipients” (Goldberg, 1995, p. 32), and verb-specific variations are realized in polysemic extensions of related senses. Additionally, despite certain verbs’ saliency to prototypical sense, such a case is attributed not solely to verbal semantics or verb-oriented arguments but also certain constraints imposed by the constructional mapping of the “construal” (Langacker, 1987).

Early accounts of constructional analysis of English ditransitive emphasized the verb sensitivity approach, in which verbs included in different semantic classes are viewed as influential over the alternation preferences of English native speakers. Further proposing that certain verbs are more frequently preferred in one or the other variant, although the exact meaning is possible with any variant of choice with any verb as long as it permits ditransitive construction (Wasow, 2002; Arnold et al., 2003; Bresnan & Nikitina, 2008). Likewise, it is emphasized that the verb has the prominent position of influencing syntactic patterning of the constructions; for instance, the verb ‘give’ is commonly realized with animate recipients and concrete theme arguments in describing the transfer of possession sense.

In line with the constructional account of dative alternations, studies acknowledged that constructional choices vary due to factors including contextual variables, demographics of speakers, additional sociolinguistic factors, and metalinguistic influences with varying degrees of significance in different situations. For instance, Bresnan et al. (2004) observed additional constraints in spoken data, which are also imposed by characteristics of other arguments in a sentence, such as discourse accessibility, relative length, pronominally, definiteness and animacy. Similarly, Jensen et al. (2018) scrutinized the Spoken BNC2014 corpus investigating English dative alternation via multivariate models considering additional sociolinguistic factors. Results indicated that with a focus on sociolinguistics factors, “a graduate or postgraduate qualification lean towards a PD recipient, as does male gender, albeit with a smaller effect” (Jensen et al., 2018, p. 23). Also, Szmrecsanyi et al. (2017) investigated dative alternation among four varieties of spoken English, namely American English (AmE), British English (BrE), Canadian English (CanE), and New Zealand English (NZE), and observed how different groups of speakers would use the same constructions in varying forms.

The study of dative verbs and alternations is not uncommon in the second language acquisition/learning environment. Different approaches have attempted to explain the dative phenomenon from learners’ perspective, investigating learners’ preferences, acquisition order, awareness of verb sensitivity, or grammatical correctness of learner constructions. For instance, Le Compagnon (1984) reported French EFL learners’ common preference for prepositional constructions. Similarly, Mazurkewich (1984) also revealed a similar pattern for French and Inuit EFL learners, further proposing that prepositional datives are acquired

first. In contrast, Tanaka (1987) applied acceptability judgment and translation tasks to Japanese EFL learners, focusing on the verb 'give' in different contexts. The results highlighted that both variants were equally frequent in learner outputs. More recently, Chang (2004) argued that intermediate-level Chinese EFL learners strongly prefer prepositional dative, and discourse or information flow (given vs new) has a limited influence on learners' choices. Finally, Marefat (2005) explained that advanced and high-intermediate Persian EFL learners are aware of information flow in the discourse and sensitive to discourse factors resulting in native-like constructions of learners. Wolk et al. (2011) also explored the verb sensitivity awareness of French EFL learners with different proficiency levels. They observed that intermediate-level students favoured prepositional object dative, while advanced learners used two variants in almost similar frequencies, indicating advanced learners' better knowledge of current constructional rules and a higher awareness of verb sensitivity. Similar findings were also noted in other studies; for instance, beginner-level Russian EFL learners (De Cuypere et al., 2014) and German EFL learners (Führer, 2009) commonly preferred prepositional object datives more than double object ones. Lastly, Zeybek (2018) also acknowledged a higher prepositional object dative frequency than double object variants in Turkish EFL learners' data.

Other studies investigating the same phenomena utilized corpus analysis and multifactorial statistics. For instance, Callies and Szczesniak (2008) studied fifteen verbs frequently used in English dative constructions over German and Polish learners' subcorpora of ICLE, comparing results to British and American university students' subcorpora of LOCNESS. They noted that learners are aware of verb bias, information structure, and syntactic weight of post-verbal arguments, and regarding verb-construction pair choices, learners preferred certain verbs in either variant resembling similarities to native speakers. Similarly, Song and Sung (2017) analyzed the Korean EFL learners' corpus and compared the results to the native speaker corpus investigating fifteen frequent alternating verbs previously cited in Callies and Szczesniak (2008) and observed that Korean EFL learners favoured prepositional datives significantly more than native speakers. Finally, Babanoğlu (2007, 2011) scrutinized the Turkish subcorpus of ICLE, and on par with the studies mentioned above, she acknowledged a higher prepositional object dative frequency compared to native speakers of English.

The most comprehensive of all is Jäschke (2016), which examined 16 learner corpora, each representing one of 16 different native languages of learners and a corpus of English native speakers. The study reported that all learners included in the study successfully acquired both prepositional and double object variants and followed the pattern for post-verbal arguments as predicted by the harmonic alignment pattern (Bresnan & Ford, 2010). However, the length difference was the most prominent factor in predicting learner constructions,

and learners used more PD constructions in their writings than native speakers. Similarly, the current study also investigated the dative constructions in learner corpus, with a focus on influential lexical arguments in priming two different alternations, namely double object constructions and prepositional constructions.

As clearly inferred from the studies mentioned above, foci of interest have frequently been on the contextual and linguistic factors proposed to be influential on variant preferences. Besides, few studies in second/foreign language dative alternation explicitly addressed the role of the interaction between the learners' first language and the target language as a possible factor controlling construction preference. Firstly, Whong-Barr and Schwartz (2002) investigated datives in child first language acquisition and examined if L1 grammar properties would be transferred to L2 grammar. In the study, Japanese L1 speaker children accepted all dative structures as grammatically correct, although some were ungrammatical. Meanwhile, in the same study, children with Korean L1 could successfully distinguish grammatically correct structures from ungrammatical ones. Therefore, the authors suggested that Japanese children overgeneralized rules from their first language to English. In another study, Al-jadani (2018) also strongly emphasized L1 influence over L2 constructions, particularly in the context of Arabic EFL learners, and stated that Arabic learners failed to acquire the double object dative variant since the corresponding structure does not exist in Arabic.

Apart from those, many studies indirectly or implicitly addressed and discussed the role of first language as a likely influencing factor. For instance, Babanoğlu (2011) noted no significant difference among proficiency levels; however, there were patterns of over- and underuse compared to native speakers, which she concluded was due to L1 transfer. Kang (2011) and Song and Sung (2017) also concluded a similar discussion with an emphasis on cross-linguistic influences, as Korean does not have an equivalent structure for English ditransitive dative. Finally, Jäschke (2016) also concluded that learners whose first languages have a similar structural organization of dative constructions and corresponding structures tend to represent a more nativelikeness in English dative use.

However, one major problem with the studies cited above is that the analyses encompassed only a selected set of verbs proposed to be frequent in English dative constructions. Consequently, a common finding has been the frequent use of prepositional dative by non-native speakers compared to English native speakers. Nonetheless, given the hypothesis of lexical bias or verb saliency, it would have been more plausible if these studies had reported findings as the observed dominant use of prepositional object construction for selected verbs only. Moreover, since each verb and additional argument are likely to impose individual constraints for different variants, reliance on the analysis of only a set of selected verbs may lead to incorrect overgeneralization. Lastly, although

many studies, except for Jäschke (2016), included one group of students who share the same first language, findings were discussed mainly around the first language influence without any empirical data.

Therefore, following the framework provided by previous studies, yet with a different methodology, this study analyzed English datives with a top-down approach, from constructions to verbs and other variables. Accordingly, the dative phenomenon was investigated not through a set of selected verbs but via a constructional perspective. In doing so, learner corpora of students with different first languages were compared to evaluate the degree of variation among them and to reveal the common patterns of dative constructions. Therefore, the study aimed to investigate the probabilities of observing two English alternating dative constructions and compare these probabilities along with significance values across corpora. Additionally, it assessed the weight of influential factors, including pronominality and classification of agents, recipients, and themes as human vs. nonhuman, in addition to the length in characters of these items.

## Method

### Corpus Data

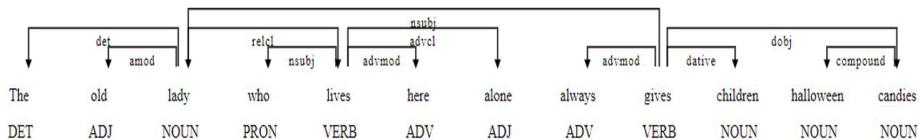
The study scrutinized ICLE version 3 (International Corpus of Learner Language) (Granger et al., 2020), a collection of essays written by learners from twenty-six different native languages on various topics. As of version 3, the corpus consists of “9,529 essays for a total number of 5,766,522 words distributed over twenty-five national subcorpora” (Granger et al., 2020, p. 33). Learners are young adults and university undergraduates with an average age range between 20 and 23 years old. Additionally, proficiency levels based on the Common European Framework (Council of Europe, 2001) vary from B1 as the lowest to C2 as the highest (Granger et al., 2020, p. 11). Approximately “61% of the sample essays were rated as advanced (C1 or C2), even reaching 100% in the case of Swedish, but can be as low as 5 or 10% in others” (Granger et al., 2020, p. 11). Therefore, there is no uniform distribution of proficiency levels across the corpus. Most texts included in the corpus are categorized as argumentative essays, while a small proportion corresponds to literary texts. Finally, there are thirty-two distinct native languages included in the corpus. However, due to their small size, Albanian, Bosnian, Arabic, Aromanian, Chinese-Mandarin, and the ones tagged as Others were not included in the analysis.

## Corpus Annotation and Dependency Parsing

Dative constructions were revealed through dependency parsing, which is the context-dependent analysis of sentence segmentation based on a sentence's grammatical and lexical structure. To specify, in dependency parsing, “the syntactic structure of a sentence is described solely in terms of the words (or lemmas) in a sentence and an associated set of directed binary grammatical relations that hold among the words’ and “the head-dependent relationship is made explicit by directly linking heads to the words that are immediately dependent on them, bypassing the need for constituent structures” (Jurafsky & Martin, 2021, p. 1). Figure 1 shows an example of dependency parsing output for a dative sentence.

**Figure 1**

### *Dependency Parsing for Sample Dative Sentence*



In Figure 1, the sentence was processed in two sets; clausal relations and modifier relations, as the headword (predicate or root), is a verb ‘give’ around which intra-sentence relationships are built, and syntactic roles are described, the word ‘children’ is the dative (DATIVE), and ‘candies’ is the direct object (DOBJ). Also, the subject of the sentence is segmented further, the noun ‘lady’ as the subject (NSUBJ) and who lives alone as the relative clause (RELCL), and the old as an adjectival modifier (AMOD) with a determiner (DET).<sup>1</sup>

The corpus was parsed, pos-tagged, and scrutinized in Python programming language via the natural language package Spacy version 3.2 (Honninger, Matthew and Montani, Ines and Van Landeghem, Sofie and Boyd, Adriane, 2020). The statistical dependency parsing model of the Spacy package has a claimed 100% tokenization and 94% dependency parsing accuracy.<sup>2</sup> For the study, two dependency patterns expressing prepositional object dative—(Subject—Predicate—Indirect Object—Preposition—Prepositional Object) and double object dative (Subject—Predicate—Direct Object—Indirect Object) constructions were defined, and corresponding sentences were extracted from the corpus via Python programming language. Next, extracted sentences were

<sup>1</sup> A detailed list of Dependency Labels can be found at [https://github.com/clir/clearnlp-guidelines/blob/master/md/specifications/dependency\\_labels.md](https://github.com/clir/clearnlp-guidelines/blob/master/md/specifications/dependency_labels.md)

<sup>2</sup> See <https://spacy.io/models/en> for more detailed accuracy evaluation.

segmented into the constituents, and subjects (Agent) and objects (Theme and Recipient) were tagged for pronominality and nonpronominality. Finally, lexical items were manually tagged for the humanness of referential entities as human (including collective nouns referring to humans) or nonhuman (animals and inanimate objects, including mass nouns). However, a note of caution is required for the variable consisting of referential entity classification, which is commonly termed as animacy in many studies. As indicated in many studies (Lakoff, 1987; Myhill, 1992; Yamamoto, 1999; Vihman, 2019), the concept of animacy is a linear scale extending from humans through animals to inanimate objects rather than a binary category of animate vs. inanimate. Also, in English, the animacy feature of animals is not clearly defined since some animals are conceptualized as more human-like (Leach, 1964) and therefore included in animates, while others are positioned on the other end of the animacy hierarchy scale and included among inanimates. Likewise, English codifies animals with the inanimate pronoun 'it' when singular. Regarding this, the referential status of subjects and objects was coded as human vs. nonhuman in the current study to avoid terminological confusion.

## Dataset and Statistical Analysis

Following Bresnan et al. (2007), the datasets were segmented into ten explanatory variables: native language, part-of-speech, and tags for referential entities of agent, theme, recipient, length of themes and recipients, and semantic classes of verbs. The numeric variables including length of theme and recipient were presented as log-scaled to adhere to normality distribution. Different from other studies, the properties of agents were also provided to the model as probable predictors. Another difference is that previous studies included semantic classes of verbs as categorical variable. However, as will be evident in the related section, the dataset was skewed regarding verb frequency, as the verb 'give' was dominantly prevalent across each corpus in both constructions. Consequently, this variable was excluded in the current study from the regression analysis, as it would have also mislead the results conditioned on the verb and their semantic classes. Finally, to investigate variation across learner corpora, the ICLE corpus was divided into subcorpora, and learners' native languages were assigned as random effect variables considering the ICLE corpus structure, where each learner language represents an individual subcorpus.

- Native Language, factor with 27 levels: Chinese, Chinese-Cantonese, Hungarian, Tswana, Korean, Swedish, German, Greek, Persian, Portuguese, Bulgarian, Japanese, Italian, Serbian, Polish, Czech, Lithuanian, Russian, Norwegian, Finnish, Macedonian, Turkish, Spanish, French, Dutch, Punjabi, Urdu, other.

- Agent, Theme, and Recipient part-of-speech tags, a categorical variable with two levels: Pronominal and Nonpronominal.
- Agent, Theme, Recipient animacy, a categorical variable with two levels: Human, Nonhuman.
- Recipient and Theme length: log-transformed, integer-valued variable.
- Alternations: a categorical variable with two levels, Double Object Dative (DO), Prepositional Object Dative (PD).

Statistical analysis consisted of descriptive analysis and Bayesian Regression as the probabilistic model. The term Bayesian indicates a different approach to the statistical inference of results; for instance, rather than significance estimates or point estimates, it provides the probabilities of predictors' effects over the response in the available data. Therefore, correlation and relationship among variables are expressed in uncertainty values termed as posterior distribution.

Bayesian regression was applied via "rstanarm" package (Goodrich et al., 2020) for the R programming language, and outputs were analyzed further with the 'BayestestR' (Makowski et al., 2019) package. MCMC (Markov Chain Monte-Carlo) was utilized with Bernoulli likelihood for categorical outcomes on normal (weekly informed) priors as the sampling method. Considering posterior distributions, incidences of centrality were reported as median, while Highest Density Interval (HDI) stated the uncertainty with posterior characterization, and Credible Interval (CI) reported the range of percentage of probable values. Finally, the probability of effect existence was expressed in Probability of Direction (pd), and the significance of the effect was estimated via Region of Practical Equivalence (ROPE). For the study, two different regression models were prepared; as the first one focuses on the differences among subcorpora regarding differences in first languages only, excluding additional variables. Meanwhile, the second model consists of influential factors as fixed effects and learners' first languages as random effects variable, that is, variation of interest conditioned on the additional factors. In other words, the first model realized the probabilities of alternating constructions given the first language differences only, while the second model represented the probabilities conditioned on the additional factors assessing the weights of each variable.

Two different regression models are as follows;

- 1) Binomial Regression with Native Language (fixed) predictor only as conditioned on categorical outcomes (DO vs PD)

**Logistics Regression Model for Native Language as Predictor only:**

Alternation  $\sim -^3 1 + \text{Native Language, family} = \text{binomial}(\text{link} = \text{'logit'}, \text{init}_r = 0.5, \text{QR} = \text{TRUE}, \text{iter} = 10000, \text{prior\_intercept} = \text{normal}(2, 0.5), \text{prior} = \text{normal}(0, 2.5, \text{autoscale} = \text{TRUE})$

---

<sup>3</sup>-1 indicates no intercept was defined.

- 2) Binomial regression with additional variables as fixed effects and Native Language as random effect predictors

**Logistic Regression Model for Additional Predictors with Native Language as Random Effect:** Alternation  $\sim$  Agent POS + Agent Humanness + Theme POS + Theme Humanness + Theme Length + Recipient POS + Recipient Humanness + Recipient Length + (1 | Native Language), <sup>4</sup>family = binomial(link = 'logit'), init\_r = 0.5, QR = TRUE, iter = 10000, prior\_intercept = normal(2, 0.5), prior = normal(0, 2.5, autoscale = TRUE)

## Statistical Analysis

### Overall Dative Construction Use across Corpora

The initial analysis included descriptive statistics for constructions, their frequencies across each corpus, the number of documents each construction was observed in, the number of verb types and tokens, and their distribution over constructions in each subcorpus.

Table 1 represents the raw frequencies of each construction and their percentages in each learner subcorpus. Initially, there were observed a total of 4011 dative constructions in the whole corpus, of which 70% ( $n = 2808$ ) were double object constructions, and 30% ( $n = 1203$ ) were prepositional object dative. Additionally, the table shows the number of unique documents in which constructions were used. For instance, considering the whole ICLE corpus, all double object constructions were dispersed over a total of 2054 (73%) individual learner essays. In other terms, the structure was used at least once in roughly seven out of ten texts consisting of double object construction. Therefore, comparing percentages of unique documents for both constructions, although the overall instances of prepositional variants were significantly lower than double object ones', the total number of unique documents with the former construction was higher ( $n = 994$ , 82.62%). The results indicate that the repeated use of prepositional constructions by the same students was lower compared to the figure for double object constructions.

---

<sup>4</sup>Learners' native languages were set as random effects variable.

**Table 1**  
*Distribution of Dative Constructions Across Learner Subcorpora*

Native Language	Alternation				NTotal
	Double Object		Prepositional Object		
	N1	N2	N1	N2	
Chinese-Cantonese	230 (61.82%)	172 (74.78%)	142 (38.17%)	117 (82.39%)	372
Japanese	134 (62.91%)	87 (64.92%)	79 (37.08%)	53 (67.08%)	213
Turkish	99 (47.82%)	74 (74.74%)	108 (52.17%)	80 (74.07%)	207
Russian	145 (76.31%)	103 (71.03%)	45 (23.68%)	41 (91.11%)	190
Swedish	147 (81.66%)	101 (68.70%)	33 (18.33%)	27 (81.81%)	180
Tswana	147 (81.66%)	111 (75.51%)	33 (18.33%)	33 (100%)	180
German	150 (85.22%)	111 (74%)	26 (14.77%)	26 (100%)	176
Greek	108 (62.79%)	80 (74.07%)	64 (37.20%)	57 (89.06%)	172
Serbian	132 (81.48%)	85 (64.39%)	30 (18.51%)	29 (96.66%)	162
Dutch	130 (81.76%)	81 (62.30%)	29 (18.23%)	24 (82.75%)	159
Norwegian	120 (76.43%)	93 (77.5%)	37 (23.56%)	30 (81.08%)	157
Punjabi	68 (49.27%)	46 (67.64%)	70 (50.72%)	46 (65.71%)	138
Korean	86 (64.66%)	58 (67.44%)	47 (35.33%)	40 (85.10%)	133
Macedonian	104 (79.38%)	78 (75%)	27 (20.61%)	22 (81.48%)	131
Czech	107 (82.30%)	73 (68.22%)	23 (17.69%)	19 (82.60%)	130
Italian	82 (63.07%)	68 (82.92%)	48 (36.92%)	41 (85.41%)	130
Persian	83 (64.34%)	69 (83.13%)	46 (35.65%)	36 (78.26%)	129
Bulgarian	93 (72.65%)	67 (72.04%)	35 (27.34%)	30 (85.71%)	128
Spanish	82 (67.21%)	58 (70.73%)	40 (32.78%)	36 (90%)	122
Polish	93 (76.85%)	71 (76.34%)	28 (23.14%)	24 (85.71%)	121
Portuguese	80 (66.66 %)	66 (82.5%)	40 (33.33%)	37 (92.5%)	120
French	84 (75%)	63 (75%)	28 (25%)	24 (85.71%)	112
Hungarian	75 (72.11%)	59 (78.66%)	29 (27.88%)	26 (89.65%)	104
Urdu	52 (54.73%)	34 (65.38%)	43 (45.26%)	28 (65.11%)	95
Chinese	65 (73.03%)	48 (73.84%)	24 (26.96%)	23 (95.83%)	89
Finnish	61 (75.30%)	54 (88.52%)	20 (24.69%)	19 (95%)	81
Lithuanian	51 (63.75%)	44 (86.27%)	29 (36.25%)	26 (89.65%)	80
<b>Total</b>	2808 (70.0%)	2054 (73.00%)	1203 (30%)	994 (82.62%)	4011

*Legend:* N1 shows the total raw frequency and the ratio of relevant variant to the sum of both constructions' frequency. N2 shows the number of unique documents and the ratio to the total frequency of the relevant construction.

Furthermore, regarding subcorpora individually, the highest rate of double object construction (85.22%) was observed in German learners' subcorpora when the figure was compared to that of prepositional construction in the same subcorpus (14.77%). On the other hand, the lowest rate of double object construction (47.82%) was observed in Turkish learners' subcorpus, followed by Punjabi learners' subcorpus (49.27%). Finally, the lowest number of unique documents of double object constructions was 62.30%, the lowest in Dutch learners' subcorpus, and the highest was 88.52% found in Finnish learners' subcorpus. To specify, the Finnish subcorpus had the lowest repetition rate by the same student as it had the highest ratio of unique documents for double object construction, while the Dutch learners' subcorpus presented the highest repetition rate. Meanwhile, in some cases, figures for prepositional construction dispersion were 100%, whereas the lowest ratio was observed in Urdu learners' subcorpus (65.11%), indicating a higher repetition rate by the same student.

**Table 2***Logistic Regression with Native Language as Predictor Only*

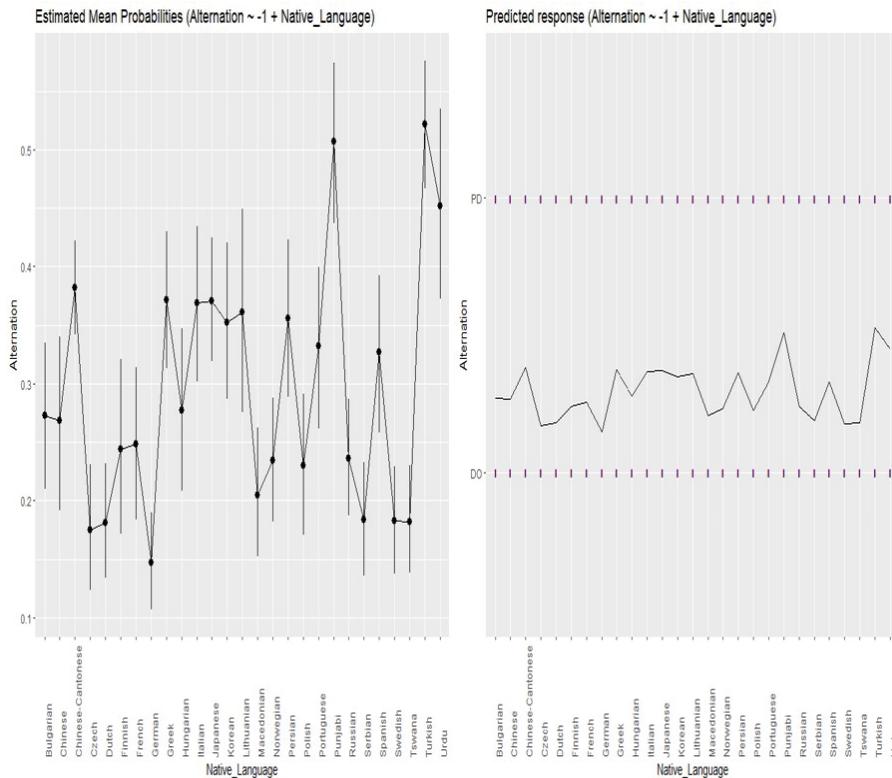
Parameters	Median	89% CI	pd	in ROPE %	BF
Bulgarian	-0.98	[-1.30, -0.68]	100%	0%	799.94
Chinese	-1.01	[-1.39, -0.62]	100%	0%	45.37
Chinese-Cantonese	-0.48	[-0.66, -0.31]	100%	0%	121.15
Czech	-1.55	[-1.93, -1.19]	100%	0%	> 1000
Dutch	-1.51	[-1.83, -1.18]	100%	0%	> 1000
Finnish	-1.13	[-1.55, -0.71]	100%	0%	178.95
French	-1.11	[-1.46, -0.76]	100%	0%	584.08
German	-1.76	[-2.11, -1.42]	100%	0%	> 1000
Greek	-0.52	[-0.77, -0.27]	100%	0%	4.96
Hungarian	-0.95	[-1.31, -0.62]	100%	0%	130.83
Italian	-0.54	[-0.83, -0.25]	100%	0%	1.23
Japanese	-0.53	[-0.75, -0.30]	100%	0%	19.82
Korean	-0.61	[-0.89, -0.32]	100%	0%	4.9
Lithuanian	-0.57	[-0.94, -0.19]	99%	3%	0.283
Macedonian	-1.36	[-1.70, -1.02]	100%	0%	> 1000
Norwegian	-1.18	[-1.49, -0.88]	100%	0%	> 1000
Persian	-0.59	[-0.88, -0.29]	100%	0%	2.84
Polish	-1.21	[-1.53, -0.85]	100%	0%	> 1000
Portuguese	-0.7	[-1.01, -0.39]	100%	0%	7.85
Punjabi	0.03	[-0.24, 0.30]	57%	74%	0.013
Russian	-1.17	[-1.45, -0.91]	100%	0%	> 1000
Serbian	-1.49	[-1.82, -1.17]	100%	0%	> 1000
Spanish	-0.72	[-1.04, -0.42]	100%	0%	9.27
Swedish	-1.5	[-1.80, -1.19]	100%	0%	> 1000
Tswana	-1.5	[-1.82, -1.20]	100%	0%	> 1000
Turkish	0.09	[-0.14, 0.31]	74%	75%	0.015
Urdu	-0.19	[-0.52, 0.14]	82%	47%	0.019

*Legend:* Credible Interval as HDI Highest Density Interval, pd = Probability of Direction, ROPE = Region of Practical Equivalence, BF = Bayes Factor

Table 2 shows the results for the second regression model, including alternation types as responses and learners' native languages as fixed-effect predictors. Note that double object construction was assigned as the reference level. Therefore, negative coefficients for most subcorpora indicate priming of DO constructions with a 100% probability of effect existence (pd). Also, based on %ROPE values, the effect was statistically significant. Nonetheless, in two subcorpora, Turkish and Punjabi, the direction was towards PD constructions. Considering these subcorpora, the effect had a roughly 57% of probability (Median = 0.03, CI[-0.24, 0.30]) in Punjabi and %74 of probability (Median = 0.09, CI[-0.14, 0.31]) in Turkish subcorpus. However, in both cases, the effect was not significant. Figure 2 shows the probabilities of alternations and predicted responses for each subcorpus.

**Figure 2**

*Estimated Means Probabilities and Predicted Alternations for each Subcorpus*



To sum up, estimated probabilities of observing DO or PD constructions show that only two subcorpora, Punjabi and Turkish, have probability values

higher than threshold 0.5, revealing a higher likelihood of observing PD. On the other hand, the probability of observing DO constructions was the highest in German learners' subcorpus, followed by Czech learners' subcorpus. Despite variations in probabilities of observing two different types of dative constructions, the effect of differences in learners' first languages was not statistically significant.

**Table 3**  
*The Most Frequent Verbs in both Constructions in ICLE*

Alternations					
Double Object			Prepositional Object		
Verb	Freq in DO	Ratio to Total Freq	Verb	Freq in PD	Ratio
<i>Give</i>	1485	52.9%	<i>Give</i>	633	52.6%
<i>Offer</i>	141	5.02%	<i>Bring</i>	106	8.81%
<i>Show</i>	134	4.77%	<i>Offer</i>	62	5.15%
<i>Bring</i>	123	4.38%	<i>Do</i>	54	4.49%
<i>Tell</i>	113	4.02%	<i>Pay</i>	52	4.32%
<i>Teach</i>	109	3.88%	<i>Send</i>	50	4.16%
<i>Provide</i>	95	3.38%	<i>Provide</i>	40	3.33%
<i>Cost</i>	73	2.60%	<i>Show</i>	29	2.41%
<i>Ask</i>	57	2.03%	<i>Sell</i>	21	1.75%
<i>Take</i>	48	1.71%	<i>Teach</i>	19	1.58%
<b>Total</b>	<b>2378/2808</b>	<b>84.68%</b>		<b>1066/1203</b>	<b>88.61%</b>

*Legend:* Freq in DO states the frequency of the verb in double object constructions, Freq in PD states frequency in the prepositional object. The ratio is the percentage of each verb to each construction.

Followingly, 147 different verbs in 4011 dative instances were observed. However, given the cut-off value of 10, this number dramatically dropped to 28 different verbs and even lower when verbs were organized into constructions. There were 126 different verbs in DO and 62 in PD constructions; however, the figures were 23 and 13 verbs, respectively, above the threshold value. Table 3 shows the top ten most frequent verbs per alternation and their percentages to the total verb frequency in each construction. Therefore, the most frequent top ten verbs accounted for approximately 85% ( $n = 2378$ ) of all double object constructions and 89% ( $n = 1066$ ) of all prepositional constructions, indicating a limited number of different verbs across the corpus. Likewise, the most common verb for both constructions was 'give', as it constituted more than 50% of each construction in the ICLE corpus individually and again almost 52% ( $n = 2118$ ) of all the constructions in the corpus.

Table 4 shows raw frequencies and percentages of part-of-speech tags along with their animacy status of components, namely agent, theme, and

recipient in ICLE. In both constructions, slightly more than half of the agents were nonpronouns, involving common nouns and proper nouns, while 66% were nonhuman entities in the double object variant and 51% were human entities in prepositional object ones. Followingly, nonpronoun themes were equally common in both constructions, and approximately 98% of all themes in each case were nonhuman. However, there was a stark contrast in recipients as pronouns were more common in double object variants and nonpronouns in prepositional object ones. Finally, almost 90% of recipients were human entities in double object constructions; the figure for prepositional ones was 64%.

**Table 4**

*The Distribution of Part-of-Speech and Animacy Tags in both Constructions*

Part-of-Speech	Alternation							
	Double Object				Prepositional Object			
	Pron.	NonPron	Ani.	Inani.	Pron.	NonPron	Ani.	Inani.
<b>Agent</b>	1240 (44.2%)	1568 (55.8%)	952 (33.9%)	1856 (66.1%)	501 (41.6%)	702 (58.4%)	622 (51.7%)	581 (48.3%)
<b>Theme</b>	117 (4.17%)	2691 (95.8%)	63 (2.24%)	2745 (97.8%)	70 (5.82%)	1133 (94.2%)	24 (1.99%)	1179 (98.0%)
<b>Recipient</b>	1987 (70.8%)	821 (29.2%)	2512 (89.5%)	296 (10.5%)	192 (16.0%)	1011 (84.0%)	767 (63.8%)	463 (36.2%)

*Legend:* *Pron.* states pronominals, including demonstrative and subject pronouns. *NonPron.* states nonpronouns, including common nouns and proper nouns. *Ani.* stands for animate while *Inani.* for inanimate.

## Estimating Weights of Influential Variables

The previous section detailed descriptive analysis of constructions with relative frequencies of verbs and other constituents. This section investigates the results of Binomial regression analysis to estimate individual factors' weights and the probabilities of observing varying alternations given the factors considered in the study.

**Table 5**  
*Summary of Posterior Distribution for Predictors in ICLE*

Parameter	Median	89% CI	pd	% in ROPE	BF
Intercept	-1.0	[-1.70, -0.31]	99.03%	2.67%	0.274
<b>Agent</b>					
PosPRON	-0.25	[-0.39, -0.09]	99.54%	24.59%	0.505
Animacy(Inanimate)	-0.98	[-1.14, -0.82]	100%	0%	> 1000
<b>Theme</b>					
PosPRON	1.11	[ 0.78, 1.42]	100.00%	0%	> 1000
Animacy(Inanimate)	0.79	[ 0.29, 1.24]	99.68%	1.71%	0.703
Length	-0.67	[-1.04, -0.29]	99.75%	1.71%	1.03
<b>Recipient</b>					
PosPRON	0	[-2.10, -1.70]	100%	0%	> 1000
Animacy(Inanimate)	0.74	[ 0.58, 0.91]	100%	0%	> 1000
Length	1.83	[ 1.40, 2.29]	100%	0%	> 1000

Legend: Credible Interval as HDI Highest Density Interval, pd = Probability of Direction, ROPE = Region of Practical Equivalence, BF = Bayesfactor

In Table 5, the summary for posterior distribution for the second regression model consisted of learners’ native language as random effects and other variables as fixed effects. Initially, given all the factors kept constant, with a probability of 99.03%, (Median = -1.0, 89% CI[-1.70, -0.31]), the double object construction (reference level) was 0.36 exp(-1.0) more likely.

**Figure 3**  
*Possibilities of Predicted Response across each Subcorpus*

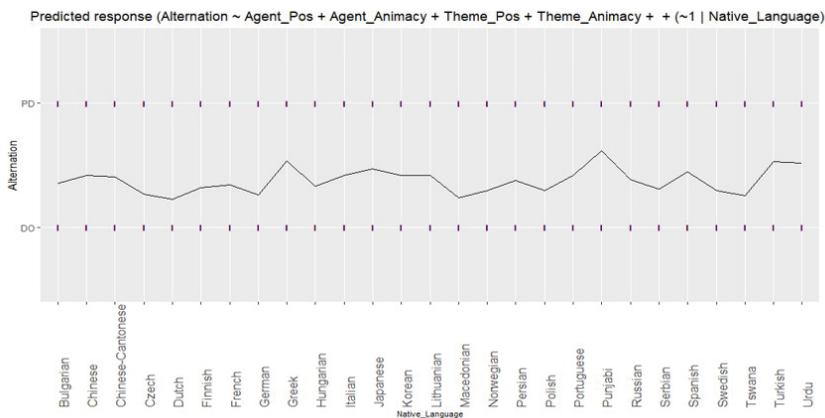


Figure 3 represents the probabilities of each alternation after conditioning the data for each reference level (human and pronominal agent, theme, recipient). It was noted that the probabilities of PD constructions were higher in Punjabi and Turkish subcorpora, followed by Greek learners' corpus. Also, in some cases, for instance, in Dutch, German, and Macedonian learners' corpora, probabilities favoured DO constructions. Nonetheless, as stated earlier in the interpretation of the previous regression model, the variation of probabilities of observing different constructions across all subcorpora was insignificant.

Considering influential items, pronominal agents with a probability of 99.54% (Median = -0.25, 89% CI[-0.39, -0.09] ) and pronominal recipients with a probability of 100% (Median = -1.90, 89% CI [-2.10, -1.70]) were in favour of DO constructions. However, the significance of the effect was undecided for the agent (24.59% in ROPE). Similarly, in frequency analysis, pronominal agents shared similar ratios in both constructions. Next, the effect was significant for the pronominal recipients (0% in ROPE), and they were  $0.14 \exp(-1.90)$  more likely to be primed in DO than PD constructions. On the other hand, the effect of pronominal themes has a probability of 100% being positive (Median = 1.11, 89% CI[ 0.78, 1.42] and can be considered significant (0% in ROPE). Therefore, pronominal themes were 3.04 times more likely to be observed in PD constructions, while nominal ones were more likely in DO. The following excerpts are from learner corpora,<sup>5</sup> and words/phrases in bold are examples of themes and recipients with more weight.

- 1) Several events have been showing **us this will to** cross frontiers. (Finnish – FIJY1049)
- 2) This Church provides **them the rest** in their heads (Czech – CZUN1018)
- 3) Firstly, First National Bank offers **you many oportunnities** and you can find it everywhere you and you can get it anytime you want. (Tswana – TSNO1399)
- 4) As a mother God has assigned **her a different role** to perform. (Punjabi – PACJ1004)
- 5)... the most common for the human kind is that we all usually don't give a **chance to someone...** (Serbian – SWUL4001)
- 6) He has a family and it is the right of his family that he also give **time to them**. (Urdu – PACJ1015)
- 7) ...but by doing so we hardly do **justice to him** or her as we simply cannot bring them back to life. (Polish – POPZ4035)
- 8) It's a sort of bliss for Cam and James: at that moment they feel that their father could ask **them anything** they would do it. (French – FRUC1030)

<sup>5</sup> The first line in brackets represents learner's first language and second line shows ICLE specific document id of the essay the sentence was extracted. Also note that errors in the learners' sentences were not corrected when sampling.

- 9) Second level, now we should tell **them useful English**. (Japanese JPKO2014)
- 10) However much applause these actors were given they owe **their success to William Shakespeare**. (Swedish – SWUL7050)
- 11) Later we have no sure that the judges will give **us justic**. (Punjabi – PAGF1066)
- 12) They can see that their work is providing **the society something good**, and this might relieve them from their feeling of alienation. (Norwegian – NOOS1045)

Considering the animacy status of variables, only nonhuman agents had more weight in DO variants, as the effect was significant (0% in ROPE) and had a probability of 100% of being negative (Median = -0.98, 89% CI[-1.14, -0.82]). Nonetheless, both nonhuman themes and nonhuman recipients significantly favoured PD constructions with a probability of 100% in both cases (Median = 0.79, 89% CI[0.29, 1.24]; 0.74, 89% CI[0.58, 0.91], respectively). The results indicate that nonhuman themes were 2.20 more likely in PD variants, whereas nonhuman recipients were 2.09 more probable in the same construction. The followings are sentences from learner corpora,<sup>6</sup> and words/phrases in bold present nonhuman agents.

- 13) **Modern means of transportation** give us the chance to travel widely, labour-saving devices make life easier and more enjoyable. (Bulgarian – BGSU1025)
- 14) Seriously, **television** gives us not only an amusement, but also good pieces of information. (Chinese – CNUK3052)
- 15) Secondly, **university degrees** not only bring people some important knowledge, but also can affect people's future. (Czech – CZPR3044)
- 16) Theoretically, **capitalism** gives people this freedom to choose. (Dutch – DBAN1004)
- 17) Because **it** will bring me money that will enable me to travel and do all the things I've dreamt of. (Finnish – FIHE1016)
- 18) **Our consumer society** offers us everyday more and more available products of all kinds: not only the necessities as food, clothes ... but also the superfluity as videos, freezers, micro-wave oven, televisions, etc... (French – FRUC2024)
- 19) **Cyber cafes** provide them a opportunity to use internet, they only pay little money to use computer in Cyber cafes. (German – GEAU2031)
- 20) First of all, **the competition** at the job market does not allow one the luxury of not having a mobile. (Hungarian – HUEL3075)

<sup>6</sup>The first line in brackets represents learner's first language and second line shows ICLE specific document id of the essay the sentence was extracted. Also note that errors in the learners' sentences were not corrected when sampling.

21) In conclusion, **the nature** sends us obvious warning messages by giving rise to these phenomena. (Japanese – JPTM1025)

Lastly, assessing the length of themes and recipients, results showed that while longer recipients were more likely in PD constructions, the opposite occurred for theme length: the longer the themes, the more likely they are to be observed in DO constructions. The effect of theme length has a probability of 99.75% of being negative (Median =  $-0.67$ , 89% CI[ $-1.04$ ,  $-0.29$ ]) and the effect of recipient length has a probability of 100% of being positive (Median =  $1.83$ , 89% CI[ $1.40$ ,  $2.29$ ]). Also, %ROPE values indicate that the effect can be considered significant in both cases.

## Discussion and Conclusion

This study investigated the English dative constructions in learner corpora with a focus on likely influential factors for patterning two different constructions, double object and prepositional one. After revealing the frequencies of constructions, verbs along with pronominality and humanness status of agents, themes and recipients, two different regression models were applied to observe probabilities of constructions. The first model focused on only the differences among learners' native languages regardless of additional factors, while the second model included additional factors and learners' first languages. Therefore, the study aimed to analyze if learners' first language or other factors have any weight in preference for any variants.

The results revealed that the differences in learners' first languages had no added value in priming a particular variant over the other. Despite varying probabilities of two different dative constructions, differences were not statistically credible, and only in two subcorpora, namely Punjabi and Turkish, the probability of observing PD construction was higher. A similar phenomenon was also observed in verb choices, as the verb 'give' was dominantly more frequent across the whole corpus in both variant types. Moreover, the top ten verbs commonly occurred in both variants consisted of 80% of all constructions, and of all constructions, more than 50% of each were structured with the verb 'give.' Lastly, DO type constructions were relatively more prone to be repeated by the same learner; however, PD type constructions had a wider dispersion range.

In general, dative alternation studies in learner languages are limited to investigating a set of selected verbs presumably frequent in dative construction in native speaker data. However, this approach might be incomplete and

misleading since studies tend to generalize findings for these selected verbs to the overall usage of the construction. Although the results pertain only to the findings of selected verbs, for instance (Babanoğlu, 2011; Jäschke, 2016; Song and Sung, 2017), prepositional object constructions were stated to be overused in learner data. Nonetheless, the correct interpretation indicates the overuse of prepositional object construction regarding only the verbs investigated in these studies, not the overuse of the construction in the given corpus.

In the current study, parameters were drawn for normally distributed data with normal priors and learner languages as random effect variables, that is, each subcorpus as a sampling source representing different groups for learners with a shared category of interest. The results showed that despite individual variations in probabilities across subcorpora, the difference was not significant, and most learners commonly primed the use of DO constructions.

Studies such as Szmrecsanyi et al. (2017) and Jensen (2018) have already revealed that constructional variation is influenced by several additional variables, including context, gender, and dialect. These studies invalidate the conception of a true preference only based on verb sensitivity or another single contextual factor, which is a case also investigated further by Chambaz and Desaguiller (2016). To specify, dative constructions are sensitive to the referential status of verbal complements (given vs. new). Also, considering syntactic complexity or end-weight preferences, the longer constituents are placed latter, while shorter ones are preferably positioned in front. Likewise, double object variants are realized with longer recipients; however, prepositional object variants consist of longer themes. Moreover, different organizations were observed for other variables; definite constituents are placed before indefinite ones, and animate items tend to occur before inanimate ones (Bresnan et al., 2007; Bresnan & Ford, 2010).

Given this fact, including the additional agent variable and excluding verb semantic class in the current study may have altered results, yet the findings were consistent with previous studies. In the current study, human recipients favoured DO constructions, considering the weight of individual variables over the construction type, pronominal, nonhuman agents, and pronominal. On the other hand, pronominal, nonhuman themes and nonhuman, nonpronominal recipients were more likely in PD constructions. Similarly, Bresnan et al. (2004) and Jäschke (2016) also reported that nonhuman recipients favoured PD constructions and Jäschke (2016) observed a higher probability of DO constructions given that recipients are pronominal and the likelihood of PD constructions with pronominal themes. Nevertheless, both studies excluded variable theme animacy from their conclusions as it was estimated to be a nonsignificant factor in regression. On the contrary, in this study, given the %ROPE value, that is, the percentages of probabilities inside the null range, parameter values of the factor fall entirely outside the rope, and theme referential feature was a sig-

nificant factor in dative construction preferences. Furthermore, similar to both studies mentioned above, a higher probability of prepositional object datives was observed with longer recipients, while longer themes were more probable in DO constructions in the current dataset. As a result, there was a common structural use of dative constructions across learner corpora. In other words, like the dominance of the verb ‘give’ in both types of constructions, priming of similar types of variables also exist in learner data regardless of first language differences.

The L1 influence is controversial in the SLA/EFL environment and even more complicated in interlanguage grammar studies. However, some studies mentioned before are biased towards native language interference as a source of variation or discrepancies in L2 outputs. They are biased in the sense that despite the lack of clear evidence of L1 interference or influence, learners’ inaccuracy, recursive use of similar patterns, or unauthentic uses were attributed to first language influence directly or indirectly. Indeed, the point of departure in one study was stated as “the influence of the disparity between English and Arabic” (Aljadani, 2018, p. 65), and in others, it was the existence or nonexistence of corresponding rules or categories in the target language (Montrul, 1997; Santoro, 2007; Cuypere et al., 2014; Szcześniak, 2017; Yang and Luo, 2017; Pongyoo, 2017; Zeybek, 2018). Nonetheless, Cuypere et al. (2014) and Pongyoo (2017) stated limited evidence favouring L1 interference and claimed that issues raised with dative alternation might be due to differences in learners’ proficiency levels. Other studies resulted in the influence of the first language on the target language, or more precisely, the transfer of L1 rules or knowledge to L2.

Moreover, Jäschke (2016) suggested that learners whose native languages share similarities with English in terms of dative construction were the ones “who most successfully master the native-like distribution of the two dative variants” (p. 166) and further claimed that “learners of those languages which have a dative alternation like English are very successful in acquiring the English-like distribution of the two competing dative constructions” (p. 167). One note of caution here is required as Jäschke (2016) realized verb senses as random effects; in other words, verb senses were provided as grouping variable where the variation of interest was centred. The data in that study was grouped based on the semantic classes of verbs used in the constructions, not based on learners’ first languages. Therefore, the observed difference in Jäschke’s (2016) study may have been due to the different verb choices of learners but not due to differences in learners’ first languages.

So far, these suggestions in the studies mentioned above are in favour of Bybee’s claim that “the acquisition of the L2 pattern in all its details is hindered by the L1 pattern” (2008, p. 232). Studies also suggested that the frequent use of particular forms in L2 may be due to first language influence (Foley & Flynn, 2018) or learners’ awareness or informed knowledge of L1

forms that would be appropriately transferred to L2 (Kellerman, 1989, as cited in Foley & Flynn, 2018). These views may explain the dominant use of the verb 'give' in dative constructions by learners in the current study; however, as acknowledged by Odlin (2018), it still bears more studies to observe the nature of the transfer and to what extent forms are transferred. Selective transfer of the lexis with high prototypicality effects is possible, yet it is not evident if it is the transfer of habitual uses, that is, transfer of meaning or the transfer of corresponding conceptual structures. The result of regression modelling in the current study supports Odlin's (2018) statement as differences in first languages were observed to have no significant effect on learners' priming of different types of dative constructions.

To conclude, in the case of this study, learners with different first languages presented similar probabilities of using the same dative variant, that is, double object construction. The result disagrees with arguments on first language influence suggested in other studies. Differences in learners' first language provided no difference at all in terms of English dative construction preferences and weight of influential factors. Regardless of the L1 variation, a common patterning was observed across learner data.

## **Limitations and Further Research**

As stated before, the construction preference may be affected due to various reasons. Similarly, the lack of genre differentiation in the ICLE corpus was an added limitation for comparison since certain genres may prime the use of particular forms as in academic writing. Also, the data lacked verb variation, which restricted the investigation of verb sensitivity in regression analysis, and predominant priming of the verb 'give' yielded skewed distribution for verb preferences, hindering a more detailed analysis of first language transfer. However, the skewness itself may indicate learners' inability to use other verbs in dative constructions due to additional reasons. For instance, it might also be indicative of external factors and learner strategies such as discourse influence, avoidance of using alternative verbs, or lack of knowledge or awareness of the dative structures as grammatical units. The result also suggests, a further analysis with alternative data sources and research methods to investigate the topic for the generalizability of the findings.

## References

- Aljadani, A. S. (2018). Second language acquisition of the English dative alternation by native speakers of Arabic. *Journal of Language and Cultural Education*, 6(1), 65–88. <https://doi.org/10.2478/jolace-2018-0005>
- Babanoğlu, M. P. (2011). İngilizce'deki Dolaylı Nesnenin İngilizce'yi İkinci Dil Olarak Edinen Türk Yetişkinler Tarafından Edinimi [The Acquisition of English Dative Alternation by Turkish Adult Learners of English]. *Dil Dergisi*, 153, 53–72. [https://doi.org/10.1501/Dilder\\_0000000156](https://doi.org/10.1501/Dilder_0000000156)
- Bresnan, J., Cueni, A., Nikitina, T., & Baayen, R. H. (2007). Predicting the dative alternation. In *Cognitive foundations of interpretation* (pp. 69–94). KNAW.
- Bresnan, J., & Hay, J. (2008). Gradient grammar: An effect of animacy on the syntax of give in New Zealand and American English. *Lingua*, 118(2), 245–259. <https://doi.org/10.1016/j.lingua.2007.02.007>
- Bybee, J. (2008). Usage-based grammar and second language acquisition. In *Handbook of cognitive linguistics and second language acquisition* (pp. 226–246). Routledge.
- Callies, M., & Szcześniak, K. (2008). Argument realization, information status and syntactic weight—A learner-corpus study of the dative alternation. *Fortgeschrittene Lernervarietäten. Korpuslinguistik und Zweitspracherwerbsforschung*, 165–187. <https://doi.org/10.1515/9783484970342.2.165>
- Chang, L. S. (2004). Discourse effects on EFL learners' production of dative constructions. *Journal of National Kaohsiung University of Applied Sciences*, 33, 145–170.
- Council of Europe. Council for Cultural Co-operation. Education Committee. Modern Languages Division. (2001). Common European framework of reference for languages: Learning, teaching, assessment. Cambridge University Press.
- De Cuyper, L., De Coster, E., & Baten, K. (2014). The acquisition of the English dative alternation by Russian foreign language learners. *PHRASIS (GENT): Studies in Language and Literature 2*, 187–212.
- Ellis, R. (1997). *SLA research and language teaching*. Oxford University Press.
- Foley, C., & Flynn, S. (2013). The role of the native language. *The Cambridge handbook of second language acquisition*, 97–113. <https://doi.org/10.1017/cbo9781139051729.008>
- Führer, M. (2009). *The dative alternation in the interlanguage of German learners of English*. GRIN Verlag.
- Gelman, A., Goodrich, B., Gabry, J., & Vehtari, A. (2019). R-squared for Bayesian regression models. *The American Statistician*, 73(3), 307–309. <https://doi.org/10.1080/00031305.2018.1549100>
- Goldberg, A. (1995). *Constructions: A construction grammar approach to argument structure*. University of Chicago Press.
- Goodrich, B., Gabry, J., Ali, I., & Brilleman, S. (2020). rstanarm: Bayesian applied regression modeling via Stan. R package version, 2.21.1 <https://mc-stan.org/rstanarm>
- Honnibal, M., Montani, I., Van Landeghem, S., & Boyd, A. (2020). spaCy: Industrial-strength Natural Language Processing in Python, Zenodo, 2020.
- James, C. 1980. *Contrastive analysis*. Longman.
- Jäschke, K., & Plag, I. (2016). The dative alternation in German-English interlanguage. *Studies in Second Language Acquisition*, 38(3), 485–521. <https://doi.org/10.1017/s0272263115000261>
- Jenset, G. B., McGillivray, B., & Rundell, M. (2018). The dative alternation revisited: Fresh insights from contemporary British spoken data. In *Corpus Approaches to Contemporary British*

- Speech* (pp. 185–208). Routledge. <https://doi.org/10.5334/johd.11>
- Juffs, A. (2000). An overview of the second language acquisition of links between verb semantics and morpho-syntax. *Second Language Acquisition and Linguistic Theory*, 170–179.
- Jurafsky, D., & Martin, J. H. (2022). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition* (3rd ed. draft). Retrieved November 12, 2022, from <https://web.stanford.edu/~jurafsky/slp3/>
- Kruschke, J. K. (2018). Rejecting or accepting parameter values in Bayesian estimation. *Advances in Methods and Practices in Psychological Science*, 1(2), 270–280. <https://doi.org/10.1177/2515245918771304>
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites* (vol. 1). Stanford University Press.
- Lakoff, G. (1987). The death of dead metaphor. *Metaphor and symbol*, 2(2), 143–147.
- Leach, E. R. (1964). Anthropological aspects of language: Animal categories and verbal abuse. *New Directions in the Study of Language*. In E. Lenneberg (Ed.) (pp. 23–63). MIT Press.
- Makowski, D., Ben-Shachar, M. S., & Lüdtke, D. (2019). bayestestR: Describing effects and their uncertainty, existence and significance within the Bayesian framework. *Journal of Open Source Software*, 4(40), 1541. <https://doi.org/10.21105/joss.01541>
- Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdtke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10, 2767. <https://doi.org/10.3389/fpsyg.2019.02767>
- Makowski, D., Ben-Shachar, M. S., Chen, S. A., & Lüdtke, D. (2019). Indices of effect existence and significance in the Bayesian framework. *Frontiers in Psychology*, 10, 2767.
- Marefat, H. (2005). The impact of information structure as a discourse factor on the acquisition of dative alternation by L2 learners. *Studia Linguistica*, 59(1), 66–82. <https://doi.org/10.1111/j.1467-9582.2005.00120.x>
- Mazurkewich, I. (1984). The acquisition of the dative alternation by second language learners and linguistic theory. *Language Learning*, 34(1), 91–108. <https://doi.org/10.1111/j.1467-1770.1984.tb00997.x>
- Melnick, R. J. (2017). *Consistency in variation: On the provenance of End-Weight* (Doctoral dissertation, Stanford University).
- Montrul, S. (1997). On the parallels between diachronic change and interlanguage grammars: The L2 acquisition of the Spanish dative case system. *Spanish Applied Linguistics*, 1, 87–113.
- Myhill, J., & Myhill. (1992). *Typological discourse analysis: Quantitative approaches to the study of linguistic function*. Blackwell.
- Odlin, T. (2008). Conceptual transfer and meaning extensions. In *Handbook of cognitive linguistics and second language acquisition* (pp. 316–350). Routledge. <https://doi.org/10.4324/9780203938560-22>
- Pongyoo, M. T. (2017). *The acquisition of dative constructions by Thai learners of English* (Doctoral dissertation, Thammasat University).
- Richards, J. C. (2015). *Error analysis: Perspectives on second language acquisition*. Routledge. <https://doi.org/10.2307/325758>
- Santoro, M. (2007). Second language acquisition of Italian accusative and dative clitics. *Second Language Research*, 23(1), 37–50. <https://doi.org/10.1177/0267658307071603>
- Song, E. J., & Sung, M. C. (2017). A corpus-based study of contextual factors influencing Korean EFL learners' dative alternation: Lexical verbs, syntactic weights, and information structures. *Australian Review of Applied Linguistics*, 40(1), 19–39. <https://doi.org/10.1075/aral.40.1.03son>

- Szczęśniak, K. (2017). Benefits of L1–L3 Similarities. The case of the dative case. *Theory and Practice of Second Language Acquisition*, 3(2), 87–110.
- Tanaka, S. (1987). The selective use of specific exemplars in second-language performance: The case of the dative alternation. *Language Learning*, 37(1), 63–88. <https://doi.org/10.1111/j.1467-1770.1968.tb01312.x>
- Touchie, H. Y. (1986). Second language learning errors: Their types, causes, and treatment. *JALT journal*, 8(1), 75–80.
- Vihman, V. A., & Nelson, D. (2019). Effects of animacy in grammar and cognition: Introduction to special issue. *Open Linguistics*, 5(1), 260–267.
- Yamamoto, M. (1999). *Animacy and reference*. [Studies in Language Companion Series, Vol. 46]. John Benjamins Publishing Company.
- Yang, C., & Luo, H. (2017). The acquisition of Mandarin dative constructions by L2 and heritage learners. *Chinese as a Second Language Research*, 6(2), 251–275. <https://doi.org/10.1515/caslar-2017-0011>
- Zeybek, G. The acquisition of dative alternation and markedness in second language learning. *Turkish Online Journal of English Language Teaching*, 3(2), 26–41.

Fatih Ünal Bozdağ

## **Probabilistische Analyse der englischen Dativkonstruktionen in akademischen Texten von EFL-Lernern**

### **Zusammenfassung**

Grammatikalische Muster in Texten der Lernenden gehören zu den am meisten untersuchten Themen im Bereich des Zweit-/Fremdspracherwerbs und erfahren dank Korpusstudien einen besonderen Aufschwung. Auch für englische Dativkonstruktionen gibt es eine umfangreiche Literatur, die das sprachliche Phänomen aus verschiedenen Perspektiven und mit unterschiedlichen theoretischen Grundlagen erklärt. Bis auf wenige Ausnahmen lag der Schwerpunkt des Interesses jedoch stets auf dem Vergleich von Lernerdaten mit Daten von Muttersprachlern, insbesondere im Hinblick auf die Häufigkeit der Verwendung in der Lernumgebung der Zweit-/Fremdsprache. Im Unterschied zu anderen Studien wurde in der vorliegenden Studie die englische Dativalternation in Lernerdaten unter einem probabilistischen Gesichtspunkt untersucht, wobei 27 Lernerkorpora von Lernenden mit unterschiedlichen L1 einer Analyse unterzogen wurden. Die Ergebnisse zeigten, dass die Unterschiede in den Muttersprachen der Lernenden einen geringen Zusatznutzen für die Variationen zwischen den Lernenden hatten. Darüber hinaus wurde eine ähnliche Tendenz wie bei dem Priming des Verbs „geben“ in Dativkonstruktionen für die anderen Variablen in der Konstruktion beobachtet.

*Schlüsselwörter:* englische Dativkonstruktionen, Lernerkorpora, Interimssprache, Bayes'sche Regression, Einfluss der Erstsprache



**Marta Sánchez-Saus Laserna**  
University of Cádiz, Spain

 <https://orcid.org/0000-0002-1123-2873>

## **Lexical Availability and Foreign Language Teaching: Main Contributions of a Growing Field\***

### Abstract

Teaching vocabulary presents a significant challenge in foreign language instruction. This article discusses the advantages that the results obtained in studies on lexical availability offer for foreign language teaching. The methodology of lexical availability, which involves extracting vocabulary that is closely associated with specific lexical categories, was first developed in France in the 1950s, and since then, it has been predominantly used for teaching Spanish as a first language. More recently, in the past 15 years, it has also been employed for teaching Spanish as a foreign language. This article examines the origins and methodology of lexical availability, with a particular focus on two significant applications related to teaching Spanish as a foreign language. From a cognitive standpoint, the article discusses studies analysing the configuration of the mental lexicon using available vocabulary data. From a didactic perspective, the article explores how lexical availability can be used to improve the selection of vocabulary in teaching materials for Spanish as a foreign language.

*Keywords:* lexical availability, vocabulary teaching, foreign language teaching, Spanish as a foreign language, cognitive linguistics

One of the primary challenges in language teaching is deciding which vocabulary should be taught at each level of learning. Lexical availability is a complementary approach to traditional frequency-based research, as it uncovers the words that are most readily accessible to speakers in a particular communicative context, rather than the most commonly used or frequent ones.

---

<sup>1</sup> This research was funded by the Andalusian ERDF Operational Program (project “Observación del Pulso Social en Andalucía a Través del Análisis Léxico (Pulso Andaluz). Primera Fase” UMA20-FEDERJA-013) and the Spanish Ministry of Science and Innovation (project “TERLEX-WEB. Aplicaciones de la lingüística digital al ámbito de la terminología: la creación de un léxico relacional bilingüe de usos terminológicos de la semántica léxica” PID2022-139201OB-I00).

Therefore, by combining both available and frequent lexicons, we can establish the essential vocabulary of a language (Ávila Muñoz, 2016, 2017; Tomé Cornejo, 2015). The initial aim of lexical availability was to complement frequency-based vocabulary research, but it has now evolved into a substantial field of study with an extensive literature and practical relevance to enhance vocabulary teaching in both native and foreign language instruction, especially for Spanish. It also holds significance for cognitive studies that explore the mental lexicon's structure in both native and non-native speakers.

Studies on lexical availability in languages other than Spanish are still not as popular as they are in this language, although in recent years these studies have notably developed for English as foreign language.<sup>1</sup> It is also worth noting that researchers from Hispanic countries have largely contributed to the development of this field, despite the fact that its origins are in France and focused on French, and that it can be applied to any language.

In this context, given the considerable developments in lexical availability research in recent years, this article seeks to provide an up-to-date overview of the field's most important findings. Our focus will be on research that addresses issues particularly relevant to foreign language teaching, such as how available lexicon is organised internally, and how such studies can greatly assist in selecting appropriate vocabulary for instructional materials. Thus, our goal is to inform researchers interested in language teaching beyond Spanish or English about the benefits and prospects of these studies. The article will provide an overview of the terminology, key authors, and major milestones in the field.

## Lexical Availability

### Origins

The concept of lexical availability was first introduced in the 1950s (Michéa, 1953), but it was not until the 1980s that research was standardised with a common approach and measurement system. The concern to determine what the fundamental lexicon of a language is—that is, which words are most commonly used in a given language and to establish which vocabulary should be taught to foreign language students—is an old one. Just as adult speakers of a language

---

<sup>1</sup> Some of the most recent and significant publications on lexical availability applied to English as a Foreign Language are, among others: Ferreira & Echeverría, 2010; Canga Alonso, 2017; Martínez Adrián & Gallardo del Puerto, 2017; Ferreira, Garrido, & Guerra, 2019; Jiménez Catalán & Fernández Fontecha, 2019; Jiménez Catalán & Canga Alonso, 2019; as well as the monograph edited by Jiménez Catalán in 2014, dedicated to applications of lexical availability to teaching Spanish and English as foreign languages.

cannot possibly know *all* the vocabulary of their own language, a restricted sample must also be selected when teaching vocabulary in a foreign language. As Grève and Van Passel (1971, p. 105) point out,

in view of the lexicological limitation [...] both in the field of literature and in that of everyday language, a similar limitation is imposed in the teaching of a language, both of the mother tongue and of foreign languages (author's own translation).

Since the end of the nineteenth century—and even before that, though mainly during that period—lexicographical work has focused on collecting the basic vocabulary of a community. Such work, however, was not always intended for language teaching purposes and many publications were criticised for referring only to written texts. Still, some of them should be considered ground-breaking work in their field.<sup>2</sup> The general belief that the most frequent lexical elements found in a text corpus are the most useful for building a language's basic vocabulary had to be reconsidered. Around 1950, a group of French researchers set up a study to provide students of French as a foreign language a solid basic knowledge of French, which led to the development of the *Français Fondamental* (Gougenheim et al., 1964). On that occasion, the selection of words was based on frequency rates, that is, number of times a given word was repeated in a text. It was soon noticed that this method labelled some widely used French words as non-existent or very rare. Well-known words did not appear on the list simply because they were not related to the topics and circumstances of conversation included in those studies (Michéa, 1953).

The absence of everyday words (or frequent lexica) such as *fork*, *subway* or *tooth* soon raised suspicions about the flaws of using word frequency as criteria for selecting the most commonly used vocabulary items of a given language. As a result, experts pursued a criterion known as vocabulary availability, which involved devising a new word-counting formula that went beyond mere frequency and incorporated a simple- and complex-dispersion rate. This formula proved more effective at determining the rate of word usage than the frequency-based methodology, as it took into account the *distribution of frequencies* in the text and avoided potential oversights.

The type of relationship observed between the resulting “basic lexicon”—or “thematic lexicon”—and the communicative situations they were mentioned in, and the subsequent critical review of previous word-selecting criteria, led to the development of a new research tool coined as *lexical availability* and the concept of *available lexicon*.

<sup>2</sup>Carcedo (2000) carries out a detailed historiographical review on the subject, based mainly on Gougenheim et al. (1964). In Sánchez-Saus (2019, pp. 15–18), we also review the basic vocabularies based on the lexical frequency previous to the *Français Élémentaire* and the *Français Fondamental*.

Michéa (1953, p. 342) was the first author to use this term, defining it as follows:

In the presence of a given situation, the first words that come to mind are those that are specifically related to that situation. An “available word” is a word, which, though not particularly frequent, is always ready for use and comes immediately to mind when needed. (author’s own translation)

The methodology used to extract available vocabulary was stimuli-based. Stimuli represented thematic areas or so-called *centres of interest* (which, from a cognitive perspective, can be considered as categories, of different types; this matter will be addressed in subsection Cognitive Approach). The term *centre of interest*, which was adopted from pedagogical studies, was first used by Michéa (1950, p. 189) despite being traditionally used by lexicography for the classification of certain lexicographic works (Sánchez-Saus, 2019). The authors of *Français Fondamental* selected 16 centres of interest, which they believed contain the most basic and universal vocabulary. However, their choice is not considered universal by later works, and the authors do not provide an explanation for their selection. The basic lexical units included in their studies are:

1. body parts;
  2. clothing (men’s and women’s);
  3. the house (except for the furniture);
  4. house furniture;
  5. food and beverages at meals (all daily meals);
  6. objects placed on the table and used at daily meals;
  7. the kitchen, its furniture and utensils therein;
  8. the school, its furniture and the school equipment;
  9. heating and lighting;
  10. the town;
  11. the village;
  12. means of transportation;
  13. field and garden work;
  14. animals;
  15. games and entertainment;
  16. professions (all kinds of professions, not those referring to a single job).
- (Gougenheim et al., 1964, pp. 152–153, author’s own translation)

The problem of how to select the centres of interest that should be used has been widely debated in recent years, initially raised by López Morales (1999, p. 33). A summary of the heterogeneity of centres of interest in related literature can be found in the work of Samper Padilla, Bellón Fernández, and Samper Hernández (2003). For more up-to-date information, refer to Fernández Smith

et al. (2012), Paredes García (2014), González Fernández (2014), and Tomé Cordero (2015). Moreover, a summary and a proposed set of centres of interest for learners of Spanish as a foreign language can be found in Sánchez-Saus (2016).

While frequency lists usually have a balance of grammatical terms, such as articles, prepositions, and conjunctions, along with verbs, adjectives, and nouns, the available lexicon mostly comprises nouns, with a few exceptions for verbs, adjectives, and adverbs. The reason for the prevalence of nouns in the available lexicon is mainly because of the methodology used to recall those words and the selection of centres of interest. As the fundamental lexicon of a language is comprised of both the frequent and available vocabulary, the two types of studies complement each other to form the fundamental lexicon of a language.

### **Method of Extraction of Available Lexicon**

Lexical availability can be defined as the flow of vocabulary used in a specific communicative situation and which, as noted above, complements the frequent lexicon to determine the basic lexicon of a language. In order to collect data from this lexicon, experts use associative tests, that is, a test where informants identify all lexical units derived from the proposed centres of interest and they note down all the words related to them.<sup>3</sup> This results in lists of vocabulary that outline the real lexicon a speaker uses for each centre of interest (Carcedo González, 1998).

The formula used to calculate the availability index of each word is that developed by the professors of the Universidad Nacional Autónoma de México, López Chávez and Strassburger (cf. López Chávez & Strassburger, 1987, 1991). From a given vocabulary list, their formulation is based on the analysis of a matrix of vectors in which the following elements are involved:

- (a) the absolute frequency of a unit,
- (b) the absolute frequency of that unit regarding its position on the list,
- (c) the number of participants in the survey,
- (d) the amount of positions the unit holds in the list; and
- (e) the unit positions, that is, the unit's availability degree, which gets lower as the word appears in lower positions.

The formula for the lexical availability index of a word is as follows:

---

<sup>3</sup> Some studies have conducted the tests orally. This was particularly the case when the informants were children. Gómez-Devís and Herranz-Llácer (2022) discuss the methodological problems that this approach entails.

$$D(P_j) = \sum_{i=1}^n e^{-2.3 \left( \frac{i-1}{n-1} \right)} \frac{f_{ji}}{I_1}$$

wherein

$n$  = unit highest position regarding a specific centre of interest,

$i$  = unit position,

$j$  = unit index,

$e$  = Euler's number, or Napier's constant (2.7182828459045...),

$f_{ji}$  = absolute frequency of the word in  $j$  in the  $i$  position,

$I_1$  = number of informants participating in the survey,

$D(P_j)$  = availability of the word in  $j$ .

Based on this formula, one can find out the lexical availability index/degree of a specific word, its frequency, that is, the number of times the word appears with respect to the total number of units, its frequency of occurrence based on the percentage of informants who have mentioned this word during the survey and the cumulative frequency, that is, the sum of all relative frequencies. This formula has now been incorporated in different software applications and it can be used to calculate lexical availability. It is included in several platforms and software applications, such as Lexidisp (Moreno Fernández, Moreno Fernández & García de las Heras, 1995), Dispolex (<http://www.dispolex.com>), Dispogen (Echeverría, Urzúa & Figueroa, 2005) and DispoCen (Ávila Muñoz, Sánchez Sáez & Odishelidze, 2021) which, unlike the others, is conceived as an R package capable of extracting lexical availability data and also lexical centrality data (topic discussed in subsection Cognitive Approach).

It should be noted, however, that this is not the only formula that has been used in studies of lexical availability and, in fact, it has been criticised by some specialists. Callealta Barroso and Gallego Gallego (2016), for example, analyse other existing indexes, formulas, their usefulness and their differences, such as the indexes for lexical availability of a word, lexical availability of a subject, competence of the subject's lexical availability and the lexical compatibility of a word. This last concept, the lexical compatibility of a word, was used by Ávila Muñoz and Sánchez Sáez (2010 and 2011) to accurately describe the word association processes produced by subjects during availability tests. In these and similar works (Ávila Muñoz & Villena Ponsoda, 2010), the need to objectively justify the use of traditional formulas is raised for the first time. These authors propose alternative ways of finding the index of lexical availability using Fuzzy Sets Theory, which also allows to suggest new lines of research.

## **The Pan-Hispanic Project**

While the initial studies adopting the lexical availability approach concentrated mainly on French and English (Dimitrijévich, 1969; Mackey, 1971; Njock, 1979), Spanish is the language that has seen the most development of these studies and has been the subject of the highest number of publications. The first works on lexical availability related to Spanish-speaking communities emerged in Central America and the Caribbean in the 1970s, primarily authored by Humberto López Morales, who played a significant role in advancing the use of lexical availability in Spanish and disseminating these studies throughout the Hispanic world.

The Pan-Hispanic Project on Lexical Availability, coordinated by López Morales, gave a final and crucial boost to the study of lexical availability in Spanish. This project, which involved many Spanish and Latin American researchers, utilised a consistent methodology for over two decades in order to create lexically available dictionaries specific to various regions of the Spanish-speaking world. As a result, academics were able to draw linguistic, ethnographic, and cultural comparisons amongst different Spanish-speaking regions, and to specify geographical areas of dissemination, a useful starting point for further analysis.

Samper Padilla (1998) provided a detailed description of the frame methodology used in the Pan-Hispanic Project. In processing and editing language materials, Samper Padilla employed similar informants (pre-university level native speakers), centres of interest (the sixteen centres selected in the first French works), variables (such as gender, sociocultural level, geographical area—urban/rural—and type of school—public/private), and guidelines for lemmatisation.

## **Lexical Availability and Spanish as a Foreign Language**

The significance of adapting the lexical availability methodology to Spanish as a foreign language became evident after the launch of the Pan-Hispanic Project on Lexical Availability. According to Carcedo (2000, p. 46), availability tests for foreign Spanish speakers are valuable in determining how social and cultural backgrounds influence vocabulary knowledge, identify which elements of Spanish vocabulary individuals are able to activate, and determine the most common errors that are made. By testing students at different levels, we can also analyse the various stages of their lexical learning process, and grade lexical units accordingly. Comparisons between native and non-native communities are now feasible because availability studies have been conducted in both.

This provides a solid lexical foundation for textbook editors and other teaching materials for foreign language instruction.

In the early 1990s, research and articles on lexical availability began to criticise the limited vocabulary traditionally included in Spanish teaching materials. Benítez (1994) conducted a thorough analysis of three Spanish textbooks used for language teaching and concluded that the textbooks contained numerous irrelevant or completely unnecessary lexical units. Similarly, Benítez and Zebrowski (1993) conducted a study on the four Spanish textbooks most frequently used for language teaching in Poland and arrived at the following conclusion:

[A]uthors do not adequately select the vocabulary to be taught and they absolutely disregard the results of Spanish lexicon studies, i.e. the ones we base our work on and the ones most recently published in several Hispanic countries. (Benítez & Zebrowski, 1993, p. 229, author's own translation)

Carcedo (2000) pioneered the study of lexical availability for foreign Spanish speakers, using Finnish speakers as his focus. Carcedo first extracts their available Spanish lexicon using the same methodology employed in the Pan-Hispanic Project. He then creates a corpus with output and analyses and describes the specific trends observed in the vocabulary of these students. He highlights its applicability for error analysis, examines the features and evolution of that vocabulary in different learning phases, and accounts for the cultural peculiarities that are reflected in that lexicon between native Spanish speakers and Finnish students. He also concludes that the available lexicon of Finnish students, although quantitatively adequate, is not in line with the real lexical needs of everyday communication among Spanish speakers. Therefore,

good curricular programming in early learning phases should include words which are highly available amongst Spanish mother tongue speakers, since these are the most used. The incorporation of other, less frequent units should be avoided. (Carcedo, 2000, p. 219, author's own translation)

Samper Hernández's (2002) monographic study on lexical availability in Spanish as a foreign language is another significant work that includes students with various mother tongues. In the study, 45 students enrolled in Spanish courses at the University of Salamanca were interviewed about the 16 centres of interest outlined in the Pan-Hispanic Project. Samper Hernández applied four variables, namely gender, knowledge of other languages, level (beginner, intermediate, advanced or higher), and mother tongue (English, Italian, Japanese, and others), to the results. The study revealed that the available lexicon of foreign speakers is quantitatively much smaller than that of native speakers. The

variables such as gender and knowledge of other languages were found to have little influence on the number of words known by the speakers, whereas the overall level of knowledge of Spanish had a significant impact. Samper Hernández's methodological innovations are particularly noteworthy. She took the criteria of the Pan-Hispanic Project (as outlined by Samper Padilla, 1998) and adapted them for non-native Spanish speakers results. In doing so, she incorporated new elements such as the presence of words from the students' mother tongues, spelling peculiarities, and other problems observed in the recorded errors.

In Sánchez-Saus (2016) we conducted a study on 322 Spanish as a foreign language students, with six different mother tongues (English, French, German, Italian, Finnish, and Polish), and categorised their Spanish knowledge into A, B, and C levels of the Common European Framework of Reference for Languages (CEFR). The methodology used for data processing followed Samper Hernández's approach but with a modified list of centres of interest. Results indicated that factors such as gender,<sup>4</sup> knowledge of languages other than Spanish, or daily use of Spanish had no influence on the number of words known by the students. Significant differences were only found between A and B level informants, not between B and C levels. The study concluded that the greatest amount of lexicon is usually learned between lower and middle levels, and that the difference in the number of words known between middle and upper levels is not significant.

As of today, experiments similar to those described above have also been reproduced in non-Spanish-speaking countries with students of Spanish, for example, in Poland, Slovenia, Romania, Iceland, the United States, Turkey, China, Portugal, Korea, Morocco, among others. Aabidi (2019) provides an exhaustive account of them, to which we refer the reader for detailed information. Akbarian and Farrokhi (2021) also conduct a similar review, but focused on EFL.

### **Cognitive Approach: Study of the Relationships between Words in a Specific Centre of Interest**

Generating words out of categories, a phenomenon called “semantic fluency” or “category instance generation,” is quite common in cognitive studies. One of the pioneering researchers working on the links between lexical availability and cognition is Natividad Hernández Muñoz (Hernández Muñoz, Izura & Ellis,

---

<sup>4</sup> As Montero Saiz Aja (2021) points out, although gender is an important variable in second language acquisition, it has received little attention in studies measuring results from tests related to active vocabulary. She conducts a comprehensive review of these studies and carries out an analysis with EFL students, reaching the same conclusion as us: gender does not have an impact on this issue.

2006; Hernández Muñoz, Izura, & Tomé Cornejo, 2013; Hernández Muñoz & Tomé Cornejo, 2017). She began by analysing which cognitive factors (semantic, lexical, or both) influence the availability of words when generated in response to certain category labels. She then concludes that typicality, concept familiarity, age of acquisition, and word frequency are the criteria determining the availability of a given word.

Following these works, many later articles have analysed the cognitive mechanisms that characterise available lexicons, and the relationships between lexical units belonging to the same category or centre of interest. Here we can name only a few, but López González (2014) carefully revisits the first published articles on this idea, while putting into practice the described methods with Spanish learners from Poland.

Cañizal Arévalo (1991), Paredes García (2006), and Manjón (2008) were the first to observe the numerous relationships between available words. Each of them, using a different set of native Spanish speakers, observed that words tend to be organised into clusters or groups of closely related words. The later development of the DispoGrafo program (Echeverría et al., 2008) promoted this type of research and led to multiple publications that illustrate the relationships between words within a centre of interest through graphs. DispoGrafo was also employed by Ferreira and Echeverría (2010) to examine the relationships present in the lexicon of native and non-native English speakers (in this case, native Spanish speakers). The participants took a lexical availability test where they had to generate words from various semantic categories. The semantic connections found were subsequently illustrated through several graphs or semantic networks, which revealed that EFL students distinguish only large semantic categories, while native speakers of English organise words into highly specific subcategories. That is to say, EFL students created clusters—which were made explicit by means of several graphs or semantic networks—with fewer words and with less frequent relationships between the words, while natives created more clusters, denser ones and with more frequent relationships between words.

Santos Díaz (2017a) shares a similar objective to that of Ferreira and Echeverría (2010). Using the same software application, Santos conducts an analysis with Spanish native speakers, non-native speakers with French as their mother tongue, and non-native speakers with English as their mother tongue. Her study concludes that the results of the lexical availability task in native speakers show, as in Ferreira and Echeverría's study, dense clusters with a great number of subcategories inside (for example, in the human body: parts of the head, limbs and parts, internal organs, etc.). When asked to complete the survey in English or French, the number of identified words and relationships established between them is lower.

In other works, the nature of the relationships established in the available lexicon has been analysed. For example, in Sánchez-Saus (2016), a study carried

out only with non-native speakers of Spanish from different mother tongues, we conducted an analysis of the relationships within each centre, also using DispoGrafo. We considered relationships such as synonymy, hypernymy-hyponymy, antonymy, parasynonymy, metonymy, and meronymy, as well as cultural oppositions and terminological relationships. We concluded that the predominance of certain types of relationships depended on the type of centre of interest. When this could be considered a natural category (for example, “parts of the body” or “animals”), the relationships tended to depend on the meaning of the units (primarily hypernymy-hyponymy between the name of the category and the words inside it: “body”—“arm,” or “animal”—“dog,” and cohyponymy: “arm”—“leg,” “dog”—“cat”). When the centres were formed as radial categories or schemas, the relationships between words tended to be designative, especially metonymy or contiguity (for example, in a centre of interest like “school and university,” a relationship like “classroom”—“chair”—“table” is very frequent).

Meanwhile, Ávila Muñoz and Sánchez Sáez (2010, 2011) have conducted work on centrality that supports the cognitive principles of lexical availability. This research has led to the development of new lines of inquiry based on the Prototypes Theory, which propose innovative strategies for identifying community categories that are shared among members of specific communities. In the adaptation of the concept of availability to that of centrality, it was considered that each stimulus or centre of interest revolves around a prototype created from the concept that the stimulus itself determines. When an individual undergoes this type of experiment, they access their lexical network from the prototype generated by the initial stimulus as an entry point. Accessibility involves entering the central core of the lexical network represented by the stimulus, and from that access point, each individual will traverse their personal lexical network. Obtaining the structure of this lexical network for each subject is an impossible and useless task, as it is supposed to be determined by a multitude of uncontrollable biographical factors. However, from the particular realisations, the structure of lexical accessibility for a population in a specific stimulus or centre of interest can be quantitatively estimated. The quantification of this accessibility is the measure of the concept of centrality of each term for each stimulus, once the information provided by all individuals in the studied sample has been integrated (Ávila, Santos, & Trigo, 2020, p. 86). Thanks to the development of the concept and measurement of lexical centrality, levels of centrality or prototypicality of each word within the centre of interest or lexical category can be established.

All these papers reviewed under the cognitive approach are of enormous interest in themselves, as we have tried to show. Nonetheless, identifying the cognitive nature of the categories (or centres of interest) used in lexical availability and the types of relations that underpin them has applications of great

relevance for language teaching, especially foreign language teaching, as we will explore in the next section.

### **Didactical Approach: Selection of Vocabulary to Be Included in Teaching Materials of Spanish as a Foreign Language**

One of the main objectives of lexical availability has always been to improve the selection of the vocabulary in language teaching materials. Despite this, research employing lexical availability data with both native and non-native speakers has gained traction only in recent times. However, this scientific interest traces back 25 years ago when some publications began examining the lexicon chosen in Spanish as a foreign language teaching materials. These works already mentioned availability studies and pointed out their many deficiencies (see the analyses by Benítez & Zebrowski, 1993; Benítez, 1994).

Later in time, Bartol Hernández (2010) listed the advantages of using the available lexicon methodology when developing new learning materials. These are:

(a) the available lexicon methodology reduces the arbitrariness of vocabulary selection, as it is based on an academic corpus resulting from an availability survey;

(b) the words under study are followed by an availability index, which results from combining the word's frequency and survey position (a measurable calculation);

(c) it helps identify the lexical organisation of what is called "semantic memory" and the "mental lexicon" of the informants tested;

(d) the word output is grouped into cognitive fields called centres of interest, these being fairly equivalent to the topics and subtopics outlined in the CEFRL;

(e) it allows for diatopic comparison between different methodologies, thus including a sociolinguistic approach in teaching materials;

(f) such studies, at the same time, relate the impact of social aspects to the lexical proficiency of the student.

Additionally, Bartol Hernández (2010) concludes that lexical production tests prove also to be highly useful for a different purpose, that is, to determine the lexical mastery of students of Spanish as a foreign language, his discovery occurring at a time when lexical availability in Spanish as a foreign language was still under-developed.

Paredes García (2015) follows the same path and considers how to apply the results of lexical availability to the selection of vocabulary in the teaching of Spanish as a foreign language. He claims the importance of using an empirical basis for the lexical selection methodology and highlights the fact that the centres of interest used by lexical availability are very similar to those in-

cluded in the CEFRL and the Instituto Cervantes Curricular Plan. Additionally, Paredes describes three more criteria for lexical selection: (1) social distribution (percentage of surveys in which a term appears); (2) relative weight (based on the accumulated frequency of a term in a centre of interest), and (3) prototypicality. This last criterion refers to the relationship between a unit's semantics and its thematic category. Ávila Muñoz and Villena Ponsoda (2010, pp. 183–84) consider that the terms in availability lists respond to four types of association: (1) nuclear elements, (2) denotational associations, (3) derivations, and (4) individual associations. When analysing the lexical associations that speakers assign to a given centre of interest, we can determine the degree of semantic compatibility of a unit in relation to its original stimulus. Assuming the principles of the theory of prototypes, Ávila and Villena resort to the mathematical concept of *diffuse set*, according to which the relationship between an element and its set is not defined in traditional belonging/non belonging terms, but according to the degree of compatibility of the element with respect to its set.

Ávila Muñoz (2016) uses the concept of fuzzy sets for lexical selection. He takes hold of the theory of prototypes to explain the way words originate in the minds of speakers and then proposes an appropriate lexical selection method based on the Fuzzy Expected Value model to determine a series of guidelines that shall result in the development of a database for teaching purposes. The model aims to help Spanish as foreign language teachers to select the to-be-taught vocabulary according to the lesson's learning objectives and the level of their students. This selection is based on the compatibility index of each of the terms regarding its notional scope, that is, the degree of accessibility of a word naturally given by its native population.

The concept of fuzzy sets has also been used by Santos Díaz (2017b) to establish criteria for selecting vocabulary for CEFRL A1 and A2 levels, which is a very useful method for designing teaching materials. To a certain extent, Santos, Trigo, and Romero (2020) complement this work by considering the semantic relationships of available words using a cognitive approach when developing tasks for teaching materials. They design tasks based on the relationships such as hypernymy-hyponymy, morphological derivation, or metonymy observed in cohesive centres of interest, which they called "prototypical centres," and other tasks based on cultural aspects for less cohesive centres of interest, called "relational centres."

In their study, Jiménez Calderón and Rufat (2019) compared the results of a lexical availability study conducted on Spanish university students with a lexical frequency repertoire and with the examples provided in the inventory of specific notions of the Instituto Cervantes Curricular Plan (ICCP). Their objective was to create a trustworthy vocabulary list for the category of "parts of the body" and establish a methodology that can be applied to other areas

covered by the ICCP for future research. The authors' ultimate goal is to develop a comprehensive lexical repertoire for teaching Spanish at all levels.

Hidalgo Gallardo and Rufat (2022) present their own approach to selecting appropriate learning vocabulary for foreign students of Spanish. Their method involves a triangulation or combination of three sources of lexical information: (1) the lexicon used by native Spanish speakers, (2) the most commonly used Spanish words, and (3) the specific notions listed in the Instituto Cervantes Curricular Plan.

Finally, the utilisation of lexical availability has been employed to produce teaching materials for Spanish language learners who are immigrants. Jiménez Berrio (2013) conducted an analysis of the lexicon related to the human body that is presented in textbooks used for immigrant students and compared it to the actual available vocabulary of those same learners. The author determined that the vocabulary selection in the textbooks was insufficient in terms of quantity and lacked many essential words that should be included in secondary school curriculum, especially those related to non-visible body parts.

## Conclusions

Lexical availability is a highly significant area of research for determining the vocabulary that should be taught, as well as for analysing certain aspects of the configuration of the mental lexicon. However, current literature mainly focuses on Spanish and has been conducted predominantly by Spanish and Latin American researchers. This article aims to raise awareness of lexical availability among the international research community, particularly researchers interested in foreign language teaching, and extend our findings to other foreign languages.

We have aimed to outline the key developments in the history of lexical availability, which emerged with the objective of complementing the frequent vocabulary to determine the fundamental lexicon of a language. The French authors who pioneered this approach in the mid-20th century had a didactic objective: to select the lexical units to teach at the initial levels of teaching French as a foreign language. We have also examined the methodology that has been used in most studies, including the mathematical formula of lexical availability, the selection of categories or centres of interest for lexical extraction, and the latest methodological updates such as the concept of lexical centrality, which strengthen the links between the available lexicon and the prototypicality of certain elements within a semantic category.

In the section on applications of availability studies, we have focused on two perspectives: cognitive and didactic. Several studies have employed lists of available vocabulary to explore the configuration of the mental lexicon, including the relationships between words inside the categories, and the differences in these relationships across various semantic categories. As a result, many articles have been published in recent years featuring both Spanish native speakers and learners of Spanish as a foreign language. One significant finding from these articles is that foreign language learners have weaker relations established in their lexicon compared to those established by native speakers.

As for the didactic perspective, the results of the lexical availability studies have been very useful for selecting the vocabulary that should be taught at the different levels of Spanish as a foreign language and, therefore, that should appear in learning materials, as well as the extent to which the guidelines of the Instituto Cervantes Curriculum Plan correspond to the vocabulary that learners are able to activate. In addition, knowing what relationships are established within the semantic categories has been used to recommend more efficient didactic approaches to vocabulary.

## References

- Aabidi, L. (2019). La disponibilidad léxica en español como lengua extranjera: dos décadas de investigación científica [The Lexical Availability in Spanish as a Foreign Language: Two Decades of Scientific Research]. *marcoELE. Revista de Didáctica Español Lengua Extranjera*, 28. Retrieved from: <https://marcoele.com/disponibilidad-lexica-dos-decadas-de-investigacion/>
- Akbarian, I., & Farrokhi, J. (2021). Lexical availability and writing ability of EFL learners. *International Journal of Applied Linguistics*, 31, 48–64. <https://doi.org/10.1111/ijal.12319>
- Ávila Muñoz, A. M. (2016). El léxico disponible y la enseñanza del español. Propuesta de selección léxica basada en la teoría de los conjuntos difusos. *Journal of Spanish Language Teaching*, 3(1), 31–43. <https://doi.org/10.1080/23247797.2016.1163038>
- Ávila Muñoz, A. M. (2017). The available lexicon: A tool for selecting appropriate vocabulary to teach a foreign language. *Iranian Journal of Language Teaching Research*, 5(1), 71–91. <https://doi.org/10.30466/ijltr.2017.20343>
- Ávila Muñoz, A. M., & Sánchez Sáez, J. M. (2010). La disponibilidad léxica. Antecedentes y fundamentos. In A. M. Ávila Muñoz & J. A. Villena Ponsoda (Eds.), *Variación social del léxico disponible en la ciudad de Málaga: diccionarios y análisis* (pp. 37–81). Sarriá.
- Ávila Muñoz, A. M., & Sánchez Sáez, J. M. (2011). La posición de los vocablos en el cálculo del índice de disponibilidad léxica: procesos de reentrada en las listas del léxico disponible de la ciudad de Málaga [Word Rank in Lexical Availability Indexes: Re-entry Processes in Malaga Town Available Vocabulary]. *ELUA*, 25, 45–74. <https://doi.org/10.14198/ELUA2011.25.02>

- Ávila Muñoz, A. M., Sánchez Sáez, J. M., & Odishelidze, N. (2021). DispoCen. Mucho más que un programa para el cálculo de la disponibilidad léxica. *ELUA*, 35, 9–36. <https://doi.org/10.14198/ELUA2021.35.1>
- Ávila Muñoz, A. M., Santos Díaz, I. C., & Trigo Ibáñez, E. (2020). Análisis léxico-cognitivo de la influencia de los medios de comunicación en las percepciones de universitarios españoles ante la COVID-19 [A Lexical-Cognitive Approach to the Influence of the Media on the Perceptions of Spanish University Students Regarding COVID-19]. *Círculo de Lingüística Aplicada a la Comunicación*, 84, 85–95. <https://doi.org/10.5209/clac.70701>
- Ávila Muñoz, A. M., & Villena Ponsoda, J. A. (Eds.) (2010). *Variación social del léxico disponible en la ciudad de Málaga. Diccionario y análisis*. Sarriá.
- Bartol Hernández, J. A. (2010). Disponibilidad léxica y selección del vocabulario. In R. M. Castañer Martín & V. Lagüéns Gracia (Eds.), *De moneda nunca usada. Estudios filológicos dedicados a José M. Enguita Utrilla* (pp. 85–107). Instituto Fernando El Católico, CSIC. Retrieved from: <https://ifc.dpz.es/recursos/publicaciones/29/95/09bartol.pdf>
- Benítez Pérez, P. (1994). Léxico real/irreal en los manuales de español para extranjeros. In S. Montesa Peydró & A. Garrido Moraga (Eds.), *Actas del II Congreso Nacional de ASELE Español para extranjeros: didáctica e investigación* (pp. 325–333). ASELE. Retrieved from: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/02/02\\_0323.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/02/02_0323.pdf)
- Benítez Pérez, P., & Zebrowski, J. (1993). El léxico español en los manuales polacos. In S. Montesa Peydró & A. Garrido Moraga (Eds.), *Actas del III Congreso Nacional de la Asociación para la enseñanza del español como lengua extranjera (ASELE) El español como lengua extranjera: de la teoría al aula* (pp. 223–330). ASELE. Retrieved from: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/03/03\\_0221.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/03/03_0221.pdf)
- Callealta Barroso, F. J., & Gallego Gallego, D. J. (2016). Measures of lexical availability: comparability and standardization. *Boletín de filología*, 51(1), 39–92. <https://doi.org/10.4067/S0718-93032016000100002>
- Canga Alonso, A. (2017). Spanish L1 speakers' and EFL learners' available lexicon. *Anuario de estudios filológicos*, 40, 5–23.
- Cañizal Arévalo, A. V. (1991). Redes semánticas y disponibilidad léxica en el español de escolares mexicanos. In C. Hernández Alonso & G. De Granda Gutiérrez (Eds.), *El español de América: actas del III congreso internacional de el español en América. Valladolid, 3 a 9 de julio de 1989* (pp. 631–641). Junta de Castilla y León, Consejería de Cultura y Turismo.
- Carcedo González, A. (1998). Sobre las pruebas de disponibilidad léxica para estudiantes de español/LE. *RILCE*, 14(2), 205–224.
- Carcedo González, A. (2000). *Disponibilidad léxica en español como lengua extranjera: el caso finlandés (estudio del nivel preuniversitario y cotejo con tres fases de adquisición)*. Turun Yliopisto.
- Dimitrijévic, N. (1969). *Lexical Availability. A new aspect of the lexical availability of secondary school children*. Julius Gross Verlag.
- Echeverría, M. S., Urzúa, P., & Figueroa, I. (2005). *Dispogen II. Programa computacional para el análisis de la disponibilidad léxica*. Universidad de Concepción.
- Echeverría, M. S., Vargas, R., Urzúa, P., & Ferreira, R. (2008). DispoGrafo: una nueva herramienta computacional para el análisis de relaciones semánticas en el léxico disponible [DispoGrafo. A New Computational Tool for the Analysis of Semantic Relations in Lexical Availability]. *RLA. Revista de Lingüística Teórica y Aplicada*, 46(1), 81–91. <https://doi.org/10.4067/S0718-48832008000100005>
- Fernández Smith, G., Sánchez-Saus Laserna, M., & Escoriza Morera, L. (2012). Studies on lexical availability: The current situation and some future prospects. In Eizaga Rebolgar,

- B. (Ed.), *Studies in linguistics and cognition* (pp. 35–56). Peter Lang, Linguistics Insights Series, 158.
- Ferreira, R., & Echeverría, M. S. (2010). Redes semánticas en el léxico disponible de inglés L1 e inglés LE [Semantic networks of words in lexical availability studies of English L1 and English FL]. *Onomázein*, 21, 133–153.
- Ferreira, R., Garrido Moscoso, J. I., & Guerra Rivera, A. (2019). Predictors of lexical availability in English as a second language. *Onomázein*, 46, 18–34. <https://doi.org/10.7764/onomazein.46.03>
- Gómez-Devis, M. B., & Herranz-Llácer, C. V. (2022). Léxico disponible de escolares de la etapa primaria o básica: bases y propuesta metodológicas [Lexical Availability in Primary or Basic School-Children: Methodological Bases and Proposal]. *Pragmalingüística*, 30, 183–204. <https://doi.org/10.25267/Pragmalinguistica.2022.i30.09>
- González Fernández, J. (2014). Idoneidad de los centros de interés clásicos en los estudios de disponibilidad léxica aplicados al español como lengua extranjera [Suitability of the Centers of Interest in Classical Studies Lexical Availability Applied to Spanish as a Foreign Language]. *Revista Nebrija de Lingüística Aplicada a la Enseñanza de Lenguas*, 8(16), 41–53. <https://doi.org/10.26378/rnlael816223>
- Gougenheim, G., Michéa, R., Rivenc, P., & Sauvageot, A. (1964). *L'élaboration du Français Fondamental (I Degré). Étude sur l'établissement d'un vocabulaire et d'une grammaire de base*. Didier.
- Grève, M., & Van Passel, F. (1971). *Lingüística y enseñanza de lenguas extranjeras*. Fragua.
- Hernández Muñoz, N., Izura, C., & Ellis, A. W. (2006). Cognitive aspects of lexical availability. *European Journal of Cognitive Psychology*, 18(5), 734–755. <https://doi.org/10.1080/09541440500339119>
- Hernández Muñoz, N., Izura, C., & Tomé Cornejo, C. (2013). Cognitive factors of lexical availability in a second language. In R. M. Jiménez Catalán (Ed.), *Lexical availability in English and Spanish as a second language* (pp. 169–186). Springer. <https://doi.org/10.1007/978-94-007-7158-1-10>
- Hernández Muñoz, N., & Tomé Cornejo, C. (2017). Léxico disponible en primera y segunda lengua: bases cognitivas. In F. del Barrio de la Rosa (Ed.), *Palabras Vocabulario Léxico. La lexicología adaptada a la didáctica y a la diacronía* (pp. 99–122). Edizioni Ca Foscari. <https://doi.org/10.14277/978-88-6969-169-0>
- Hidalgo-Gallardo, M., & Rufat, A. (2022). Selección del vocabulario meta en español: propuesta metodológica basada en la triangulación de fuentes léxicas [Target Vocabulary Selection: A Methodological Proposal Based on the Triangulation of Lexical Sources]. *Onomázein*, 58. <https://doi.org/10.7764/onomazein.58.07>
- Jiménez Berrio, F. (2013). *Léxico disponible de inmigrantes escolares no hispanohablantes*. Servicio de Publicaciones de la Universidad de Navarra.
- Jiménez Calderón, F., & Rufat, A. (2019). Estudio de léxico disponible para la selección de vocabulario meta del español en enseñanza universitaria. In R. Roig-Vila (Ed.), *Investigación e innovación en la Enseñanza Superior. Nuevos contextos, nuevas ideas* (pp. 234–251). Octaedro. Retrieved from: [https://rua.ua.es/dspace/bitstream/10045/98882/1/Investigacion-e-innovacion-en-la-ES\\_023.pdf](https://rua.ua.es/dspace/bitstream/10045/98882/1/Investigacion-e-innovacion-en-la-ES_023.pdf)
- Jiménez Catalán, R. M. (Ed.). (2014). *Lexical availability in English and Spanish as a second language*. Springer.
- Jiménez Catalán, R. M., & Canga Alonso, A. (2019). The available English lexicon of male and female Spanish adolescents. *ELIA: Estudios de Lingüística Inglesa Aplicada*, 19, 157–176. <https://doi.org/10.12795/elia.2019.i19.07>

- Jiménez Catalán, R. M., & Fernández Fontecha, A. (2019). Lexical availability output in L2 and L3 EFL Learners: Is there a difference? *English Language Teaching*, 12(2), 77–87. <https://doi.org/10.5539/elt.v12n2p77>
- López Chávez, J., & Strassburger Frías, C. (1987). Otro cálculo del índice de disponibilidad léxica. In *Presente y perspectivas de la investigación computacional en México. Actas del IV Simposio de la Asociación Mexicana de Lingüística Aplicada* (pp. 1006–1014). UNAM.
- López Chávez, J., & Strassburger Frías, C. (1991). Un modelo para el cálculo del índice de disponibilidad léxica individual. In H. López Morales (Ed.), *La enseñanza del español como lengua materna: actas del II Seminario Internacional sobre Aportes de la Lingüística a la Enseñanza del Español como Lengua Materna* (pp. 91–112). Editorial de la Universidad de Puerto Rico.
- López González, A. M. (2014). Desarrollo de los estudios de disponibilidad léxica en Español Lengua Extranjera (ELE). In N. M. Contreras Izquierdo (Ed.), *La enseñanza del español como LE / L2, Actas del XXIV Congreso Internacional ASELE* (pp. 411–421). ASELE. Retrieved from: [https://cvc.cervantes.es/ensenanza/biblioteca\\_ele/asele/pdf/24/24\\_411.pdf](https://cvc.cervantes.es/ensenanza/biblioteca_ele/asele/pdf/24/24_411.pdf)
- López Morales, H. (1999). *Léxico disponible de Puerto Rico*. Arco/Libros.
- Mackey, W. F. (1971). *Le vocabulaire disponible du Français*, 2 vols. Didier.
- Manjón-Cabeza Cruz, A. (2008). Determinación de redes semánticas naturales por medio de encuestas de disponibilidad léxica. In A. Moreno Sandoval (Ed.), *El valor de la diversidad (meta)lingüística: Actas del VIII congreso de Lingüística General* (pp. 1–20, 64–65). Universidad Autónoma de Madrid. Retrieved from: <http://www.illf.uam.es/clg8/actas/index.html>
- Martínez Adrián, M., & Gallardo del Puerto, F. (2017). The effects of language typology on L2 lexical availability and spelling accuracy. *International Journal of English Studies*, 17(2), 63–79. <https://doi.org/10.6018/ijes/2017/2/256411>
- Michéa, R. (1950). Vocabulaire et culture. *Les Langues Modernes*, 44, 188–189.
- Michéa, R. (1953). Mots fréquents et mots disponibles. Un aspect nouveau de la statistique du langage, *Les Langues Modernes*, 47, 338–344.
- Montero Saiz Aja, A. (2021). Gender-based Differences in EFL Learners' Language Learning Strategies and Productive Vocabulary. *Theory and Practice of Second Language Acquisition*, 7(2), 83–107. <https://doi.org/10.31261/TAPSLA.8594>
- Moreno Fernández, F., Moreno Fernández, J. E., & García de Las Heras, A. J. (1995). Cálculo de disponibilidad léxica. El programa LexiDisp. *Lingüística*, 7, 243–250.
- Njock, P. E. (1979). *L'univers familier de l'enfant africain*. CIRB.
- Paredes García, F. (2006). Aportes de la disponibilidad léxica a la psicolingüística: una aproximación desde el léxico del color. *Lingüística*, 18, 19–55.
- Paredes García, F. (2014). A vueltas con la selección de centros de interés en los estudios de disponibilidad léxica: para una propuesta renovadora a propósito de la disponibilidad léxica en ELE [On Selection of the “Centers of Interest” in Classical Studies Lexical Availability Applied to Spanish as a Foreign Language]. *Revista Nebrija de Lingüística Aplicada*, 8(16). <https://doi.org/10.26378/rnlael816224>
- Paredes García, F. (2015). Disponibilidad Léxica y enseñanza de ELE: el léxico disponible como fuente curricular y como recurso en el aula [Lexical Availability and Spanish as a Second Language (ELE): Available Lexicon in the Syllabus and as Classroom Resource]. *Linred*, XIII. Retrieved from: <http://hdl.handle.net/10017/25257>
- Samper Hernández, M. (2002). *Disponibilidad léxica en alumnos de español como lengua extranjera*. ASELE.
- Samper Padilla, J. A. (1998). Criterios de edición del léxico disponible: sugerencias. *Lingüística*, 10, 311–333.

- Samper Padilla, J. A., Bellón Fernández, J. J., & Samper Hernández, M. (2003). El proyecto de estudio de la disponibilidad léxica en español. In R. Ávila, J. A. Samper, H. Ueda, et al. (Eds.), *Pautas y pistas en el análisis del léxico hispano(americano)* (pp. 27–140). Iberoamericana Vervuert.
- Sánchez-Saus Laserna, M. (2016). *Léxico disponible de los estudiantes de español como lengua extranjera en las universidades andaluzas*. Editorial Universidad de Sevilla, colección Lingüística.
- Sánchez-Saus Laserna, M. (2019). *Centros de interés y capacidad asociativa de las palabras*. Editorial Universidad de Sevilla, colección Lingüística.
- Sánchez-Saus Laserna, M. (2022). Redes semánticas, léxico disponible y didáctica del vocabulario en ELE: un análisis por niveles de español [Semantic Networks, Available Lexicon and Vocabulary Didactics in SFL. An Analysis in Different Spanish Learning Levels]. *Tejuelo*, 35(3), 167–204. <https://doi.org/10.17398/1988-8430.35.3.167>
- Santos Díaz, I. C. (2017a). Organización de las palabras en la mente en lengua materna y lengua extranjera (inglés y francés) [Word Organization in the Mind of Native Speakers and Foreign Language Learners (English and French)]. *Pragmalingüística*, 25, 603–617. Retrieved from: <https://revistas.uca.es/index.php/pragma/article/view/2406>
- Santos Díaz, I. C. (2017b). Selección del léxico disponible: propuesta metodológica con fines didácticos [Methodological Proposal for the Selection of the Available Vocabulary]. *Porta Linguarum*, 27, 122–139. Retrieved from: [https://www.ugr.es/~portalin/articulos/PL\\_numero27/9\\_Inmaculada.pdf](https://www.ugr.es/~portalin/articulos/PL_numero27/9_Inmaculada.pdf)
- Santos Díaz, I. C., Trigo Ibáñez, E., & Romero Oliva, M. F. (2020). La activación del léxico disponible y su aplicación a la enseñanza de lenguas [Activation of the Available Lexicon and Its Application to Language Teaching]. *Porta Linguarum*, 33, 75–93. Retrieved from: [http://www.ugr.es/~portalin/articulos/PL\\_numero33/5\\_Inmaculada%20Clotilde.pdf](http://www.ugr.es/~portalin/articulos/PL_numero33/5_Inmaculada%20Clotilde.pdf)
- Tomé Cornejo, C. (2015). *Léxico disponible. Procesamiento y aplicación a la enseñanza de ELE*. Doctoral Thesis. Universidad de Salamanca. Retrieved from: [https://gredos.usal.es/bitstream/handle/10366/128287/DLE\\_Tom%C3%A9CornejoC\\_L%C3%A9xicodisponible.pdf?sequence=1&isAllowed=y](https://gredos.usal.es/bitstream/handle/10366/128287/DLE_Tom%C3%A9CornejoC_L%C3%A9xicodisponible.pdf?sequence=1&isAllowed=y)

Marta Sánchez-Saus Laserna

## **Lexikalische Verfügbarkeit und Fremdsprachenunterricht: Hauptbeiträge eines wachsenden Bereichs**

### **Zusammenfassung**

Die Vermittlung von Vokabular stellt eine große Herausforderung im Fremdsprachenunterricht dar. In diesem Artikel werden die Vorteile erörtert, die die Ergebnisse von Studien zur lexikalischen Verfügbarkeit, bei denen es um die Extraktion von Vokabeln geht, die eng mit bestimmten lexikalischen Kategorien verbunden sind, für den Fremdsprachenunterricht bieten. Die Methode der lexikalischen Verfügbarkeit wurde erstmals in den 1950er Jahren in Frankreich entwickelt und wird seitdem vor allem für den Unterricht von Spanisch als Erstsprache eingesetzt. In den letzten 15 Jahren wurde sie auch für den Unterricht von Spanisch als Fremdsprache eingesetzt. In diesem Artikel werden die Ursprünge und die Methodik der lexikalischen Verfügbarkeit untersucht, wobei der Schwerpunkt auf zwei wichtigen Anwendungen im Zusammenhang mit dem Unterricht

von Spanisch als Fremdsprache liegt. Vom kognitiven Standpunkt aus betrachtet, behandelt der Artikel Studien, die die Konfiguration des mentalen Lexikons anhand von verfügbaren Vokabeln analysieren. Aus didaktischer Sicht untersucht der Artikel, wie lexikalische Verfügbarkeit genutzt werden kann, um die Auswahl von Vokabeln in Lehrmaterialien für Spanisch als Fremdsprache zu verbessern.

*Schlüsselwörter:* lexikalische Verfügbarkeit, Wortschatzunterricht, Fremdsprachenunterricht, Spanisch als Fremdsprache, kognitive Linguistik



**Zuzanna Wnuk**

Adam Mickiewicz University, Poznań, Poland

 <https://orcid.org/0000-0002-5883-7917>

## **Graphemic Awareness Development of Polish Learners of Chinese as a Foreign Language**

### **Abstract**

Metalinguistic orthographic awareness plays a crucial role in the development of Chinese language proficiency: writing, reading, and learning new words. However, few studies have explored the topic using qualitative methods. The aim of this work is to explore how the perception of Chinese characters of Chinese foreign language learners changes along with their proficiency. The responses were collected from 43 Polish university students of Chinese with the use of a graphemic awareness test during which participants were to decide on correctness of presented pseudo-characters. The results of this study show that beginners focus first and foremost on correctness of the strokes, the intermediate learners' attention shifts to character elements and their correct position, and the advanced learners analyze the characters taking into consideration both character elements and structure of characters. It is suggested that the overall rapid development of graphemic awareness in the study group might be due to the explicit instruction. The conclusion can be drawn that the described shift in learners' perception is a result of them learning to direct their attention towards the critical orthographic aspects of the characters.

*Keywords:* characters, Chinese as foreign language, Chinese characters learning, metalinguistic awareness, graphemic awareness

## **Introduction: Learning the Chinese Writing System as a Second Writing System**

The writing system is considered to be the most difficult aspect of learning Chinese as a foreign language (Chuang & Ku, 2011; Everson, 1998; Shen, 2004; Xu 2013). Multiple reasons for the difficulties have been indicated: the fundamental difference in the way Chinese and alphabetic writing systems encode linguistic information is one of them. The Chinese writing system is logographic, or more specifically, morphosyllabic—generally one grapheme encodes one syllable and one morpheme, while in alphabetic writing systems generally one grapheme encodes one phoneme. The Chinese writing system orthography is also different from the alphabetic one, with a large number of graphical units and complex ways according to which they are combined to compose a single grapheme.<sup>1</sup> The manner in which graphic components combine to form Chinese characters and in which characters construct words differ from the way letters construct words in alphabetic writing systems. Therefore, more morphological awareness is necessary when reading a text in Chinese (McBride, 2016; Wang et al., 2004). Other difficulties posed by the Chinese writing system noted by previous research are: the number of characters that needs to be learned, and lack of obvious correspondence between sound and script (Shen, 2005; Shen & Ke, 2007).

The process of learning a foreign language writing system is not the same as learning the rules of the one's first writing system. While the learner of a foreign language already knows how writing systems work in general, they may have to understand and internalize the different ways the foreign language writing system encodes linguistic information. To shift between a phonological writing system, as in the case of the Polish alphabet, and a morphosyllabic Chinese writing system, it is necessary for the learner to understand the different fundamental rules of the new writing system and develop the different aspects of metalinguistic orthographic awareness.

Currently, seven Polish universities offer Chinese Studies as a major, with six universities providing dedicated Chinese writing system classes. The duration of instruction hours during the first year of study varies between programs, ranging from 30 to 120 hours. A typical curriculum<sup>2</sup> includes

<sup>1</sup> More details on Chinese writing system orthography (structure and elements of characters) can be found in articles written by Wang et. al. (2004), Shen & Ke (2007).

<sup>2</sup> Data based on study programs as well as syllabi published by universities (in alphabetical order): Adam Mickiewicz University in Poznań, College of Modern Languages in Poznań (private university), Jagiellonian University in Kraków, John Paul II Catholic University of Lublin, University of Gdańsk, University of Warsaw, University of Wrocław. Only programs called “Chinese studies” or “sinology” were taken into consideration.

instruction on the structure of characters, including strokes, radicals, and other elements of characters, stroke order, basic information concerning the history of the Chinese writing system, and the ability to use dictionaries. The number of characters taught and the choice between simplified or traditional characters are dependent on the teaching program and didactic materials used by each university. Additionally, some programs may also include the basic principles of calligraphy and the role of characters in word formation. In case when dedicated Chinese writing system classes are not offered, the basic language competences needed to use Chinese characters are included in general Chinese language classes.

## Study Aim

There is significant interest in the metalinguistic orthographic knowledge of the Chinese writing system in the context of Chinese as a foreign language. However, research on this topic has predominantly been conducted using quantitative methods. Some researchers have addressed that problem; for example, Wang et al. (2004) had pointed out that the addition of post-experimental qualitative questions would lead to a better understanding of the orthographic awareness development process. This study was carried out to provide new insights into this topic using a qualitative method based on the phenomenographic approach to answer the questions:

- How graphemic awareness of Chinese writing system develops in Chinese as a foreign language learners coming from alphabetic background?
- How learners' perception of Chinese characters changes along with language proficiency?

Phenomenography approach is fitting to answer the abovementioned research questions, as its aim is to explore the qualitatively different ways in which people experience, conceptualize and understand various aspects of phenomena.<sup>3</sup> By employing the think-aloud protocol alongside a designed metalinguistic orthographic awareness assessment test, it was possible to explore the answers for the provided task and the reasoning behind the learners' choices. The aim of this paper is to report on the findings of the study investigating the development of graphemic awareness of the Chinese writing system of learners of Chinese as a foreign language coming from an alphabetic language background—adult Polish native speakers.

---

<sup>3</sup> More on phenomenography and its uses can be found in Orgill (2012) and Ornek (2008).

## Literature Review

### Phenomenography

The phenomenography approach emerged in the early 1970s with the publications of Marton, who was searching for new ways to conduct research in education and pedagogy. The main objective was to determine why some learners performed better than others, and to improve the teaching process (Jurgiel-Aleksander, 2016, p. 269). The aim of phenomenography was not examining a particular phenomenon, but rather exploring how it is experienced by people and how these experiences differ from each other (Moroz, 2013, p. 33; Orgill, 2012; Ornek, 2008; Tse et al., 2010, p. 81).

In the 1990s, the “new phenomenography,” also known as variation theory, emerged. According to this theory, each phenomenon is associated with a large amount of information, and due to the limitations of human cognition, it is impossible to perceive every aspect of a phenomenon at the same time. Learning involves changing mental representations of phenomena to qualitatively better ones (cf. Orgill, 2012). The main interest of the new phenomenography therefore shifted from describing different ways of experiencing a particular phenomenon to explaining the causes of different ways of experiencing it and utilizing this knowledge during teaching (Orgill, 2012).

The perspective of phenomenography research is referred to as a second-order perspective. The aim of phenomenography research is not to answer the question “what the phenomenon is” but rather “how the phenomenon is perceived” (Jurgiel-Aleksander, 2016, p. 270; Moroz, 2013, pp. 34–35). As such, phenomenography makes it possible to probe into mental representations of phenomena. The result of phenomenography research is the creation of a “map” showing a limited number of ways in which the particular phenomenon is perceived (Jurgiel-Aleksander, 2016, p. 270; Lam, 2010, p. 53; Moroz, 2013, p. 33; Ornek, 2008).

A semi-structured individual interview is the preferred method of data collection in phenomenography research. Its aim is to encourage the participants to reflect on their experiences related to the particular phenomenon. In educational research, participants are often asked to solve tasks using the “think-aloud protocol,” which allows researchers to gain insight into the participants’ thought process. This allows researchers to gather data that is essential for phenomenography, and often inaccessible with the use of other methods (Jurgiel-Aleksander, 2016; Moroz, 2013; Orgill, 2012; Ornek, 2008). Standard procedure for data analysis and interpretation, which aims to determine the various ways of experiencing the researched phenomenon, is described by Jurgiel-Aleksander (2016) and Ornek (2008).

The phenomenography approach was employed in previous research on learning the Chinese writing system to investigate the development of orthographic

awareness in Chinese children (Lam, 2006, 2010). According to Lam (2010, p. 53), “[according to variation theory], in order to experience the characters in a powerful way, children need to develop a structure of awareness, that is, to direct their focal attention to certain critical aspects of the characters and to be simultaneously aware of them.” Lam’s research specifically focused on exploring children’s knowledge of part-whole relations (relations between the character’s elements and the character) and part-part relations (internal relations between character elements), as well as its didactic implications (Lam, 2006, 2010).

In addition, the phenomenographic approach to teaching and learning—bringing learners’ attention to key aspects—has been tested in the context of Chinese writing system acquisition by Chinese children (Tse et al., 2010). The study introduced an integrated approach to teaching Chinese characters, which involved directing learners’ attention to key aspects of Chinese characters. The effectiveness of this new approach was compared to that of traditional teaching methods over the course of one academic year, in three primary schools in Hong Kong. The reported findings indicated that the proposed integrated approach was more effective in teaching the Chinese writing system to native learners (Tse et al., 2010).

## **Metalinguistic Orthographic Knowledge**

Metalinguistic awareness is generally understood as the ability to identify, analyze, and manipulate language units. Although researchers generally agree that it plays a crucial role in the process of learning to read and write, the conceptualization and exact definitions differ between studies. It can be understood as the ability to reflect on and manipulate the structure of the language; the ability to analyze the language and exhibit conscious control of it; perceiving the language as an object of potential analysis; a way to treat language during the process of production or comprehension (Kuo & Anderson, 2008, pp. 39–41).

In this article, the term metalinguistic awareness will be used to refer to the set of interconnected abilities that concern different structural aspects of the language, and the relations between the language and the writing system. Metalinguistic awareness can be therefore split into the following aspects: phonological, semantic, morphological, graphemic, graphophonological, and graphomorphological awareness. Throughout this paper, the term metalinguistic orthographic awareness will represent three aspects of metalinguistic awareness that refer to the writing system, that is, graphemic, graphophonological, and graphomorphological awareness.

Due to the different ways in which language units are encoded, different aspects of metalinguistic awareness develop in learners of different languages and writing systems. For example, in the case of alphabetic writing systems, such

as the Polish alphabet, the graphophonological awareness will be strongly developed. In the case of logographic writing systems, like Chinese characters, the graphomorphological awareness will be strongly developed. Previous research confirms that the phonological awareness is more important in the case of learning the English than the Chinese writing system, while in the case of the Japanese and Chinese writing systems, the morphological awareness plays a crucial role in reading development (Cook & Bassetti, 2005, p. 18; McBride, 2018, p. xi).

The graphemic awareness is an aspect of metalinguistic awareness that develops during the process of learning the logographic writing systems. It was first defined in the context of research on learning and teaching Japanese *kanji* as the “awareness that *kanji* can be segmented into graphemes and that graphemes can be the subject of analysis” (Toyoda, 1998, p. 156). For the purpose of this study, the graphemic awareness of the Chinese writing system is understood as the awareness of the characters structure as well as the position and function of characters’ elements. As only some of the character elements can function independently, they are not all considered graphemes but repeatable stroke clusters that characters are composed of instead.

Most research on metalinguistic orthographic knowledge of the Chinese writing system has focused on its connection to literacy, reading, writing, and learning new words. It has been shown that both learners of Chinese as a foreign language and native speakers made use of morphological, semantic, and phonological information provided by elements of characters (Tong & Yip, 2015). Studies on both groups have also indicated that the graphemic awareness is related to the ability to read characters and learning new ones (Anderson et al., 2013; Shen & Ke, 2007; Tong & Yip, 2015; Williams, 2013). The development of metalinguistic orthographic awareness has also been a topic of research. It has been established that it can develop without explicit instruction after repeated exposure to characters; however, the graphomorphological awareness of Chinese as a foreign language learners developed better after explicit instruction (Wang et al., 2004). In the same vein, Loh et al. (2018) state that knowledge concerning the position of character elements and the legitimacy of characters is implicit and, therefore, difficult to acquire. It has been also determined that encouragement to perceive characters as sets of meaningful parts, that is, using “Meaningful Interpretation and Chunking,” enhances students’ retention and memorization of Chinese characters (Xu 2013). Shen and Ke (2007) draw the conclusion that three different skills—knowledge of a character’s elements, perception of these elements, and their application—do not develop simultaneously, and each of them shows a unique developing trend. Loh et al. (2018) findings are also consistent with those results, as they confirmed that the three types of orthographic awareness develop in an asynchronous manner.

## Methodology

### Main Study Design

The data analyzed was collected during a larger study on the development of metalinguistic orthographic awareness of the Chinese writing system in Polish learners of Chinese as a foreign language. The approach is mainly based on the new phenomenography (variation theory) approach as it is suitable for addressing the research questions:

- How graphemic awareness of Chinese writing system develops in Chinese as a foreign language learners coming from alphabetic background?
- How learners perception of Chinese characters changes along with language proficiency?

To collect data on the perspectives of learners on characters, their knowledge, and skills in regarding characters, the think-aloud protocol was used during the test. The investigator did not use any terminology that had not been previously used by the study participants: the only term used was Chinese characters (Polish: chińskie znaki).

### Graphemic Awareness Assessment Test

The designed test concerned primarily the graphemic awareness of the Chinese writing system. The participants' task was to decide whether or not the presented characters were correct and provide the reasoning for their answers. The stimuli for this task were the following ten characters:

Figure 1

*Graphemic Awareness Task—Stimuli*



Because the aim of the study was to collect data on learners' orthographic awareness and not their vocabulary knowledge, nine out of ten Chinese characters presented are pseudo-characters that do not exist in the context of the Chinese writing system. Some of the pseudo-characters followed the structural rules of the Chinese writing system—the elements of the pseudo-characters were in their correct place, that is, pseudo-character 2, 3, 8, and 10. The rest of the pseudo-characters had specific elements in incorrect positions: pseudo-characters 1, 4, 5, 7, 9. The stimuli were designed with the use of common character elements and structures. The character number 6 was an existing character and was used as a control item to check whether participants answers would be different in case of real and pseudo-characters.

## Participants

A total of 43 Chinese foreign language learners from three course groups participated in this study. Criteria for selecting participants were as follows:

- participants had to be adult learners of Chinese as a foreign language;
- participants had to be Chinese language majors at university with 10 to 14 hours of Chinese language instruction per week with designated Chinese writing system classes;
- participants had to be learners who had started learning Chinese at university and had not previously learned other languages that use Chinese characters.

All of the criteria had to be fulfilled in order for learners to be eligible for participation in the current study. All participants were native Polish speakers learning Chinese at university as their major with little to no previous experience with learning the Chinese. The first-year curriculum of all of the participants included 60 to 70 hours of classes on the Chinese writing system. One participant was excluded from the current study on the basis of previous experience in learning Japanese. The participants were divided into three groups according to the number of hours of Chinese language learning, while the precise moment of data collection varied:

**Table 1**

### *Participant Groups*

Year of studies	First			Second		Third
Hours of instruction	100+	200+	400+	500+	600+	800+
Number of participants	7	9	6	5	5	10

Informed consent was obtained from all participants.

## Data Collection And Analysis

The responses of the participants were recorded and then transcribed. The notes of the participants on the task sheet were also collected as additional data. Data collection was carried out in Polish; examples from the collected data for this article were translated into English. During the data collection process, member checking was employed: restating the participants' responses and requesting confirmation or clarification of specific statements to verify the accuracy of the data. The transcripts were analyzed using the phenomenographic method in the following steps:

- Data familiarization (reading through the collected data);
- Data compilation (sorting the data thematically);
- Data condensation (finding the most significant elements in participants' answers);
- Preliminary grouping (grouping the determined elements together);
- Comparison (comparing and naming the determined categories);
- Creating outcome space (contrastive comparison of determined categories).

The data, including transcripts and notes of the participants, were analyzed using the MAXQDA Analytics Pro (ver. 2022) software. Themes were derived from the data in an inductive manner. The raw data, in Polish, is available at request from the corresponding author.

## Results

### Beginner

**Table 2**

*Beginner Learners' Answers on Pseudo-Character Correctness<sup>4</sup>*

Stimuli	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Correct	17	13	14	1	2	17	0	16	0	15
Incorrect	0	4	3	16	15	0	17	1	17	2
Strokes	1	6	5	3	2	2	1	3	3	6

<sup>4</sup>Key: PC: pseudo-character, C: character, number refers to Figure 1.

Table 3

*Detailed Responses with Learners' Comments on Strokes Marked in Grey<sup>5</sup>*

Participant	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
F 11	I	C	C	I	I	C	I	C	I	C
F 12	I	I	I	I	I	C	I	C	I	C
F 13	I	I	C	C	I	C	I	I	I	C
F 14	I	C	C	I	I	C	I	C	I	C
F 15	I	C	C	I	I	C	I	C	I	C
F 16	I	C	I	I	C	C	I	C	I	I
F 17	I	I	C	I	C	C	I	C	I	C
F 21	I	C	C	I	I	C	I	C	I	C
F 22	I	C	C	I	I	C	I	C	I	C
F 23	I	C	C	I	I	C	I	C	I	C
F 24	I	C	C	I	I	C	I	C	I	C
F 25	I	C	C	I	I	C	I	C	I	I
F 26	I	C	C	I	I	C	I	C	I	C
F 27	I	I	I	I	I	C	I	C	I	C
F 28	I	C	C	I	I	C	I	C	I	C
F 29	I	C	C	I	I	C	I	C	I	C

There were a total of 16 participants at the beginner level, 12 of whom identified as female and four of whom identified as male. The age range of the participants was 18 to 24 years, with a median age of 20 years and a mean age of 20.31 years. The results of the graphemic awareness test are presented in Table 2 and Table 3. Table 2 provides an overview of the beginner level learners' answers on correctness and incorrectness of presented pseudo-characters with an indication of how many of the participants elaborated on stroke correctness. Table 3 provides detailed answers that contain mentions of the correctness of the strokes indicated with grey color. From Table 3 it can be noted that multiple learners paid attention to strokes and their correctness during this task. All participants decided that C6 is correct. Participants comments did not differ between C6 and pseudo-characters. Through the analysis of the responses of the beginners, three themes were identified: strokes, elements, and uncertainty. These three themes are analyzed in detail in the following sections.

<sup>5</sup> Key: PC: pseudo-character, C: character number refers to Figure 1; F: first year of study, number refers to particular participant (numbers starting with 1: 100+ hours of instruction, starting with 2: 200+ hours of instruction); I: incorrect, C: correct.

### ***Theme One: Strokes***

Beginners during their analysis focused on general orthographic correctness of the presented pseudo-characters, that is, on strokes angles, length, their relation to other strokes. Examples of comments are as follows:

- F12 “This should be more slanting to the left, here to the right.”
- F13 “It looks like it was written with one stroke and it should be two.”
- F17 “It doesn’t stick out here, the horizontal line should stick out a little bit from the left and right sides.”

Some learners did not mention the incorrect position of the elements, their only reason for determining that the character is incorrect being the strokes position. In other cases, learners commented on the strokes first and then after some thought added comments concerning elements, for example:

- F12 “This stroke doesn’t touch that one, so this is incorrect” [referring to PC4, no mentions of elements].
- F14 “In this first element, this stroke shouldn’t touch the bottom one” [referring to PC4, no mentions of elements].
- F27 “I’ll start with that what I noticed first—this stroke should connect to this one” [skips the PC1, starts from PC2].

The learners also made comments on the general visual outlook of the characters, for example:

- F16 “I’ve never seen strokes in such a configuration.”
- F17 “The number of dots doesn’t fit, if there were two it would be more correct, but maybe three are fine too...”
- F23 “I’m not sure if it’s a problem, I’d write them closer, maybe more symmetrical.”

### ***Theme Two: Elements***

The beginners also noted the errors connected to the elements of the characters. Some of the comments concerning the elements were general remarks about their position, as in the following examples:

- F12 “It seems to me that the last three [dots/strokes], that last element, I, like, don’t know these characters, but as a rule, when I was learning, this last element was standing at the beginning of character, I think it opens the character.”
- F16 “This character arises my suspicions, the element on the left usually is on the right.”
- F17 “It seems to me that this element could also be on this side, but on the right side it would be more correct.”

The participants also mentioned the names of the semantic elements that they had recognized: *water* [ɣ̣], *person* [ɹ̣], *heart* [ɥ̣], and *speech* [ɨ̣]. Examples

of learners commenting on incorrect positions of elements while referencing their names are:

- F11 “It seems natural to me for *standing person* to be on the left side, this version seems unnatural to me.”
- F12 “I’m not sure with this one, *heart*, this element on the top, shouldn’t it be on the bottom?”
- F23 “According to my knowledge at this point it [this character] seems incorrect because I’d put *three drops of water* on this side, as I’m saying I don’t remember them to be on the right side.”

Beginner-level learners tended to have problems with recognition of the elements: they were not sure of their names, stated they did not know the elements, or they mistook elements for each other. In some instances, only after some time and thought did they point out the incorrect position of element in pseudo-characters:

- F11 “There should be one more dot on the right side [referring to the *knife* [!]] element in PC5, probably mistaking it for *heart* [!]] element].”
- F17 “*Heart* element was always on the bottom of character, and this element, I’ve never seen it before, so I don’t know if it’s correct.”
- F27 “I don’t have any concerns with this one [after a while], oh wait, that *standing person* should be rather on the left.”

### ***Theme Three: Uncertainty***

Another theme identified in the beginner-level learners’ statements is general uncertainty. The learners often used the phrases “for me,” “in my opinion,” “it seems to me” in their statements. In some cases, their explanation on their answers consisted only of their general intuition, and they could not elaborate on their reasoning:

- F12 “I’d say that this one could function.”
- F13 “This one seems good to me.”
- F17 “I think that this one could be seen as correct.”

Beginner-level learners also made comments on their limited knowledge of the characters and the Chinese language:

- F14 “It might be because of my lack of character knowledge, these three little strokes up and down, they are maybe called *three drops of water*...”
- F16 “I’m not a specialist on the Chinese language, but I think this one is incorrect.”
- F23 “As far as I know at this point, this character seems incorrect.”

## Intermediate

**Table 4**

*Intermediate Learners' Responses on Pseudo-character Correctness<sup>6</sup>*

Stimuli	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Correct	16	12	14	1	3	15	0	15	2	10
Incorrect	0	4	2	15	13	1	16	1	14	6
Strokes	0	3	2	2	0	0	1	1	2	2

**Table 5**

*Detailed Answers with Learners' Comments on Strokes Marked in Grey<sup>7</sup>*

Participant	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
S 11	I	I	C	I	I	C	I	C	I	I
S 12	I	C	C	I	I	C	I	C	I	C
S 13	I	C	C	I	C	C	I	C	I	C
S 14	I	C	C	I	I	C	I	C	I	C
S 15	I	C	C	I	I	C	I	C	I	I
S 16	I	I	I	C	I	C	I	C	C	C
S 21	I	I	I	I	I	C	I	I	I	C
S 22	I	C	C	I	I	C	I	C	I	C
S 23	I	C	C	I	I	C	I	C	I	C
S 24	I	C	C	I	I	C	I	C	I	C
S 25	I	C	C	I	I	C	I	C	I	I
S 31	I	C	C	I	I	C	I	C	I	I
S 32	I	C	C	I	I	C	I	C	C	C
S 33	I	C	C	I	C	I	I	C	I	I
S 34	I	C	C	I	I	C	I	C	I	I
S 35	I	I	C	I	C	C	I	C	I	C

<sup>6</sup>Key: PC: pseudo-character, C: character, number refers to Figure 1.

<sup>7</sup>Key: PC: pseudo-character, C: character number refers to Figure 1; S: second year of study, number refers to particular participant (numbers starting with 1: 400+ hours of instruction, starting with 2: 500+ hours of instruction, starting with 3: 600+ hours of instruction); I: incorrect, C: correct.

There were a total of 16 participants at the intermediate level, 11 of whom identified as female and five of whom identified as male. The age range of the participants was 19 to 28 years, with a median age of 21 years and a mean age of 21.19 years. Tables 4 and 5 provide an overview of the data from the test conducted on intermediate learners. The responses of the learners containing mentions of strokes' correctness are summarized in Table 4, while detailed answers are indicated with grey color in Table 5. What is noticeable here is that more than half of the participants do not mention strokes at all, focusing more on general character structure and position of the elements instead. All but one participant decided that C6 is correct. Participants comments did not differ between C6 and pseudo-characters. Four general themes were identified through the analysis of intermediate learners' answers: elements; structure of characters; uncertainty; stroke correctness. These themes are presented in detail in the following sections.

### ***Theme One: Elements***

Intermediate-level learners during their analysis of the pseudo-characters often referred to the elements by their Polish names, but sometimes they would also use their Chinese names. Sometimes only the word 'element' was used—but mostly in the cases of elements that are not radicals and therefore do not have names. Examples of learners mentioning the names of elements are:

- S13 “Again, in this character, this element, it’s *speech* I believe, I would put it on the left side of the character, under this, under this lid [referring to *roof* element].”
- S23 “Next character, it’s not good, here there is a *speech* character *yan*, on it should be on the left side, beginning the character.”
- S25 “The fourth character is probably not correct, because this element on the right side should be on the left side.”

The elements that learners recognized and mentioned in name are: *water* [氵], *person* [亻], *heart* [心], *speech* [讠], *roof* [宀], *knife* [刂]. Interestingly, only a few learners mentioned the *mouth* [口] element, although it is one of the most basic and common ones.

- S12 “And maybe here *heart*, I think heart usually is located on the bottom part of the character rather than top part.”
- S15 “The same with this *knife*, it can only be on the right side.”
- S24 “The fourth [character] is not [correct], because *person*, I think it should stand on the other side, I can’t remember any character in which *standing person* was on the right side.”

Intermediate learners used different terms to refer to elements, and sometimes the same participant used different terms throughout the task. The terms

used by the participants are “radical,” “element,” “character.” There was no significant pattern in the collocation of terms with particular elements—the same element could be called “element” or “radical”:

- S11 “The elem... radical *three water drops* is on the left side, so like, this character is incorrect for sure.”
- S21 “It’s rather strange that this *three water drops* element, that it is in this position.”
- S22 “The first [character] is incorrect, because the *water* character should be on the left side.”

In some cases, learners were not sure if they recognized the element correctly, forgot the name of the element in question, or mistook the element for another one. The learners most often had problems with *knife* [刂] element—instead of recognizing it and pointing out the incorrect position, learners would rather mistook it for *heart* [忄] element or *small* [小] element.

- S21 “At the end [pointing to the right side of character] there is something connected to the sea, it’s fish maybe... [referring to pseudo-character 10].”
- S22 “It’s the same here, because of... this... [pointing at the *speech* element], it should be on the left side if anything.”
- S23 “If I see correctly, there is a vertical line with a hook, so it should be *heart* or *xiao* [*small* element], but I’ve never seen a character with something like that standing in the front.”

### ***Theme Two: General Layout and Structure***

Intermediate level learners commented on the general layout, structure, and outlook of the characters. In some cases, the comments were connected mostly to how the participant felt about the character and whether or not it looked complete and orderly or chaotic and incorrect:

- S11 “This one looks too chaotic, this left side has this thing..”
- S14 “There are elements, that have their set places, it helps us decide if character is correct or not.”
- S16 “It looks full, it looks like a complete character, I like it a lot.”

Apart from general comments on character structure, participants often commented on particular elements of the characters—some learners would only express that the element did not feel right in the place it was in, but most of them would refer to their knowledge of correct element positions:

- S14 “I have to say that this character is incorrect, looking at the element that is located on the bottom right, it’s also an element that is never on the inside of the character, usually on the outside part, on the left.”
- S16 “I think that this heart fits at the top, but it would fit better at the bottom.”
- S24 “Something seems off, like this *heart* on the top doesn’t speak to me.”

What is also noticeable is that some participants not only analyzed positions of single elements in a character, they also took into consideration whether or not particular elements fitted each other:

- S14 “The combination of these two elements would be rather strange.”
- S16 “A lot of things don’t fit here, for sure *speech* on the left side, it also doesn’t go well with the character with a *roof*.”
- S24 “Something in it looks... I know what it is, this part that is under *roof*, it would look better if that *knife* and *mouth* would be put on the other side, and this element [referring to *speech* element], *sound* or something like that put on the other side.”

Therefore, it can be presumed that intermediate level learners start to perceive characters as a set of elements that compose a coherent single unit—a character.

### ***Theme Three: Uncertainty***

During the task, intermediate learners were sometimes uncertain of their answers, they would hesitate and use disclaimers such as stating that the particular character looked correct to them or referencing their little knowledge:

- S12 “[hesitation] The rest of the character looks... no, usually this *person* is also on the left side.”
- S21 “This element, I always see it on the other side, maybe I know too few characters.”
- S25 “The fifth character doesn’t look too correct for me, these two strokes on the left, or maybe after all it could be correct, I’m not sure.”

On the other hand, some of the participants showed greater confidence, giving strong “matter of fact” statements, quickly pointing out mistakes, and being more sure of their choices:

- S14 “I might not know it all, but from my current knowledge I can say that they [elements] don’t show up in this kind of place by themselves.”
- S16 “Solely from the fact that usually *water drops* are on the left side of a character, it’s a thing I simply remember since the first classes.”
- S21 “At a first glance, the top part of this character is incorrect, I mean that stroke, *knife*, doesn’t have to be so rounded and also that *speech* looks strange.”

In one case (S12), a participant would not analyze characters one by one, but rather focus and comment on certain characters that they deemed incorrect.

### ***Theme Four: Strokes***

Although comments on strokes were not prevalent in the responses from intermediate learners, some of them did mention strokes. The strokes could

also be a reason why the participant decided that the particular character was incorrect:

- S15 “I don’t know if it’s a matter of stylistics, but this stroke should be horizontal.”
- S21 “This falling left and downwards stroke is OK, but it should be sticking more closely to this vertical stroke.”
- S25 “This stroke at the bottom seems too long to me.”

While commenting on strokes, intermediate learners also mentioned the strokes with direct reference to the elements, in some cases exhibiting knowledge how a single stroke can change one element into another one:

- S16 “This character is almost perfectly complete, it misses one dot though, but maybe it could exist with *ice* [referring to *ice* and *water* elements].”
- S21 “This second character, *nü woman*, it not written completely correctly, I would write it in this way [writes the element].”
- S21 “At a first glance, the top part of this character is incorrect, I mean that stroke, *knife* doesn’t have to be so rounded and also that *speech* looks strange [not mentioning the incorrect position of elements].”

## Advanced

**Table 6**

*Advanced Learners’ Responses on Pseudo-character Correctness<sup>8</sup>*

Stimuli	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
Correct	0	10	10	0	1	10	0	8	1	6
Incorrect	10	0	0	10	9	0	10	2	9	4
Strokes	0	1	0	0	0	0	0	0	0	4

**Table 7**

*Detailed Answers with Learners’ Comments on Strokes Marked in Grey<sup>9</sup>*

Participant	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
T 11	I	C	C	I	I	C	I	C	C	C
T 12	I	C	C	I	I	C	I	C	I	I
T 13	I	C	C	I	I	C	I	I	I	C

<sup>8</sup>Key: PC: pseudo-character, C: character, number refers to Figure 1.

<sup>9</sup>Key: PC: pseudo-character, C: character number refers to Figure 1; T: third year of study, number refers to particular participant (numbers starting with 1: 800+ hours of instruction); I: incorrect, C: correct.

Participant	PC1	PC2	PC3	PC4	PC5	PC6	PC7	PC8	PC9	PC10
T 14	I	C	C	I	I	C	I	C	I	I
T 15	I	C	C	I	I	C	I	I	I	I
T 16	I	C	C	I	C	C	I	C	I	I
T 17	I	C	C	I	I	C	I	C	I	C
T 18	I	C	C	I	I	C	I	C	I	C
T 19	I	C	C	I	I	C	I	C	I	C
T 110	I	C	C	I	I	C	I	C	I	C

There were a total of ten participants at the advanced level, all of whom identified as female. The age range of the participants was 20 to 29 years, with a median age of 21 years and a mean age of 22.40 years. The results of the task conducted on advanced learners can be seen in Table 6, which presents general responses, and Table 7, which presents detailed responses of the participants. What is noticeable, in both Table 6 and Table 7, the comments concerning character strokes are sparse, while four out of five comments concern the pseudo-character 10. All participants decided that C6 is correct. Participants comments did not differ between C6 and pseudo-characters. In general, four themes were identified in the case of advanced learners: elements, structure, strokes, and certainty. These themes are described in detail in the following sections.

### ***Theme One: Elements***

The advanced level learners showed extensive knowledge when it comes to the correct position of the character elements, especially when it came to the semantic elements. While sometimes they would refer to the elements simply by pointing at them and not elaborating on their names, they would rightly adjust their position:

- T11 “This one is usually in the middle or at the bottom, it shouldn’t be here.”
- T18 “The fourth one doesn’t do it for me, I would change it, the left element I would put on the right side if anything.”
- T19 “The next element... next character [correcting themselves] looks correct because this element [referring to the *heart* element] appearing on the left side is correct.”

However, participants often recognized and used the correct names of the elements they were referring to. Advanced level learners mentioned the following elements: *water* [水], *erson* [人], *heart* [心], *heart* [忄], *speech* [讠], *roof* [宀], *knife* [刀], *hand* [扌], *ice* [冫], *fish* [鱼], *grass* [艹], *sun* [日], *big* [大], *woman* [女]. The examples of participants commenting on the elements are as follows:

- T12 “I’ve never seen *heart* at the very top, [only] in the middle or at the bottom.”

- T13 “It looks correct, there is *hand* in its place, everything that is on the right can be there, *grass, day, big*, I can’t see any mistakes.”
- T15 “I know that these two [strokes] mean *ice* and that one is *fish*.”

What is interesting, two of the participants during their elaboration on the characters correctness used the term “semantic element”—most of the participants would simply say “element” or “radical”:

- T13 “This is semantic element and it should be on the left.”
- T18 “So, this first one is incorrect, here on the right there are *three drops of water*, and it is usually a semantic element and it appears on the left.”

Although it was not the objective of the task, one of the participants tried to infer the meaning of the whole character from the recognized elements:

- T18 “The last one... it can be [correct], I don’t know it, it could mean frozen fish.”

### ***Theme Two: General Structure***

Advanced learners often made comments about the general structure and layout of the characters, commenting on how the character elements fit or do not fit each other. In some cases, they would include an explanation how they would fix the incorrect structure of the elements:

- T12 “The second one seems ok, everything fits.”
- T17 “This one seems to be fine in terms of the layout.”
- T18 “The second seems totally mixed, if it comes to, uuh, I haven’t seen a character which would have that *speech* element under the *roof*, and if it was to function, I would exchange the position of the elements under the *roof* with each other.”

In some cases, participants would not point to the elements of the character, instead referring to the general feeling the particular character was giving them or commenting that the whole character looked ‘good’ or ‘bad’:

- T11 “Hmmm... that last character also doesn’t look too good.”
- T13 “This is bad, it’s strangely written, this thing here.”
- T17 “I haven’t seen a character like that, but it’s probably ok, it doesn’t look super bad or anything.”

### ***Theme Three: Strokes***

While the strokes themselves did come up sparsely in the participants’ comments, it is still worth noting that four out of five comments concerning them were made while analyzing the tenth pseudo-character. Apart from pseudo-character 10, stroke correctness was not a factor in deciding whether the whole character is correct or not. It is also worth noting that the comments concerning strokes were directly connected to the elements:

- T12 “And this one here, if *fish* is on the right, it should have a straight stroke.”
- T13 “I’m trying to remember if in the *niang* character [娘], there was a dot at the top.”
- T16 “If there is nothing after it then this stroke here should be straight not diagonal [referring to the bottom stroke of *fish* element of PC10].”

The only instance when an incorrect stroke was the reason for stating that a character is incorrect was the case of PC10—each participant that decided that this pseudo-character was incorrect did so due to the bottom stroke of the *fish* element of this character.

### ***Theme Four: Certainty***

There was little to no hesitation in the answers of advanced learners; participants also took little to no time reflecting on their responses and moving quickly through the entire task. Some of the participants would skip characters that they consider correct:

- T14 “The second one is written correctly, [moves to the next one] this one too, [moves to the next one] this one is not, because the radical should be here, [moves to the next one] this one is also incorrect because it should be on this side.”
- T16 “In the first place, this [laugh], not on this side, so incorrect.”
- T17 “This one seems fine if it comes to the layout, [moves to next one] this *speech* here is good, [moves to next one] I would place *heart* on the bottom, it’s not ok.”

The participants were also rather confident about their knowledge, often stating that the particular element always appeared in a certain position, sometimes referencing other characters as examples. In some cases, the participants stated that they had not seen a character like that, but they still considered a character correct based on their current knowledge:

- T13 “This character also looks correct, there is *standing heart* on the left side, there is no problem, that component [referring to the right side of PC3] I remember [it] from the *jian* [检] character from the word *jiance* [检测].”
- T19 “This character, the second one, it seems correct, because these two elements, from what I remember, can be on both sides of the character, there is no rule that *nü* element has to be on the left or right side only.”
- T20 “Correct, but I’ve never seen a collocation like this.”

## Discussion

In this section, the following aspects of the study are summarized and discussed. First, the development of graphemic awareness of Chinese as a foreign language learner, specifically the knowledge of strokes, elements of character, and character structure, is addressed. The answers to research questions are also addressed in this section. The first subsection discusses the changes of learners' perception of characters, progressing from focusing on strokes to elements of character and overall character structure, as their language proficiency improves. The subsequent subsection presents the conclusions regarding the development of graphemic awareness of Chinese writing system of foreign language learners coming from an alphabetic language background. Finally, the effectiveness of the chosen research method, particularly the think-aloud protocol, is discussed, followed by an overview of the limitations of the current study and possible future research directions.

### Changing Perception: Shifting Focus from Strokes to Elements and Structure

As shown in the Results section, there was a notable shift in the responses. The beginners focused first and foremost on the strokes; they often mentioned stroke angles, lengths, and positional relation to other strokes. While beginners sometimes mentioned the general visual outlook of particular characters, they still focused on strokes. In some cases, for example PC4, the only mentioned reason for deciding the character was incorrect was the strokes—not the incorrect position of the *person* element [人]. It was noticeable that beginners would often start from the stroke analysis and only afterwards mention the elements.

This pattern changes when examining the responses of intermediate level learners—the most striking difference revealed was that more than half of intermediate learners did not mention strokes. They tended to focus more on general character structure and position of the elements. Nevertheless, some of the participants did mention strokes in their reasoning, and in some rare cases the strokes could be a reason why the character was deemed incorrect. What was also beginning to show in intermediate learners' answers and was more prevalent in advanced-level learners' answers was the fact that while commenting on the strokes, learners would do it in direct connection to the character elements. As for the advanced level learners, they only mentioned strokes in a few cases, in connection to the bigger structure and position of the element. An example of this is PC10: advanced-level learners who decided the character was incorrect did so due to the bottom stroke of the *fish* element [鱼].

Overall, the results show that the learners' attention gradually shifted from the strokes to character elements and their overall structure. The reasons for that change may be the following: along with more experience with the Chinese language and exposure to Chinese characters, learners develop the metalinguistic graphemic awareness implicitly, remembering the repeatable patterns as well as internalizing the rules that they were taught explicitly during Chinese writing classes. Learners also get more exposure to the handwritten characters and the differences between handwritten and printed characters—some aspects of the strokes and relations between them stop being perceived as errors and start being perceived as acceptable mistakes or variants of a character. It can be particularly seen in the answers of advanced-level learners as they pay attention to strokes when they could make a difference between characters.

The present study provides insight on the process of how learners' experience of Chinese characters changes along with the development of their graphemic awareness and increasing language proficiency. Similar to the case of native Chinese children learning Chinese characters described by Lam (2010), from the phenomenographic perspective, learners of Chinese as a foreign language also need to be aware of critical aspects of characters in order to use them proficiently. The present study results reveal that learners first attempt to distinguish as many aspects of the character as possible (focusing on strokes), then they gradually shift their attention to the critical orthographic aspects that differentiate characters or allow to distinguish between correct and incorrect characters. This shift in learners' perception could allow them to perceive Chinese characters in a more efficient manner: only paying attention to their critical aspects.

### **Graphemic Awareness: Knowledge of Elements and Character Structure**

During the current study, it was possible to observe how the number of comments regarding both elements and structure gradually rose along with the level of proficiency. Beginners mentioned elements less often than intermediate and advanced learners, as they focused mainly on the strokes. The attention of intermediate learners turned to character elements, but they sometimes forgot their names or mistook them for others. The advanced level learners demonstrated extensive knowledge of the elements in the case of both their position and meaning. Interestingly, learners themselves were overall aware of their knowledge—beginners took more time solving the task, often hesitated, and used phrases that pointed to their limited knowledge; on the other hand, intermediate learners were also sometimes uncertain of their answers and would use disclaimers, but they also exhibited more confidence in their answers;

advanced learners showed very little hesitation, answered quickly, and mostly gave “matter-of-fact” statements.

The difference in level of graphemic awareness could also be considered in regard to the structure of the characters. Beginners rarely commented on the structure on the level of individual elements. Intermediate-level learners commented on the general layout of the character, whether or not a particular element fits with others in a character. The comments concerning the structure sometimes were not based on logical reasoning but on the feeling about the particular character. The comments of the advanced learners were similar to those of the intermediate learners, but they explained their choices based on knowledge rather than their feelings. What is noticeable here is the change in perception of characters—from combination of strokes to combination of elements.

These results are in line with those obtained by Wang et al. (2004), as learners exhibited the sensitivity to character internal structure and quickly learned the correct position of individual elements. The results are also consistent with the trends reported by Loh et al. (2018) concerning quick development of element awareness as well as the positive influence of explicit instruction. Shen and Ke (2007) also showed in their research that beginner learners quickly become aware of the complexity of character structure, even when their knowledge of the elements is still limited.

Although this study confirmed the results of previous research, there is a difference in the way beginners who participated in the study perceived characters. In most works it is often claimed that beginners perceive characters as a whole rather than as a set of parts (e.g., Shen & Ke, 2007). However, in the current study even learners with the least experience seem to be perceiving the characters as sets of strokes rather than as an indivisible whole. The reason for this quick development might be both the presence of Chinese writing classes in the curriculum and the explicit instruction. While Loh et al. (2018) stated that the knowledge on elements of characters and the correctness of their position is implicit and difficult to acquire, it seems that explicit instruction facilitates the development of graphemic awareness. What might also be an important factor is the exposure to enough characters for learners to internalize the rules of the Chinese writing system that have been taught explicitly.

### **Effectiveness of the Think-Aloud Protocol**

The think-aloud protocol used during the current study proved to be an effective way to gain insight into the reasoning of the learners. Although learners of different proficiency levels provided the same responses in some cases, the chosen data collection method allowed for the determination of differences

in learners' reasoning behind a particular answer. Examples of such characters are: PC1, PC2, PC3, PC4—with strokes and not the position of individual elements being the reason for deciding a character was incorrect. The think-aloud protocol also allowed to differentiate reasons of learners who decided PC10 was incorrect: for example, beginner and intermediate learners would point to the left element wanting to correct it into *three water drops*, while advanced learners would point to the bottom stroke of the right element. There have been no previous studies exploring the underlying reasons for participants to determine the correctness of presented stimuli. Therefore, it would be beneficial to employ the think-aloud protocol in further research to gather additional data and validate the findings of the current study.

### **Limitations of the Study and the Directions of Future Research**

The current study was limited to a specific type of learner of Chinese as a foreign language—all participants were students of Chinese language majors at university level. While none of the participants had significant prior experience learning the Chinese language, their curriculum included classes on the Chinese writing system. The quick development of graphemic awareness and perception of characters as sets of strokes in the case of beginner learners might be due to the university curriculum. Further studies are needed to investigate the role of instruction to better understand the process of explicit and implicit learning. It would be interesting to compare how graphemic awareness develops without explicit instruction and how learners within such an environment perceive characters. Another question that remains unanswered is how quickly beginner learners start to perceive characters as sets of strokes rather than indivisible units, as in the current study, the first group of beginner learners had already received about 100 hours of Chinese language instruction. Further research could focus on exploring in detail the development of graphemic awareness in this short initial period.

### **Conclusions**

The results of this study reveal changing patterns in the perception of Chinese characters in the learners of Chinese as a foreign language. It was determined that beginner level learners perceive characters mostly as stroke clusters, while intermediate and advanced level learners shift their focus to individual elements and the overall character structure. Perception of what is

an error also changes, which is possibly due to more exposure to and understanding of the possible variations of handwritten characters. The conclusion can be drawn that the described shift in learners' perception is a result of them learning to direct their attention towards the critical orthographic aspects of the characters. The present study has also highlighted the importance of using a qualitative approach as it was possible to collect new significant data on the topic of metalinguistic orthographic awareness of Chinese writing system with the use of think-aloud protocol.

The findings of this study have several implications for the pedagogy of Chinese and the Chinese writing system. Firstly, the study confirmed that graphemic awareness develops rapidly among learners of Chinese as a foreign language, even those without prior experience with Chinese characters. Secondly, the explicit instruction of elements and structure of characters appears to facilitate and accelerate the development of graphemic awareness. It would be beneficial to direct learners' attention towards the significant relationships between strokes in characters that result in character changes, enabling them to identify the focal parts of characters. Learners' attention should be also directed towards graphically similar elements of characters, highlighting their differences to improve their ability to distinguish them. Explicit instruction on strokes, elements, and character structures, as well as exercises, such as distinguishing graphically similar characters or pseudo-characters, are not the only things that benefit the development of graphemic awareness of the Chinese writing system—correcting mistakes in strokes or elements, and categorizing characters based on their structure or elements can also play a beneficial role. Additional research comparing groups of learners that are not taught explicitly the rules of Chinese writing would help to determine the possible beneficial influence of the instruction on graphemic awareness development.

## References

- Anderson, R. C., Ku, Y.-M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. *Scientific Studies of Reading*, 17(1), 41–56. <https://doi.org/10.1080/10888438.2012.689789>
- Chuang, H.-Y., & Ku, H.-Y. (2011). The effect of computer-based multimedia instruction with Chinese character recognition. *Educational Media International*, 48(1), 27–41. <https://doi.org/10.1080/09523987.2011.549676>
- Cook, V., & Bassetti, B. (2005). Chapter 1: An introduction to researching second language writing systems. In V. Cook & B. Bassetti (Eds.), *Second language writing systems* (pp. 1–68). Multilingual Matters. <https://doi.org/10.21832/9781853597954-003>

- Everson, M. E. (1998). Word recognition among learners of Chinese as a foreign language: Investigating the relationship between naming and knowing. *The Modern Language Journal*, 82(2), 194–204. <https://doi.org/10.2307/329208>
- Jurgiel-Aleksander, A. (2016). Fenomenografia i jej poznawcze konsekwencje w świetle projektu na temat doświadczeń edukacyjnych dorosłych. Refleksja badacza. *Rocznik Andragogiczny*, 23, 267–282. <https://doi.org/10.12775/RA.2016.014>
- Kuo, L., & Anderson, R. C. (2008). Conceptual and methodological issues in comparing metalinguistic awareness across languages. In K. Koda & A. M. Zehler (Eds.), *Learning to read across languages* (pp. 51–79). Routledge. <https://doi.org/10.4324/9780203935668-10>
- Lam, H. C. (2006). *Orthographic awareness in learning Chinese characters*. The University of Hong Kong. [PhD dissertation]. <http://sunzi.lib.hku.hk/hkuto/record/B3762734X>
- Lam, H. C. (2010). Orthographic awareness. In F. Marton, S. K. Tse, & W. M. Cheung (Eds.), *On the learning of Chinese* (pp. 53–73). Sense Publishers.
- Loh, E. K. Y., Liao, X., & Leung, S. O. (2018). Acquisition of orthographic knowledge: Developmental difference among learners with Chinese as a second language (CSL). *System*, 74, 206–216. <https://doi.org/10.1016/j.system.2018.03.018>
- McBride, C. (2018). Foreword. In H. K. Pae (Ed.), *Bilingual processing and acquisition: Vol. 7. Writing systems, reading processes, and cross-linguistic influences: Reflections from the Chinese, Japanese and Korean languages* (Vol. 7, pp. ix–xii). John Benjamins Publishing Company. <https://doi.org/10.1075/bpa.7.fore>
- McBride, C. A. (2016). Is Chinese special? Four aspects of Chinese literacy acquisition that might distinguish learning Chinese from learning alphabetic orthographies. *Educational Psychology Review*, 28(3), 523–549. <https://doi.org/10.1007/s10648-015-9318-2>
- Moroz, J. (2013). Fenomenografia jako metoda badania treści świadomościowych. *Pedagogika Szkoły Wyższej*, 1, 33–44. <https://doi.org/10.18276/psw>
- Orgill, M. (2012). Variation theory. In N. Seel, *Encyclopedia of the sciences of learning* (pp. 2608–2611). Springer.
- Ornek, F. (2008). An overview of a theoretical framework of phenomenography in qualitative education: An example from physics education research. *Asia-Pacific Form on Science Learning Teaching*, 9(2), 1–14.
- Shen, H. H. (2004). Level of cognitive processing: Effects on character learning among non-native learners of Chinese as a foreign language. *Language and Education*, 18(2), 167–182. <https://doi.org/10.1080/09500780408666873>
- Shen, H. H. (2005). An investigation of Chinese-character learning strategies among non-native speakers of Chinese. *System*, 33(1), 49–68. <https://doi.org/10.1016/j.system.2004.11.001>
- Shen, H. H., & Ke, C. (2007). Radical awareness and word acquisition among nonnative learners of Chinese. *The Modern Language Journal*, 91(1), 97–111. <https://doi.org/10.1111/j.1540-4781.2007.00511.x>
- Tong, X., & Yip, J. H. Y. (2015). Cracking the Chinese character: Radical sensitivity in learners of Chinese as a foreign language and its relationship to Chinese word reading. *Reading and Writing*, 28(2), 159–181. <https://doi.org/10.1007/s11145-014-9519-y>
- Toyoda, E. (1998). Issues in the teaching and learning of Japanese. *Australian Review of Applied Linguistics. Series S*, 15, 155–168. <https://doi.org/10.1075/ara15s.15.10toy>
- Tse, S. K., Marton, F., Ki, W. W., & Loh, E. K. Y. (2010). Learning characters. In F. Marton, S. K. Tse, & W. M. Cheung (Eds.), *On the learning of Chinese* (pp. 75–102). Sense Publishers.
- Wang, M., Liu, Y., & Perfetti, C. A. (2004). The implicit and explicit learning of orthographic structure and function of a new writing System. *Scientific Studies of Reading*, 8(4), 357–379. [https://doi.org/10.1207/s1532799xssr0804\\_3](https://doi.org/10.1207/s1532799xssr0804_3)

- Williams, C. (2013). Emerging development of semantic and phonological routes to character decoding in Chinese as a foreign language learners. *Reading and Writing, 26*(2), 293–315. <https://doi.org/10.1007/s11145-012-9368-5>
- Xu, X., & Padilla, A. M. (2013). Using meaningful interpretation and chunking to enhance memory: The case of Chinese character learning. *Foreign Language Annals, 46*(3), 402–422.

Zuzanna Wnuk

## **Entwicklung des graphemischen Bewusstseins polnischer Lerner von Chinesisch als Fremdsprache**

### **Zusammenfassung**

Das metalinguistische orthografische Bewusstsein spielt eine entscheidende Rolle bei der Entwicklung von Sprachkompetenzen in Chinesisch: Schreiben, Lesen und Lernen neuer Wörter. Es gibt allerdings nur wenige Studien, die dieses Thema mit qualitativen Methoden untersucht haben. Ziel der vorliegenden Arbeit ist es, zu erforschen, wie sich die Wahrnehmung chinesischer Schriftzeichen durch Fremdsprachenlerner im Zusammenhang mit ihrem Sprachniveau verändert. Die Antworten wurden von 43 polnischen Chinesischstudenten mit Hilfe eines Tests des graphemischen Bewusstseins gesammelt, bei dem die Teilnehmer über die Korrektheit der präsentierten Pseudozeichen entscheiden sollten. Die Ergebnisse der Studie zeigen, dass sich Anfänger in erster Linie auf die Korrektheit der Striche konzentrieren, während sich die Aufmerksamkeit der fortgeschrittenen Lernenden auf die Zeichenelemente und deren korrekte Position verlagert, so dass sie die Zeichen unter Berücksichtigung sowohl der Zeichenelemente als auch der Struktur der Zeichen analysieren. Es wird vermutet, dass die insgesamt rasche Entwicklung des graphemischen Bewusstseins in der Lerngruppe auf den expliziten Unterricht zurückzuführen sein könnte. Daraus lässt sich schließen, dass sich die beschriebene Veränderung in der Wahrnehmung der Lernenden daraus ergibt, gelernt zu haben, ihre Aufmerksamkeit auf die kritischen orthografischen Aspekte der Zeichen zu richten.

*Schlüsselwörter:* chinesische Schriftzeichen, Chinesisch als Fremdsprache, Lernen chinesischer Schriftzeichen, metalinguistisches Bewusstsein, graphemisches Bewusstsein





**Hiromu Okamura**  
Shizuoka University, Japan

 <https://orcid.org/0000-0001-6425-6564>

**Tomohiko Shirahata**  
Shizuoka University, Japan

 <https://orcid.org/0000-0003-1280-5642>

## Factors Causing Overpassivisation of Unaccusative Verbs by Japanese Learners of English

### Abstract

It has been reported that second language (L2) learners of English including Japanese learners of English (JLEs) overpassivise unaccusative verbs although it is a kind of intransitive verbs. In order to account for the phenomenon, several assumptions have been proposed. However, so far, it is unclear which assumption is the most effective for explaining the overpassivisation of unaccusatives. Thus, this study tries to examine which of the three factors, animacy of subjects, existence of conceptualizable agents, or telicity of verbs, the most strongly affects overpassivisation of unaccusative verbs by JLEs. In this study, we conducted two experiments to examine this question. Study 1 was conducted with 100 university JLEs to compare the effect of animacy of subjects with that of the existence of conceptualizable agents. As a result, it was found that the animacy of subjects more strongly affected the overpassivisation of unaccusatives than the existence of conceptualizable agents. We conducted Study 2 with 101 university JLEs to examine which of the two factors, the existence of conceptualizable agents or the degree of telicity, is more influential. The results showed that the former was influential, but the latter was less so on the overpassivisation of unaccusatives. From the results of the two experiments, we concluded that the animacy of subject is the strongest influential factor among the three. On the other hand, the telicity of verbs hardly influences any errors.

*Keywords:* overpassivisation, unaccusative verbs, animacy, conceptualizable agents, telicity

It has been reported that second language (L2) learners of English overpassivise unaccusative verbs (Balcom, 1997; Hirakawa, 1995, 2006; Ju, 2000; Kondo, 2009, 2019; Kondo & Shirahata, 2015, 2018; Kondo et al., 2016, 2020; Montrul, 2000, 2001; Oshita, 1997, 2000; Sato, 2015; Shirahata et al., 2019, 2020; Yip, 1995; Yusa, 2003; Zobl, 1989). Because unaccusative verbs are a kind of intransitive verb, they must not be used as passives. Let us take an example of a verb *arrive*.

- (1) a. A letter arrived.
- b. \*A letter was arrived.

Since the verb *arrive* is an unaccusative verb, the active voice like (1a) is grammatical, but L2 learners erroneously passivise it as shown in (1b). This error is called an overpassivisation error.

A number of studies so far have argued why L2 learners including Japanese learners of English (JLEs) make such an error, and they have discussed and proposed plausible factors to account for the phenomenon. For example, some studies have supported the NP Marker Movement Hypothesis proposed by Zobl (1989), which supposes that L2 learners inappropriately associate the subject movement of unaccusatives with that of passive (Hirakawa, 1995; Oshita, 1997; Yip, 1995). Some claimed that existence of external agents in the discourse affected the overpassivisation of unaccusatives (Ju, 2000; Sato, 2015). Others stated that the first language (L1) transfer of the morphology was considered to be the factor causing the overpassivisation (Montrul, 2000, 2001; Kondo, 2009). Moreover, Yusa (2003) suggested that the property of telicity which unaccusatives have was influential on the error. More recently, Shirahata et al. (2020) argued that animacy of the subjects had influence on the overpassivisation, and also the degree of telicity of unaccusatives partially affected the error. There are also some studies that examined the overpassivisation from pedagogical perspectives such as the influence of instruction in L2 classrooms and in English textbooks (Kondo & Shirahata, 2015, 2018; Kondo et al., 2016, 2020; Shirahata et al., 2019).

In this way, various plausible factors for the phenomenon have been proposed in the previous studies. However, few of them have compared these factors at the same experiment. Thus, so far, it is unclear which factor can most strongly affect the overpassivisation of unaccusatives. Therefore, to clarify which factor is the most influential among the several factors proposed so far, we will focus on three factors in this study: They are the animacy of subjects, the existence of agents in the discourse, and the degree of verb telicity.

## Research Background

### Classification of English Verbs

Verbs appear in various types of syntactic structures. On the basis of whether they need an object or not, they are categorized into the following three types, intransitive verbs, transitive verbs, and verbs used as both an intransitive and a transitive. Intransitive verbs like *happen*, *appear*, and *sneeze* do not need an object, so they only take one argument of subject, that is, a subject. On the other hand, transitive verbs such as *resemble*, *discuss*, and *persuade* need two arguments, a subject and an object. Thus, they appear in the form of a DP(Determiner Phrase)-V(Verb)-DP structure. While intransitive verbs in the form DP-V-DP and transitive verbs in the form of DP-V are ungrammatical, some verbs like *break*, *open*, and *melt* are used both with the form of an intransitive and a transitive.

Intransitive verbs have two subcategories: unergative verbs and unaccusative verbs. Although they seemingly have the same syntactic structure, the semantic role that the subject of each verb bears is different. When unergative verbs such as *swim*, *run*, and *sleep* are used, the subject of the sentence generally becomes an agent of the action, who has a will to do the action like the subject *Ken* as in (2a). By contrast, unaccusative verbs like *fall*, *happen*, and *appear* require the semantic role of theme or patient for the subject. For instance, the subject *Tom* in (2b) does not intend to do the action, but the event happens without his will. In this setting, the semantic role of the subject is a theme.

- (2) a. Ken swam.  
b. Tom fell into the lake.

Since these two types of verbs give the sentential subject a different semantic role, their derivations of the syntactic structures are also different. In the process of generating the structure of unergatives as in (3a), the DP *Ken* merges with the verb *swam* and receives the semantic role of agent at the specifier of verb phrase (VP) from the verb *swam*. Then, the DP *Ken* moves to the specifier of the tense phrase (TP) to get assigned nominative case. Thus, the DP *Ken* becomes the subject with the semantic role of agent.

- (3) a. Ken swam.  
b.  $[_{TP} \text{Ken} [_T \emptyset [_{VP} \text{Ken} [_V \text{swam}]]]]$
- 

On the other hand, for unaccusative structures such as ‘The earthquake happened’ in (4a), *the earthquake* does not possess its own will to carry out an event, so the semantic role is a theme, not an agent. In a syntactic structure of unaccusatives as in (4b), the DP *the earthquake* is originally at the complement of the VP. Then, it receives the semantic role of theme there. After that, the DP moves to the specifier of the TP to get nominative case and becomes a sentential subject.

- (4) a. The earthquake happened.  
 b. [<sub>TP</sub> The earthquake [<sub>T</sub>  $\emptyset$  [<sub>VP</sub> [<sub>V</sub> happened ] ~~the earthquake~~]]]
- 
- The diagram consists of a horizontal line with a vertical arrow pointing upwards from its center. The line starts under the phrase 'the earthquake' in the VP complement and ends under the phrase 'The earthquake' in the TP specifier.

Therefore, since there is a difference of semantic roles of the subjects between unergative and unaccusative structures, these sentential subjects are originally at different positions in the VP. The idea that the subject of unaccusative verbs is originally at the position of VP complement (object) in the syntactic structure was first articulated in the Unaccusative Hypothesis (Perlmutter, 1978).

### Passive in English and Japanese

Transitive verbs can be used for passive as in (5b). In that case, the DP at the VP complement position moves to the subject position, and the auxiliary verb *be* is used. The agent *Ken* can be expressed with a preposition phrase “by + DP”

- (5) a. Ken wrote the book.  
 b. The book was written by Ken.  
 c. [<sub>TP</sub> The book [<sub>T</sub> was [<sub>VP</sub> [<sub>V</sub> written ] ~~the book~~]]]
- 
- The diagram consists of a horizontal line with a vertical arrow pointing upwards from its center. The line starts under the phrase 'the book' in the VP complement and ends under the phrase 'The book' in the TP specifier.

On the other hand, unergatives and unaccusatives have only one argument, a subject, so they cannot be passivised. Thus, unaccusatives with a passive form like (6b) is ungrammatical.

- (6) a. The ball fell into the lake.  
 b. \*The ball was fallen into the lake.

Let us discuss the usage of unaccusatives in Japanese. Similar to English, since unaccusatives in Japanese are used in active voice, the structure becomes ungrammatical when they are used in passive voice. For instance, the sentence

with an unaccusative *todoku* (*arrive* in English) is ungrammatical if it is used in passive as in (7b) (in Japanese, passive voice is formed by adding morphological maker ‘*reru*’ or ‘*rareru*’).

- (7) a. Tegami-ga Ken-no ie ni todoi-ta.  
 手紙が ケンの 家 に 届いた  
 a letter-Nom Ken’s house at arrive-past  
 ‘A letter arrived at Ken’s house.’
- b. \*Tegami-ga Ken-no ie ni todoka-re-ta.  
 手紙が ケンの 家 に 届かれた  
 a letter-Nom Ken’s house at arrive-passive-past  
 ‘\*A letter was arrived at Ken’s house.’

Therefore, unaccusatives are used in active voice both in English and in Japanese. However, even though unaccusatives are used in a similar construction in both languages, the overpassivisation error like (6b) is often observed in JLEs’ production of English sentences. Why does this happen? In the next section, we will examine previous studies dealing with the overpassivisation of unaccusatives by L2 learners.

## Previous Studies

### NP Movement Marker Hypothesis

As a pioneering study investigating the overpassivisation of unaccusatives by L2 learners of English, Zobl (1989) examined a corpus of written productions of L2 learners. From the data, he found that overpassivisation errors happened more frequently to unaccusatives than unergatives.

Zobl (1989) argued that L2 learners inappropriately associated the NP movement (i.e., NP at the VP complement position moves to the TP specifier) of passive with that of unaccusatives. More specifically, while unergatives have the subject bearing the semantic role of agent, both passive and unaccusatives lack a logical subject (an agentive subject) in the D-structure. Thus, the NP movement is applied in order to fill the lack of the logical subject. Zobl suggested that because passive and unaccusatives are similar in that the NP at the complement position of VP moves to the TP-spec position, L2 learners overextend the passive rule to the unaccusative rule once they have acquired the passive structure.

We call this idea that L2 learners confuse the NP-movement of passive with that of unaccusatives the “NP Movement Marker Hypothesis” (Zobl, 1989). So far, quite a few studies have examined overpassivisation of L2 learners from this perspective, and some of them support this hypothesis (e.g., Hirakawa, 1995; Oshita, 1997; Yip, 1995).

On the other hand, other studies have reported that the rate of overpassivisation of unaccusatives varies depending on verbs, subjects, and discourse type. If the NP Movement Marker Hypothesis is valid, the overpassivisation of unaccusatives should occur at the same rate for any unaccusatives even under different conditions. In the next section, we will discuss some other studies examining overpassivisation of unaccusatives from different points of view.

### Auxiliary Selection Hierarchy

As a different approach to overpassivisation of unaccusatives from the NP Movement Marker Hypothesis, some studies have examined the influence of telicity of unaccusatives (Hirakawa, 2006; Kondo, 2019; Shirahata et al., 2020; Yusa, 2003). For example, Yusa (2003) considered the property of telicity that unaccusatives have as a factor causing overpassivisation. Telicity refers to how clearly the endpoint of the activity of verbs is expressed. For instance, as to the verb *arrive* like (8a), we can understand that the action of *arrive* is completed when *Mary* reached the station. So we can imagine the clear endpoint (we call this property ‘telic’). On the other hand, for *belong* as in (8b), there is no endpoint (atelic). Thus, we can say that the degree of telicity of *arrive* is higher than that of *belong*, and *arrive* is a typical telic unaccusative and *belong* is an atelic unaccusative.

- (8) a. Mary arrived at the station.  
b. The land belongs to Mr. Suzuki.

Sorace (2000) suggested that unaccusatives and unergatives can be hierarchized according to the degree of the telicity and agentivity as shown in (9). This is called the Auxiliary Selection Hierarchy (ASH). ASH shows that the higher a verb is located in the hierarchy, the more telic and the less agentive it is. Sorace proposed that, although unaccusatives are more telic than unergatives, there is a difference of the degree of telicity even among unaccusatives. Verbs at the top of the hierarchy ‘Change of location’ such as *arrive* and *fall* are core unaccusatives, and verbs in the middle position like ‘Existence of state’ are more peripheral as unaccusatives. Unergatives are located in the lower place of the hierarchy. Thus, the verbs in the bottom ‘Controlled non-motional

process' like *play* and *work*, which are the most atelic and the most agentive, are core unergatives.

- (9) Auxiliary Selection Hypothesis (ASH, based on Sorace, 2000)
- |   |              |
|---|--------------|
| Core Unaccusative   | selects BE   |
| Change of Location (e.g., arrive, fall, come)             |              |
| Change of State (e.g., happen, appear)                    | ↑            |
| Continuation of a Pre-Existing State (e.g., remain, stay) |              |
| Existence of State (e.g., exist, belong)                  | ↓            |
| Uncontrolled Process (e.g., cough, sneeze)                |              |
| Controlled Motional Process (e.g., walk, run)             |              |
| Controlled Non-Motional Process (e.g., play, work)        |              |
| Core Unergative   | selects HAVE |

Sorace (2000) stated that, in many European languages, auxiliary verbs, *be* and *have* are often used when generating perfective aspects. In those cases, the auxiliary verb, *be* or *have*, conforms to the ASH. For the verbs being located in the higher position in the ASH, *be* is used. On the other hand, as verbs go down in the ASH, *have* is used.

Yusa (2003) claimed that if the auxiliary selection is not only applied to European languages but also to all languages, and if Universal Grammar (UG) functions in L2 acquisition as well as in L1 acquisition, it is not impossible to suppose that any L2 learners inherently have the knowledge of the ASH. Therefore, he assumed that L2 learners are more likely to accept overpassivisation with telic unaccusatives than with less telic ones and unergatives. He conducted a grammaticality judgement task including unergatives and unaccusatives with JLEs. As a result, he reported that they were more likely to accept overpassivisation with highly telic unaccusatives than less telic ones and unergatives. He concluded that the ASH can universally function in L2 acquisition as well as L1 acquisition, and overpassivisation occurs due to the restriction of the ASH.

Hirakawa (2006) also examined whether the degree of telicity of unaccusatives would affect overpassivisation by JLEs. In her study, she speculated that, if the error of overpassivisation reflects the auxiliary selection of perfective aspects and past tenses, this error will not be observed in present tenses. Thus, she conducted a grammaticality judgement task to examine this point. The results indicated that there was no difference in the rate of overpassivisation between telic and atelic unaccusatives and also no difference between past/perfective tenses and present tenses. Therefore, her results did not support Yusa's (2003) argument that the telicity of unaccusatives had an influence on overpassivisation of unaccusatives.

Kondo (2019) also conducted a grammaticality judgement task with JLEs to investigate the influence of telicity on overpassivisation. She pointed out some experimental problems of Hirakawa (2006), one of which was that Hirakawa used only three verbs for each type and Kondo believed that they were not enough. Thus, Kondo (2019) used unaccusatives and unergatives categorized in more detail. Concretely, she divided unaccusatives into three groups: ‘change of location,’ ‘change of state,’ and ‘existence of state’ based on the ASH. She also divided unergatives into three groups: ‘uncontrolled process,’ ‘controlled process; motional,’ and ‘controlled process; no-motional.’ She used five verbs for each group. From the experiment, she got two findings. First, JLEs tended to accept overpassivisation with unaccusatives more frequently than unergatives, and this result supported Zobl’s (1989) hypothesis. Second, there was no difference of test scores between telic and atelic unaccusatives. Therefore, her results did not support Yusa’s (2003) argument, so she concluded that the degree of telicity was not influential on overpassivisation of unaccusatives.

These three studies examined the overpassivisation with a different approach from the NP Movement Marker Hypothesis, but now, some new factors that can affect overpassivisation of unaccusatives have been suggested. For example, Shirahata et al. (2020) claim that animacy of subjects can affect the overpassivisation of unaccusatives, while Ju (2000) considers that conceptualizable agents can cause overpassivisation. We will discuss these two studies in the next sections.

## Animacy of Subjects

Shirahata et al. (2020) attempted to verify whether the animacy of subjects would cause overpassivisation of unaccusatives. They speculated that overpassivisation will occur more frequently when the subject is an immovable inanimate than when it is an animate. The reason is that, as a property of human languages, when the subject is an animate, we naturally feel that the subject actively does the action whatever semantic role it has. For instance, as in (10a), a person who runs in the park is undoubtedly *Ken*, and he actively does the action with his will. Likewise, as in (10b), a person who arrived at Shizuoka Station is *Hanako*. Even though a semantic role of the subject (*Hanako*) is not an agent but a theme in this case, we can imagine the situation where *Hanako* moved from somewhere and did the action (for example, she left Tokyo Station and reached Shizuoka Station) rather than she was made to do it by someone. By contrast, when the subject is an immovable inanimate such as *a letter* or *a book*, we feel that the subject would be passively involved in the event. As in (10c), since the subject *a letter* does not move actively by itself, we imagine that *a letter* is delivered by someone. Thus, Shirahata et al. (2020) supposed that

this property of human languages can cause JLEs to assume that the sentence would be passive when the subject is an immovable inanimate.

- (10) a. Ken runs in the park.
- b. Hanako arrived at Shizuoka Station.
- c. A letter arrived at Mary's house.

Note that the word “immovable” is important in their argument. Shirahata et al. (2020) mentioned that inanimate nouns are divided into two, movable inanimate and immovable inanimate. Movable inanimate is a noun that has the potential to move such as *a car*, *a train*, and *a storm*. By contrast, immovable inanimate is one that cannot move such as *a letter* and *a book*. In their previous experiments, they had observed that overpassivisation errors more often occurred when the subject was an immovable inanimate than when it was a movable inanimate and animates (Kondo, et al., 2015a, 2015b).

Shirahata et al. (2020) pointed out that previous studies which examined the influence of the telicity of unaccusatives on overpassivisation (e.g., Hirakawa, 2006; Kondo, 2019; Yusa, 2003) conducted the experiment without adequately considering the influence of the animacy of subjects. Thus, they argued that the issue of telicity should be investigated in an experiment after distinguishing animate and immovable inanimate subjects.

On the basis of the assumption mentioned above, Shirahata et al. (2020) set up the following two hypotheses and conducted an experiment to examine them.

- (11) Hypotheses by Shirahata et al. (2020)
  - a. In the course of L2 acquisition, since JLEs use the learning strategy, “when the subject is an animate, the sentence is active, while when it is an immovable inanimate, the sentence is passive,” overpassivisation is more likely to happen in the sentence whose subject is an immobile inanimate than in the sentence whose subject is an animate or a movable inanimate.
  - b. All L2 learners follow the ASH. Thus, JLEs as well as other L2 learners will overpassivise telic unaccusatives more frequently than atelic ones.

Shirahata et al.'s (2020) participants were 98 Japanese university students who learned English as a part of their general education subjects at their university. Their English proficiency level was lower intermediate (The TOEIC score range is from 380 to 420, which is equivalent to the CEFR A2).

Five unaccusatives, *arrive*, *fall*, *disappear*, *appear*, and *belong* were selected from the three categories of the ASH as in (12). Type 1 was verbs expressing ‘change of place’ such as *arrive* and *fall*, and the degree of telicity is high. The verbs of Type 2 were *appear* and *disappear*, which express ‘change of state,’ and the degree of telicity is lower than Type 1. Type 3 is ‘existence of state’ such as *belong* and the degree of telicity is the lowest of the three.

- (12) Type 1 change of place: *arrive* and *fall*  
 Type 2 change of state: *appear* and *disappear*  
 Type 3 existence of state: *belong*

Each of the five verbs was used for two sentences: one whose subject was an animate and the other whose subject was an immovable inanimate as in (13a–b). The participants were asked to read contexts written in Japanese and select which form, active or passive, they thought was grammatical for the target unaccusatives.

- (13) a. inanimate subject

The magician is staring at a big ball in the glass box. Then, (written in Japanese)

The big ball (disappeared/was disappeared) from the box.

- b. animate subject

Ken loves baseball and practices it every day.

Ken (belongs/is belonged) to the baseball club at school.

(Shirahata et al., 2020, p. 48)

As a result, the average score of animate subjects was 4.19 of 5, and that of immovable inanimate subjects was 2.47. A statistically significant difference between animate and immovable inanimate was observed. There was also interaction between animacy and verbs. For all the five verbs, the rate of overpassivisation with immovable inanimate subjects was significantly higher than that of animate subjects. Therefore, the results supported Hypothesis 1, and JLEs were more likely to overpassivise unaccusatives when the subject was an immovable inanimate than when the subject was an animate.

As for verb types, when the subject was an animate, the correct percentage of *fall* was significantly lower than that of *arrive* and *belong*. But, except for that, no difference was found among the other verbs. This result contradicted Hypothesis 2 because if Hypothesis 2 was correct, the correct percentage of Type 1 should have been lower than that of Type 2, and the percentage of Type 2

ought to have been lower than that of Type 3. On the other hand, when the subject was an immovable inanimate, the correct percentages of *arrive* and *fall* were significantly lower than those of *disappear*, *appear*, and *belong*. The percentage of *disappear* was also significantly lower than those of *appear* and *belong*. These results, to a large part, conformed to the prediction of Hypothesis 2. Thus, Hypothesis 2 was partially supported and partially rejected. That is, only when the subject was an immovable inanimate, the degree of verb telicity affected the overpassivisation of unaccusatives.

Shirahata et al. (2020) concluded that the animacy of subjects has a strong influence on the overpassivisation of unaccusatives. But only when the subject is an immovable inanimate, the verb telicity can affect the overpassivisation. Thus, they claimed that the verb telicity has a weaker influence on the overpassivisation of unaccusatives than the animacy of the subject.

### Conceptualizable Agents

Ju (2000) also challenged Zobl's (1989) hypothesis, the NP Movement Marker Hypothesis. She tackled overpassivisation of unaccusatives by Chinese learners of English with a different approach from the other studies. She assumed that the rate of overpassivisation would vary depending on the quality of the contexts previously provided. Considering this issue, she predicted that an agent in the discourse can be influential on the overpassivisation of unaccusatives.

This prediction is based on her following assumption. According to her, an agentless passive like (14a) and an unaccusative as in (14b) are similar because they both have nonagentive subjects. But the difference is that agents are omitted in an agentless passive, whereas agents inherently do not exist for unaccusatives.

- (14) a. A window broke.
- b. A window was broken.

However, even for unaccusatives, an agent-like DP can be pragmatically expressed in the discourse. For instance, as in (15a–b), the same sentence (*The ship sank slowly*) is used. In the preceding sentence of (15a), the DP '*a fighter jet*' causes the ship to sink. On the other hand, in (15b), the cause that the ship sank is in the ship itself, not because of other external factors.

- (15) a. A fighter jet shot at the ship. The ship sank slowly.
- b. The rusty ship started breaking up. The ship sank slowly.

(Ju, 2000, p. 92)

Ju (2000) called this externally conceivable agent “conceptualizable agent.” She also called a context with a conceptualizable agent like (15a) “externally caused event,” and a context without such an agent like (15b) “internally caused event.” She wondered whether the availability of conceptualizable agents in the discourse would have any effect on the overpassivisation of unaccusatives.

Based on the argument above, Ju (2000) set up the following hypothesis.

(16) Hypothesis by Ju (2000)

L2 learners of English are more likely to passivise unaccusatives when there is an externally caused event than when there is no such an event.<sup>1</sup>

Participants were 35 Chinese learners of English. They were graduate students at a university in America and their TOEFL score was in the range of 550–575, which is equivalent to the CEFR B1. She used five unaccusatives, *appear*, *die*, *disappear*, *emerge*, and *vanish* in a force-choice task. The participants were asked to read contexts and select a more grammatical form, active or passive, for the target accusative. All of the verbs were used for two types of contexts; External causation like (17a) and Internal causation like (17b). In sentences of external causation, there was a conceptualizable agent in the preceding sentence, while the sentences of internal causation did not have such an agent. In all the sentences, inanimate subjects were used.

(17) a. External causation

The magician did a trick with a coin.

The coin (vanished/was vanished) instantly.

b. Internal causation

A coin fell into the mud.

The coin (vanished/was vanished) instantly.

(Ju, 2000, p. 96)

As a result, it was found that the rate of overpassivisation errors in the external causation type was significantly higher than that of the internal causation type. This supported her hypothesis mentioned in (16). From the result, Ju (2000) concluded that cognitive factors played a key role in the acquisition of unaccusatives, and overpassivisation was more likely to occur when a conceptualizable agent was available in the discourse than when there was not such an agent.

<sup>1</sup> Although she also examined the difference between unaccusatives with transitive counterparts and those without, it is not related to our study, so we omit these details.

### Some Deficiencies of Shirahata et al. (2020) and Ju (2000)

Shirahata et al. (2020) and Ju (2000) investigated the phenomena of overpassivisation of unaccusatives that cannot be explained by the NP Movement Marker Hypothesis. However, it seems that there were some deficiencies in these studies. First of all, one problem of the Shirahata et al. (2020) study is that, although they examined the influence of the animacy of the subjects and that of the degree of verb telicity, they did not eliminate the influence of conceptualizable agents in the discourse, suggested by Ju (2000). In their experiment, they unsystematically used both sentences with a conceptualizable agent in the discourse and those without such an agent. For example, see sentences (18a–b). These are sentences used in their experiment. There is the possibility that the DP ‘the magician’ in the discourse of (18a) can be considered as an agent of an unaccusative *disappear*. On the other hand, there is no external agent that causes the event in (18b).

- (18) a. The magician is staring at a big ball in the glass box. Then, (written in Japanese)

The big ball (disappeared/was disappeared) from the box.

- b. Late last night, on my way home from the station, I was startled. This was because,

A big white object (appeared/was appeared) in front of me.

(Shirahata et al., 2020, p. 48)

In fact, the average score of (18a) was 0.47 and that of (18b) was 0.62. There was a significant difference of correct percentages between (18a) and (18b), although these two verbs, *appear* and *disappear*, are categorized into the same group of the ASH ‘change of state.’ They were expected to show no difference according to Shirahata et al.’s (2020) speculation. Thus, since not only the animacy of subjects and the degree of telicity of unaccusatives but also conceptualizable agents had a potential to affect the overpassivisation, their experiment should have been conducted by eliminating the factor of conceptualizable agents in the discourse.

As for Ju (2000), the problem is attributed to the ambiguous definition of conceptualizable agents in her experiment. For instance, see sentence (19), which was used in her experiment.

- (19) The police were searching for a jewelry box thrown into the river.  
The box (emerged/was emerged) suddenly.

(Ju, 2000, p. 110)

She used the sentence as an internally caused event, in which an agent should not exist in the discourse. But we are wondering if the participants thought that the DP ‘the police’ in the discourse was the agent of the event. Thus, we believe that no potential agent must show up in the test questions for internally caused events.

In addition, Ju (2000) did not take the degree of verb telicity into account. She used five unaccusatives, but in the data analysis she did not subcategorize them according to the ASH. Since Yusa (2003) and Shirahata (2020) suggested that there is a possibility that telicity of unaccusatives can affect the overpassivisation errors, an experiment should be conducted with the verbs subcategorized by the ASH and the factor of telicity should be eliminated.

Moreover, the findings of Shirahata et al. (2020) and Ju (2000) have raised an important question: which is the most influential factor causing overpassivisation, the animacy of subject, the degree of telicity, or the existence of conceptualizable agents in the discourse? So far, there have been no studies that have compared these three factors. Thus, this study will correct these experimental deficiencies and investigate this question.

In order to examine the research questions, we will conduct two experiments. One is to compare the influence of the animacy of subject with that of existence of conceptualizable agents (Study 1). The other is to investigate the influence of conceptualizable agents and the degree of verb telicity (Study 2). In the next two sections, we will show the outline and the results of the two experiments.

## Study 1

### Participants of Study 1

Participants were 103 Japanese university students who were learning English as a general education subject. The score range on the Oxford Quick Placement Test is from 28 to 39, which are categorized into the CEFR A2 and B1.

## Procedure

The participants of Study 1 were asked to read two sentences for each question. The first one set up a context written in Japanese, and the second one was written in English and included a target unaccusative. Then, they were asked to select a more grammatical form, active or passive, for the unaccusative.

There were four types of sentences as in (20). In Type 1 and Type 2, an immovable inanimate was used for the subject like ‘his textbook’ and ‘many historical buildings’ as in (20a–b). In the discourse of Type 1, there was a conceptualizable agent such as ‘Taro,’ but there was not in Type 2. In Type 3 and Type 4, the subject was an animate, for example, ‘Mr. Suzuki’ and ‘Ken’ as in (20c–d). A conceptualizable agent like ‘his doctor’ in (20c) was available in Type 3, but not in Type 4.

(20) a. Type 1: [-animate] and [+conceptualizable agent]

Ken didn’t bring his textbook to his house from school last. Friday.  
So,  
(written in Japanese)

His textbook (remained/was remained) in the classroom for the week-end.

b. Type 2: [-animate] and [-conceptualizable agent]

Japan lost a lot of historical buildings due to WWII. But fortunately,

Many historical buildings in Kyoto (survived/were survived).

c. Type 3: [+animate] and [+conceptualizable agent]

Mr. Suzuki suffered from a heavy disease. So, his doctor operated on him.

As a result,

Mr. Suzuki (survived/was survived) for a long time.

d. Type 4: [+animate] and [-conceptualizable agent]

Ken’s family had a trip, but he was about to have an entrance examination for a university. So,

Ken (stayed/was stayed) in his house.

As we have mentioned before, the deficiency of Ju (2000) is that there were potential agents even in the proceeding sentences of internally caused events. So as to avoid the deficiency, in this study we tried to ensure that the proceeding sentences in Type 2 and Type 4 did not include a person that had the potential to affect events of the target sentences.

Since Study 1 will examine the animacy of subjects and the existence of conceptualizable agents, we have to eliminate the influence of telicity of unaccusatives in the experiment. Thus, we have used three verbs, *stay*, *remain*, and *survive*, of the same category ‘continuation of a pre-existing state’ in the ASH (see (9)). These three verbs were used in every type. Since we had 12 target sentences with 12 fillers, there were 24 sentences in total.

Moreover, to exactly investigate the JLEs’ grammatical knowledge of unaccusatives, the methodology of Study 1 must be based on the premise that the participants know the meanings of the target unaccusatives. Therefore, we conducted a vocabulary test, where they were asked to select a suitable translation of the target unaccusative into Japanese from five choices. Then, we eliminated participants from the analysis who selected wrong choices of the target unaccusatives.

Three of the participants selected incorrect choices of the target unaccusatives in the vocabulary test. So, we removed them and finally analyzed 100 participants.

## Hypotheses of Study 1

With the study, we will examine the following three hypotheses.

(21) a. Hypothesis 1:

As Shirahata et al. (2020) suggested, JLEs are more likely to overpassivise unaccusatives when the subject is an immovable inanimate than when it is an animate.

b. Hypothesis 2:

As Ju (2000) claimed, JLEs are more likely to overpassivise unaccusatives when there is a conceptualizable agent in the discourse than when there is not such an agent.

c. Hypothesis 3:

In a sentence in which the subject is an immovable inanimate and there is a conceptualizable agent in the discourse, JLEs are most likely to overpassivise unaccusatives than in a sentence in which the subject is an inanimate without a conceptualizable agent or in a sentence in which the subject is an animate with a conceptualizable agent.

If Hypothesis 1 is correct, we predict that the correct percentage of Type 1 will be lower than that of Type 3. Also, the correct percentage of Type 2 will be lower than that of Type 4. Based on Hypothesis 2, it is expected that the correct percentage of Type 1 will be lower than that of Type 2, and the correct percentage of Type 3 will be lower than that of Type 4. From Hypothesis 3, we assume that the correct percentage of Type 1 will be lower than those of Type 2 and Type 3, and that of Type 4 will be the highest.

## Results and Discussion of Study 1

We show the percentages of correct responses (the number of correct answers divided by the number of all sentences, 300 (three verbs  $\times$  100 participants)) of the four types in Table 1 and Figure 1. The correct percentage of Type 1 was the lowest of the four types at 38.7% (116/300). The correct percentage of Type 2 was 73.7% (221/300), and that of Type 3 was 92.0% (276/300). The percentage of Type 4 was the highest of the four types at 97.0% (291/300), so participants were almost completely able to choose correct answers in Type 3 and Type 4.

Repeated measures ANOVAs revealed significant effects of sentence types ( $F(3, 297) = 122.370, p < .001$ ). Multiple comparisons showed that the correct percentage of Type 1 was significantly lower than that of Type 3. Also, the correct percentage of Type 2 was significantly lower than that of Type 4. This result is consistent with Hypothesis 1, and JLEs are more likely to overpassivise unaccusatives when the subject is an immovable inanimate than when the subject is an animate.

In addition, there was a significant difference between the correct percentage of Type 1 and that of Type 2. However, there was no significant difference between the correct percentage of Type 3 and that of Type 4, since participants can choose correct answers at a quite high rate in both types. This finding partially supported and falsified Hypothesis 2. That is, only when the subject is an immovable inanimate, the existence of conceptualizable agents can affect the overpassivisation of unaccusatives. But, from the results that the percentages of both Type 3 and Type 4 are relatively high, agents in the discourse were not influential when the subject is an animate. Therefore, even if conceptualizable agents exist in the discourse, JLEs are less likely to overpassivise unaccusatives unless the subject is an immovable inanimate.

Moreover, the correct percentage of Type 1 was significantly lower than that of Type 2 and that of Type 3. Thus, Hypothesis 3 has been supported. When the two factors, the animacy of subjects and the existence of conceptualizable

agents, are available, the overpassivisation of unaccusatives happens more frequently than when either of the two factors is involved.

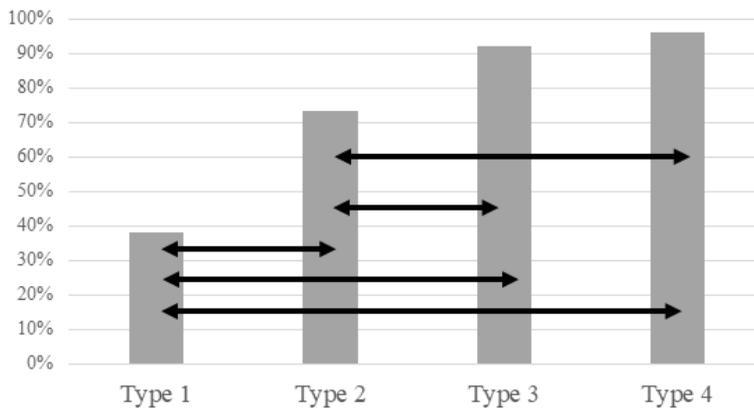
**Table 1**

*Correct Percentage of the Four Types in Experiment 1*

Sentence Type	Mean	SD
Type 1	38.7%	1.120
Type 2	73.7%	0.700
Type 3	92.0%	0.531
Type 4	97.0%	0.377

**Figure 1**

*Correct Percentage of the Four Types in Experiment 1*



While the animacy of subject affected the overpassivisation of unaccusatives regardless of whether a conceptualizable agent exists or not, the existence of conceptualizable agents can be influential only when the subject is an immovable inanimate. This finding suggests that the animacy of subjects is a stronger factor causing the overpassivisation than the existence of conceptualizable agents.

As we explained in the procedure, we chose the three unaccusatives from the same category of the ASH ‘continuation of a pre-existing state,’ so the degree of telicity of the verbs would not have influenced the overpassivisation according to the ASH. Now in the next section of Study 2, we will compare the influence of the existence of conceptualizable agents with that of the telicity of unaccusatives to examine which is a stronger factor causing overpassivisation.

## Study 2

### Participants of Study 2

Participants of Study 2 were 122 Japanese university students who learned English as a general education subject. They were different participants from those of Study 1. The score range on the Oxford Quick Placement Test is from 28 to 45, most of which are categorized into the CEFR A2 and B1.

We also conducted a vocabulary test like Study 1 to remove participants who did not know the meanings of the target unaccusatives. Twenty-one of the participants selected incorrect choices of the target unaccusatives in the vocabulary test, so we removed them from the analysis and finally analyzed 101 participants.

### Procedure

Similar to Study 1, the participants were asked to read two sentences for each question. The first one set up a context written in Japanese. The second one was written in English and included a target unaccusative. Then, JLEs were asked to select a more grammatical verb form, active or passive, for the unaccusative.

We had four unaccusatives, *arrive*, *come*, *stay*, and *remain*, and they were divided into two categories based on the ASH. See (9), *arrive* and *come* were used as verbs belonging to ‘change of place’ in the ASH, and their degree of telicity is high. *Stay* and *remain* belong to ‘continuation of a pre-existing state,’ and the degree of telicity of these verbs is lower than that of *arrive* and *come*.

There were four types of sentences as in (22). In Type 5 and Type 6, the two telic unaccusatives, *arrive* and *come* were used. In the preceding sentences of Type 5, there was a conceptualizable agent like ‘postman’ as in (22a). But there was not in Type 6 like (22b). On the other hand, Type 7 and Type 8 used two atelic unaccusatives, *stay* and *remain*. There was a conceptualizable agent in Type 7 like ‘Taro’ as in (22c) while there was not in Type 8 like (22d). In order to avoid the influence of the animacy of subjects, we used immovable inanimate subjects in all the sentences of the four types. Since there were eight target sentences with 13 fillers, we conducted Study 2 with 21 sentences in total.

- (22) a. Type 5 [+agent] and [+telicity]: *arrive* and *come*  
I heard the sound of a postman’s bike. Apparently, (written in Japanese)

A letter (came/was come) to my house.

b. Type 6 [-agent] and [+telicity]: *arrive* and *come*

The river near my house was flooded with the storm, and my house was inundated with water. What was worse,

The water (came/was come) up to the second floor of my house.

c. Type 7 [+agent] and [-telicity]: *stay* and *remain*

Taro didn't bring his textbook to his house from school last Friday. So,

His textbook (remained/was remained) in the classroom for the weekend.

d. Type 8 [-agent] and [-telicity]: *stay* and *remain*

A lot of plants in the region were extinct due to air pollution. But,

Only the trees called 'Meliaceae' (remained/were remained) in the region.

## Hypotheses of Study 2

We will conduct Study 2 to test the following three hypotheses.

(23) a. Hypothesis 4:

JLEs are more likely to overpassivise unaccusatives when there is a conceptualizable agent in the discourse than when there is not such an agent.

b. Hypothesis 5:

JLEs are more likely to overpassivise telic unaccusatives than atelic ones.

c. Hypothesis 6:

When the two factors are available, that is, when there is a conceptualizable agent with a telic unaccusative, JLEs are more likely to overpassivise than when one of the factors is involved, that is, when there is a conceptualizable agent with an atelic unaccusative or when there is not such an agent with a telic unaccusative.

If Hypothesis 4 is correct, we predict that the correct percentage of Type 5 will be lower than that of Type 6. Also, the correct percentage of Type 7 should be lower than that of Type 8. According to Hypothesis 5, it is assumed that the correct percentage of Type 5 will be lower than that of Type 7, and the percentage of Type 6 will be lower than that of Type 8. Based on Hypothesis 6, we can expect that the correct percentage of Type 5 will be lower than those of Type 6 and Type 7, and that of Type 8 will be the highest.

## Results and Discussion of Study 2

Table 2 and Figure 2 indicate the percentages of correct responses of the four types. The correct percentage of Type 5 was 73.8% (149/202) and that of Type 6 was 89.6% (181/202). The correct percentage of Type 7 was the lowest of the four types at 51.5% (104/202) and that of Type 8 was 82.7% (167/202).

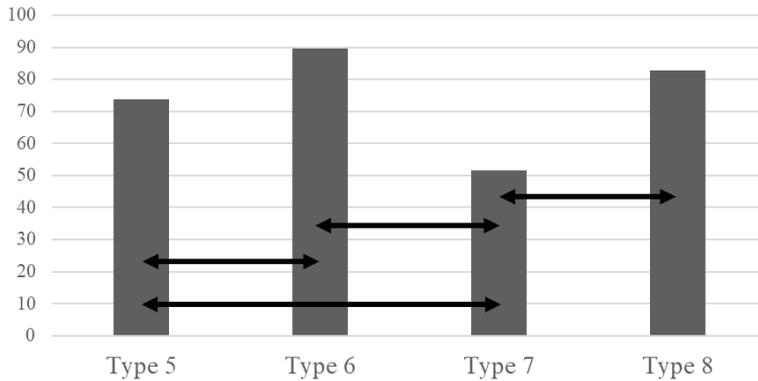
Repeated measures ANOVAs revealed significant effects of sentence types ( $F(3, 297) = 122.370, p < .001$ ). Multiple comparisons showed that the correct percentage of Type 5 was significantly lower than that of Type 6. The correct percentage of Type 7 was also significantly lower than that of Type 8. This result supports Hypothesis 4, and JLEs are more likely to overpassivise unaccusatives when there is a conceptualizable agent in the discourse than when there is no agent.

On the other hand, the correct percentage of Type 5 was significantly higher than that of Type 7, and there was no significant difference between the percentage of Type 6 and that of Type 8. This is the opposite result from the prediction of Hypothesis 5. At the same time, from the result that the correct percentage of Type 5 was not the lowest of the four, Hypothesis 6 was rejected. Thus, these findings show that the degree of verb telicity has little influence on the overpassivisation of unaccusatives.

**Table 2**

*Correct Percentage of the Four Types in Experiment 2*

Sentence Type	Mean	SD
Type 5	73.8%	0.654
Type 6	89.6%	0.452
Type 7	51.5%	0.789
Type 8	82.7%	0.570

**Figure 2***Correct Percentage of the Four Types in Experiment 2*

The result that the existence of conceptualizable agents had influence on overpassivisation of unaccusatives was consistent with the results of Study 1 and Ju (2000). In Study 1, we found that the existence of conceptualizable agents is a weaker factor than the animacy of subjects. But, Study 2 has revealed that, as long as the subject is an immovable inanimate, a conceptualizable agent can affect overpassivisation regardless of the degree of verb telicity.

Shirahata et al. (2020) argued that the degree of telicity becomes influential only when the subject is an immovable inanimate. However, the result of Study 2 indicates that telicity does not affect overpassivisation even though the subject is an immovable inanimate. Then, a new question arises; why was the influence of telicity observed in Shirahata et al. (2020), while it was not in the present study? We assume that the cause will be in the sentences of the two verbs, *disappear* and *arrive*, whose correct percentages were especially lower than those of *appear* and *belong*, in Shirahata et al. (2020). See the two sentences used in their experiment.

(24)

a. *disappear*:

A magician stared at a big ball in the glass box. Then,  
The big ball (disappeared/was disappeared) from the box.

b. *arrive*:

Taro opened the post box of his house and found a letter from America.  
The letter from America (arrived/was arrived) at Taro's house.

(Shirahata et al., 2020, p. 48)

As a problem of their experiment which we mentioned before, they did not take conceptualizable agents into account. Thus, as in (24a), the DP ‘a magician’ can be recognized as an agent and affect overpassivisation. In fact, according to Shirahata et al. (2020), the average score of *disappear* (0.47) was significantly lower than that of *appear* (0.62) although they were categorized into the same group ‘change of state’ of the ASH. If the degree of telicity affects overpassivisation of unaccusatives, the rate of the errors of the two verbs will not differ.

We believe that the same thing can account for the low average score of *arrive* (0.33). Though there was not an explicit agent in the discourse of *arrive* as in (24b), we can imagine an agent implicitly. More specifically, the incident ‘the letter arrived’ always involves some agents; there should be a person who wrote the letter and one like a postman who brought the letter to the post box. Thus, even if JLEs do not recognize an explicit agent, they are likely to overpassivise unaccusatives in a context where they can imagine an implicit agent like (24b). Therefore, given the results of our Study 2, we assume that the overpassivisation of unaccusatives in Shirahata et al.’s (2020) study can simply be affected by such an agent, not by the verb telicity.

We have examined the three factors, the animacy of subjects, the existence of conceptualizable agents, and the degree of verb telicity. As another possible factor causing overpassivisation of unaccusatives, a morphological property of L1 has been mentioned in some studies (Montrul, 2000 and 2001; Kondo, 2009). Also, Szcześniak (2020) examined L2 learners’ knowledge of unaccusatives from a perspective of association between unaccusatives and subjects. Moreover, since our experiments were conducted only with low-intermediate level learners, it is unclear whether learners’ English proficiency affects overpassivisation. In the present research, we removed these factors since it would be too complicated otherwise. Thus, in our future studies, we will examine how strongly these factors mentioned above affect overpassivisation.

## Conclusion

We have investigated which factor most strongly affects overpassivisation of unaccusatives: the animacy of subjects, the existence of conceptualizable agents, or the degree of verb telicity. From the results of Study 1, we have found that the animacy of subject can cause overpassivisation, whether conceptualizable agents exist or not. On the other hand, we have also discovered that the existence of conceptualizable agents becomes influential only when the subject is an immovable inanimate. These results suggest that the animacy of subjects is a stronger factor causing overpassivisation than the existence of conceptualiz-

able agents. Study 2 has revealed that the existence of conceptualizable agents can affect overpassivisation of unaccusatives regardless of the degree of verb telicity. However, we have also clarified that telicity of unaccusatives has little influence on overpassivisation of unaccusatives. This result contradicts Yusa (2003) and Shirahata et al. (2022).

Based on these results, we can conclude that the animacy of subjects is the most influential factor of the three, and the existence of conceptualizable agents is the second most influential, but the telicity of unaccusatives does not affect the overpassivisation. In the present study, we removed other factors like a morphological property of L1 Japanese, association between unaccusatives and subjects, and learners' English proficiency since it would be too complicated. But, in our future studies, we will investigate how strongly these factors mentioned above affect overpassivisation.

## References

- Balcom, P. (1997). Why is this happened? Passive morphology and unaccusativity. *Second Language Research*, 13, 1–9.
- Hirakawa, M. (1995). L2 acquisition of English unaccusative construction. In D. MacLaughlin & S. McEwen (Eds.), *Proceedings of the 19th Boston University Conference on Language Development* (pp. 291–302). Cascadilla Press.
- Hirakawa, M. (2006). 'Passive' unaccusative errors in L2 English revised. In R. Slabakova, S. Montrul, & P. Prevost (Eds.), *Inquiries in linguistic development*. John Benjamins.
- Ju, M. K. (2000). Overpassivization errors by second language learners. *Studies in Second Language Acquisition*, 22, 85–111.
- Kondo, T. (2009). *Argument Structure Morphosyntactic Links in the Second Language English of Adult Speakers* (Doctoral dissertation). University of Essex, UK.
- Kondo, T. (2019). The acquisition of intransitive verbs by second language learners. In T. Shirahata & K. Suda (Eds.), *Application of language acquisition research (Monograph series 3 for second language research)* (pp. 31–68). Kuroshio Publishers (in Japanese).
- Kondo, T., & Shirahata, T. (2015). The effects of explicit instruction on intransitive verb structures in L2 English classrooms. *Annual Review of English Language Education in Japan (ARELE)*, 26, 93–108.
- Kondo, T., & Shirahata, T. (2018). Explicit instruction on English verb structures in L2 classrooms. In I. Walker, D. K. G. Chan, M. Nagami, & C. Bourguignon (Eds.), *New perspectives on the development of key competencies in foreign language education* (pp. 157–179). Mouton De Gruyter.
- Kondo, T., Otaki, A., Suda, K., & Shirahata, T. (2015a, June). Japanese learners' usage of *be* + *-en* Form with English unaccusative verbs. Talk at the Japan Second Language Association 15th Annual Conference. Hiroshima University, Japan.
- Kondo, T., Otaki, A., Suda, K., & Shirahata, T. (2015b, July). Animate and inanimate contrast in the acquisition of unaccusative verbs. Talk at the 17th Annual International Conference the Japanese Society for Language Science. Beppu City, Japan.

- Kondo, T., Otaki, A., Suda, K., & Shirahata, T. (2016). Occurrences of unaccusative verbs in English textbooks and their acquisition. *Journal of the CELES*, 45, 53–60.
- Kondo, T., Shirahata, T., Suda, K., Ogawa, M., & Yokota, H. (2020). Effects of explicit instruction on intransitive and transitive verbs in second language English: with a special focus on non-instructed verbs. *Annual Review of English Language Education in Japan (ARELE)*, 31, 81–96.
- Montrul, S. (2000). Transitivity alternations in L2 Acquisition: Toward a modular view of transfer. *Studies in Second Language Acquisition*, 22, 229–273.
- Montrul, S. (2001). First language-constrained variability in the second language acquisition of argument-structure-changing morphology with causative verbs. *Second Language Research*, 17, 144–194.
- Oshita, H. (1997). “*The accusative trap*”: *L2 acquisition of English intransitive verbs* (Doctoral dissertation) University of Southern California, Los Angeles.
- Oshita, H. (2000). What is happened may not be what appears to be happening: A corpus-study of ‘passive’ unaccusatives in L2 English. *Second Language Research*, 16, 293–324.
- Perlmutter, D. M. (1978). Impersonal passives and the Unaccusative Hypothesis. Proceedings of the 4th Annual Meeting of the Berkeley Linguistics Society. UC Berkeley. pp. 157–189.
- Sato, K. (2015). *Why does overpassivisation of unaccusative verbs happen?—focus on \*An accident was happened*. Keisuisya (in Japanese).
- Shirahata, T., Kondo, T., Suda, K., Otaki, A., Ogawa, M., & Yokota, H. (2019). The learning and teaching of inanimate subjects: The case of Japanese learners of English. *Annual Review of English Language Education in Japan (ARELE)*, 30, 1–16.
- Shirahata, T., Kondo, T., Ogawa, M., Suda, K., Yokota, H. & Otaki, A. (2020). A study of overpassivisation of English unaccusatives by Japanese learners of English. In T. Shirahata & K. Suda (Eds.), *The spillover effect of second language acquisition research—from core grammar to L2 utterance (Monograph series 4 for second language research)* (pp. 31–55). Kuroshio Publishers (in Japanese).
- Sorace, A. (2000). Gradients in auxiliary selection with Intransitive Verbs. *Language*, 76, 859–890.
- Szcześniak, K. (2022). Memory of formulaic sequences in L2 retention of chunks featuring inchoative uses of unaccusative verbs. *LINGUISTICA SILESIANA*, 43, 43–64.
- Yip, V. (1995). *Interlanguage and learnability: From Chinese to English*. John Benjamins.
- Yusa, N. (2003). ‘Passive’ Unaccusatives in L2 Acquisition. In P. M. Clancy (Ed.), *Japanese/Korean linguistics, 11* (pp. 246–259). CLSI.
- Zobl, H. (1989). Canonical typological structures and ergativity in English L2 acquisition. In S. Gass & J. Schachter (Eds.), *Linguistic perspectives on second language acquisition* (pp. 203–221). Cambridge University Press.

Hiromu Okamura, Tomohiko Shirahata

## **Faktoren für die Überpassivierung von unakkusativischen Verben bei japanischen Englischlernern**

### **Zusammenfassung**

Es wurde berichtet, dass Lerner von Englisch als Zweitsprache (L2), einschließlich japanischer Englischlerner (JLE), unakkusativische Verben überpassivieren, obwohl es sich um eine Art intransitiver Verben handelt. Um dieses Phänomen zu erfassen, wurden einige Annahmen vorgeschlagen. Bisher ist jedoch unklar, welche Annahme die Überpassivierung von unakkusativischen Verben am effektivsten zu erklären vermag. In der vorliegenden Arbeit soll daher untersucht werden, welcher der drei Faktoren – Belebtheit der Subjekte, Vorhandensein konzeptualisierbarer Agens bzw. Telizität der Verben – die Überpassivierung von unakkusativischen Verben durch japanische Englischlerner (JLE) am stärksten beeinflusst. Um das Problem zu erforschen, wurden zwei Experimente durchgeführt. Studie 1 wurde mit 100 Universitätsstudenten (JLE) durchgeführt, um den Effekt der Belebtheit von Subjekten mit dem der konzeptualisierbaren Agens zu vergleichen. Als Ergebnis wurde festgestellt, dass die Belebtheit der Subjekte die Überpassivierung von unakkusativischen Verben stärker beeinflusst als das Vorhandensein konzeptualisierbarer Agens. Studie 2 wurde mit 101 Universitätsstudenten (JLE) durchgeführt, um zu untersuchen, welcher der beiden Faktoren – das Vorhandensein konzeptualisierbarer Agens oder der Grad der Telizität – eine stärkere Einwirkung hat. Die Ergebnisse zeigten, dass der erste einen Einfluss auf die Überpassivierung von unakkusativischen Verben hat, der andere jedoch weniger. Aus den beiden Experimenten lässt sich schließen, dass die Belebtheit des Subjekts die stärkste Einwirkung unter den drei Faktoren hat. Die Telizität der Verben hat dagegen kaum Einfluss auf vorkommende Fehler.

*Schlüsselwörter:* Überpassivierung, unakkusativische Verben, Belebtheit, konzeptualisierbare Agens, Telizität



**Thiri Soe**

Graduate School of International Cultural Studies  
Tohoku University, Japan

 <https://orcid.org/0009-0006-6161-696X>

## **Investigating Critical Thinking in ELT Textbooks: A Systematic Literature Review of Textbook Evaluation Studies**

### **Abstract**

The pivotal role of critical thinking and its integration in English language education are synthesised through systematic literature reviews and meta-analysis studies in the areas of pedagogical interventions, teaching methods, and assessment methods. However, there is little evidence of systematic literature review on the integration of critical thinking into English language teaching materials such as textbooks. In this study, 41 empirical textbook evaluation studies published between 2010 and 2021 were reviewed from critical thinking perspective. The purpose of this study is to provide an overview of English textbook evaluations on critical thinking integration, to identify critical thinking frameworks applied, and to look for possible research gaps among the studies. An extensive literature search was conducted by applying Xiao and Watson's (2017) eight-step systematic literature review method. From a collection of 41 empirical studies, this review focuses on such four aspects as applied critical thinking frameworks, findings, recommendations, and the language levels and content areas of the textbooks studied. The results revealed that (1) Bloom's Taxonomy and Bloom's Revised Taxonomy were most frequently applied critical thinking frameworks, (2) integration of higher-order thinking skills in textbooks was less common than lower-order thinking skills, (3) textbook assessment occurred primarily at the middle and high school levels more than at other levels, and (4) methodological considerations regarding the reliability and validity of the coding process in textbook content analysis received little attention. This study contributes a synthesised literature background of English textbook evaluation with recommendations for methodological improvements in future studies.

*Keywords:* systematic literature review, critical thinking, ELT textbook evaluation, higher order thinking skills, lower order thinking skills

## **Critical Thinking Integration in ELT Textbooks**

Mastery of critical thinking skills is a premium goal in education sector. Success in a profession is directly related to a person's superior critical thinking skills (Cottrell, 2017). In the age of technology where information is readily available at one mouse's click, educational goals are directed from rote learning toward enhancing transferable skills like critical thinking across subject specific contexts (Puig et al., 2019). As such, teachers, researchers, and education specialists have focused on improving learners' critical thinking skills through specialised course content that incorporates critical thinking activities. Critical thinking across curricula in subject-specific contexts is highly recommended for the development of learners in accordance with educational objectives (Ennis, 2018). Therefore, materials and activities used in language classrooms have become key factors in enhancing language learners' critical thinking skills (Shirkhani & Fahim, 2011). A teacher's choice of teaching methods and materials also directly affects thinking practices and training students receive (Thompson, 2011). Examining the degree of emphasis on critical thinking in ELT textbooks uncovers potential contribution of the textbooks to language learning (Birjandi & Alizadeh, 2013; Solihati & Hikmat, 2018). Thus, textbook analysts and pedagogues of English language teaching are increasingly examining the integration of critical thinking skills in language teaching materials. Despite the growing interest in critical thinking skills and activities developed in ELT materials, there is a lack of systematic literature review in ELT textbook evaluation studies from critical thinking perspective. Research syntheses are necessary since they summarise the breadth and depth of the existing literature and can also indicate research gaps that raise new research questions for further methodologically and theoretically improved studies. For these reasons, this systematic literature review study attempts to survey the existing literature on the analysis and evaluation of ELT textbooks critical thinking integration perspective.

### **Literature Review**

Critical thinking is known to have a variety of proposed definitions. In the course of its evolution, scholars, philosophers, psychologists, and educationalists have each attempted to define the concept of critical thinking skills in their own way, and there are overlapping characteristics among these definitions (Nilson, 2021). There is neither a universal agreed-upon set of critical think-

ing skills nor an educational approach for implementing them. The scope and depth of critical thinking in literature is so massive that it cannot be easily grasped as a specific idea that can be applied in any discipline in any context. In 2013, Moore conducted a qualitative study in three academic disciplines such as philosophy, history, and literary and cultural studies to define the term critical thinking as the basis for understanding and application by scholars. These disciplines were intentionally chosen because of their close relationship to education. As a result of the interview responses, Moore concluded four main concepts of critical thinking. These concepts are (1) the ability to judge ideas, (2) a skeptical and provisional view of knowledge, (3) originality of ideas or modest contribution to a body of knowledge, and (4) careful and sensitive reading of texts or input of information. In addition, other concepts peripheral to critical thinking included rational thinking, adopting an ethical and activist stance, and self-reflection.

In 2018, Ennis, a prominent figure in the field of critical thinking education, said that scholars in different disciplines define critical thinking differently depending on their stance in their field of expertise. However, as a way to cut the same pie from different angles, their definitions and assumptions are all broadly similar. Ennis clarified the concept of critical thinking by summarising all of his research over the past 30 years in one expression as, “rational reflective thinking focused on deciding what to believe and do (p. 166).” In an attempt of helping students develop their critical thinking skills, Halpern (1999, p. 70) elaborated on the concept of critical thinking as follows.

Critical thinking refers to the use of cognitive skills or strategies that increase the probability of a desirable outcome. Critical thinking is purposeful, reasoned, and goal-directed. It is the kind of thinking involved in solving problems, formulating inferences, calculating likelihoods, and making decisions.

Halpern’s core definition of critical thinking skills includes five key elements: (1) reasoning skills, (2) hypothesis testing, (3) analysis of arguments, (4) analysis of possibilities and uncertainties, and (5) decision making and problem solving. As discussed so far, several characteristics of the concept of critical thinking have been proposed by different authors from different disciplines. Thus, the practice of critical thinking in the process of language acquisition implies a rational and open-minded acceptance of language input and output. Moreover, effective communication and critical thinking skills are rated as the most important intellectual skills demanded by employers in the 21st century job industry. In an attempt to enquire into the interpretation of the meaning of critical thinking and language competency, Jones et al. (1995) surveyed a total of 600 teachers, employers, and policy makers about their definitions

of writing, speaking, listening, and critical thinking and their expectations for how college graduates should develop these skills to become effective employees at work and informed citizens in society. The survey result stated that for language input, learners should be able to identify biased or misleading information. Learners require the ability to see through indirect persuasion, information and facts that are irrelevant to an argument, and information and questions that evoke speakers' preferred responses. For language output through writing, learners should have knowledge of audience awareness (considering potential readers of their writing), purposeful writing, and text organisation. For language output through speaking, learners should be able to exchange information appropriately, manage conversations and group communication. Behar-Horenstein and Niu (2011) also supported this idea by explaining that critical thinkers must not only examine and evaluate information beforehand in their critical thinking processes, but must also take responsibility for challenging and monitoring their own thinking and arguments. These key concepts of critical thinking in the language learning aspect are also consistent with Moore and Parker (2009), who delved deeply into the concept of critical thinking and developed a long list of 21 characteristics as a definition of critical thinking. According to their definition, which summarises these 21 characteristics, critical thinking can be divided into two main categories: using logical thinking to evaluate the credibility of information and using logical thinking to construct one's arguments. Therefore, we posit that from the domain of the language learning, the concept of critical thinking skill can be boiled down into two aspects: language learners' application of rational thinking in language input and output processing, in other words, receptive and productive language processing.

### **The Roles of Critical Thinking in English Language Education**

The role of critical thinking is essential in English language learning for a number of reasons. Increasing access to the internet and information through digital media reinforces the urgent need for learners' critical thinking development. Dummett and Hughes (2019) posited that through critical thinking activities, language learners acquire four different types of literacy: information literacy, media literacy, visual literacy, and intercultural literacy, all of which foster learners' effective and constructive communication. For authentic communication, it is not enough for learners to memorise and repeat chunks of a given language examples. To communicate effectively through the target language, learners must analytically investigate given information, personalise it, and solve problems, all of which require critical thinking skills (Hughes, 2014). Through critical thinking integrated activities and tasks in language classes, learners are able to perceive and process information critically. Long-

term memory is also enhanced when learners deeply and critically explore language input through critical thinking activities (Dummett, 2016). Learners' learning memories are categorised as implicit and explicit memories, also called procedural and declarative memories, respectively. In first language acquisition, learners acquire their native language through implicit memory, which is unconscious and instinctive in nature. In foreign language acquisition, however, learners acquire the target language through explicit memory, or conscious learning. Therefore, by encouraging learners to engage in critical thinking tasks and practices over time, they move from a deep critical exploration of language input to a more instinctive sense of the target language being experienced. As a result, critical thinking activities reinforce learners' long-term memory in the target language acquisition process and move learners toward instinctive and implicit memory (Dummett & Hughes, 2019).

Not only are critical thinking skills and English language proficiency parallel skills needed to meet the demands of the 21st century employment industry, but there is a strong positive correlation between the two. According to the literature on the role of critical thinking in ELT education, pedagogical approaches for language learning that incorporate critical thinking have been widely discussed, and its role continues to expand in English language teaching programs. Integration of critical thinking in English language teaching is established in two areas of interest: language teacher's instructional methods that incorporate critical thinking and development of language teaching materials that reinforce critical thinking (Shirkhani & Fahim, 2011). In the aspect of instructional approaches, specialists and researchers provided teaching instructions and methods to help learners engage in critical thinking fostering activities in language classrooms. For example, Hughes (2014) presented a list of 20 ELT classroom activities and demonstrated how different problem solving activities can be integrated at different language proficiency levels. To help college students develop the ability to evaluate the reliability of information intake, Halpern (2002) set up a four-part critical thinking instructional model consisting of critical thinking dispositions, critical thinking skills approach, structural training, and metacognitive monitoring. Alnofaie (2013) proposed a framework for incorporating critical thinking activities into four language skills. Thus, English language teachers have become aware of effective teaching-learning techniques that enhance language learners' thinking skills and language skills simultaneously. English language teachers' classroom experiments and research findings provide guidance and effective recommendations on classroom activities, and areas that teachers should focus on to enhance learning opportunities. Notable examples include Lin's (2018) critical thinking infusion approach in Chinese high school learners' English composition, which reportedly improved both critical thinking skills and writing performance. Golpour (2014) found that Iranian EFL learners with high critical thinking

skills showed significant performance in various writing modes, including expository and argumentative writing. In Yang and Gamble's (2013) classroom experiment using critical thinking-enhanced activities such as debate and peer critique, the experimental group made significant gains in reading and listening proficiency compared to the control group that received no critical thinking integration instruction. Fahim et al. (2012) found in their study that critical thinking training had a facilitative effect on reading performance of high and low proficiency EFL students. Sanavi and Tarihat (2014) showed that explicitly introducing critical thinking skills improved learners' speaking skills compared to learners who did not receive critical thinking reinforcement.

From the perspective of language instruction, as discussed above, experimental studies in English classrooms have emphasised that incorporating critical thinking skills into classroom interventions improves language proficiency in all four skills. However, the development of language learners' critical thinking from the area of language teaching materials has not yet been well explored, especially in terms of a systematic literature review. Furthermore, there are still no fixed parameters or predominantly established working models applied to the evaluation of language materials. Therefore, the purposes of this study are to scrutinize the existing literature on ELT textbook evaluation research from critical thinking perspective, to identify widely used critical thinking frameworks in the evaluation, and to identify possible research gaps in the area.

## **Evaluating Critical Thinking Incorporation in ELT Textbooks**

Throughout our discussion of critical thinking in language teaching materials, we refer interchangeably to ELT course books, ELT textbooks, English language textbooks, and English language teaching materials as the same concept in the context of this study. We also used the two terms of *textbook evaluation* and *textbook analysis* interchangeably for examining and evaluating textbooks' contents. Research on the evaluation of English language teaching materials can be summarised into two dimensions: (1) how to evaluate and (2) what to evaluate. There are three main stages of evaluation methods for teaching materials: pre-use evaluation, in-use evaluation, and post-use evaluation (McGrath, 2002). These stages differ with respect to the time they are conducted and the objectives of the evaluation. Pre-use evaluation is conducted in selecting teaching materials before a tentative textbook is actually determined for use. Evaluators like teachers and concerned professionals observe the potential that their tentatively chosen textbooks can promise. In the in-use evaluation stage, evaluators actually use the textbook while observing whether it fits the actual classroom situation and meets its claimed potential. The post-use evaluation phase is conducted after a textbook has been in use for a period

of time to help teachers figure out what of the possibilities the textbook initially promised actually worked and what did not. This third stage can involve learners' performance assessment. It is the stage where teachers reflect upon the difficulties and gaps that need to be supplemented and adapted to learners' needs and learning situations. Ellis (1997) called pre-use evaluation a *predictive evaluation* (examining course materials before using) and post-use evaluation as a *retrospective evaluation* (examining materials after using). For the predictive evaluation, language teachers either evaluate textbooks by themselves or rely on other expert reviewers. Post-use or retrospective evaluation poses some challenges and difficulties for teachers to carry out because they are time-consuming as empirical case studies.

In this section, we will only illuminate the fact that there is no place yet for critical thinking in the evaluative criteria established so far in the literature. We cannot cover the complete literature background regarding the criteria and parameters used for ELT textbook evaluation. Here are a few examples among the huge number of ELT textbook evaluation parameters or checklists theorised by experts in the field. Williams (1983) proposed an evaluative scheme that consists of four criteria: up-to-date methodology, guidance for non-native teachers, meeting the learning goals of language learners, and relevance to the socio-cultural environment. Sheldon (1988) proposed a 17-item-list of common core factors for ELT textbook assessment, which mainly focuses on practicability, flexibility, accessibility, and layout of teaching materials. McDonough et al. (2013) proposed external and internal evaluations for the analysis of textbook contents. External evaluation looks into the contents and organisation of textbooks and other general factors such as availability of teachers' books, cultural and gender biases, layout and presentation, etc. If the results of the external evaluation indicate that the subject material is appropriate, the evaluation proceeds to internal evaluation, where the sequence of materials and exercises, the relationship between exercises and tests, and the suitability of the exercises and texts for the learning style are examined in detail. In 2015, Brown and Lee proposed a list of criteria to evaluate an ELT textbook. Their criteria are focused on the perspectives of program/course, teaching approaches, language skills, the practicality of the materials, and the availability of the supplementary materials. These aforementioned examples show how evaluation criteria all focus on the authenticity and relatedness between the teaching materials and learners' real life experiences. In other words, they mean that appropriate tasks and activities that activate learners' engagement and practical application to real life situations are highly valued in ELT material development. Learners' language practices and language input should be closely related to real-life problem solving and experiences that require learners to apply critical thinking.

Even though researchers and experts in the field of materials development have developed a rich literature on evaluation criteria for language materials,

their methods and applications are specific to each language learning situation and cannot be transferred directly to other situations without modifications (Tomlinson, 2012). However, there is one basic principle that is common to all learning situations and to all learning materials developed in different teaching and learning situations. The principle is that applied language instructions and language teaching materials should meet the actual needs and goals of a particular group of language learners in a particular context (McDonough et al., 2013). As supported by classroom experiments and case studies on language teaching methods that incorporate critical thinking, the development of critical thinking skills is essential in the process of language acquisition, regardless of different learning goals or learning situations. Still, there are currently no parameters that explicitly target the inclusion of critical thinking activities in ELT textbooks. Littlejohn (2011) also emphasised the need for evaluation criteria that critically question about the status of learning autonomy, engagement in problem-solving tasks, and emphasis on learner-centred approaches. Prior to the development of such assessment criteria, research on the integration of critical thinking into ELT curriculum is a foundation step in a future roadmap for developing assessment criteria for ELT textbooks from critical thinking perspective.

## **Materials and Methods**

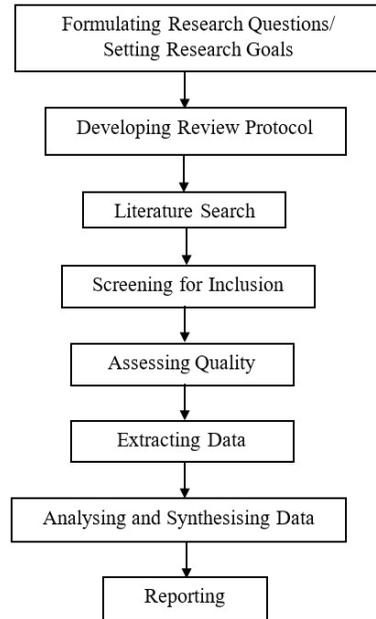
In our systematic literature review we applied Xiao and Watson's (2017) eight-step approach. Figure 1 depicts the flow of the eight-step review process, including key decisions and protocols made based on the availability of relevant literature and the purpose of the review.

### **Step 1: Formulating Research Questions or Setting Research Goals**

Critical thinking plays a prominent role in English language education worldwide. In addition, evaluation criteria for ELT textbooks have long been developed to enhance the quality and sustainability of the teaching materials. However, evaluating ELT textbooks and learning materials from critical thinking integration perspective is relatively new in its field. Consequently, there are no summarised and organised literature review studies on the topic so far.

**Figure 1**

*Flow of the Systematic Literature Review Process Adapted from Xiao and Watson (2017)*



It is worth examining the depth and scope of the existing literature, research findings, and applied critical thinking frameworks and methods in EFL textbook evaluation. In order to scrutinise the current state of ELT textbook evaluation studies from critical thinking integration aspect and research trends, and to identify possible research gaps in this domain of study, a systematic review of previous literature is necessary. For these reasons and objectives, this study is conducted by critically synthesising the previous empirical studies within the last eleven years. The parameters to be observed were determined by referring to Cooper's (1988) *focus* of the literature review. According to Cooper, the criterion called *focus* encompasses such factors as (1) research outcomes, (2) research methods, (3) theories, and (4) practices or applications. With reference to these four characteristics for the focus of literature review, we set the research objectives to examine the following five points in our textbook evaluation study:

- critical thinking frameworks used in the ELT textbook evaluation studies;
- results of the studies on the percentage of critical thinking integrated tasks in textbooks;
- levels and language areas of ELT textbooks being examined;
- validity and reliability of the studies, and;
- recommendations made by the studies.

## **Step 2: Developing Review Protocol**

This is the planning stage for the literature review process, including re-search objectives, questions, search strategies, inclusion criteria, data extraction, analysis, synthesis and interpretation, and writing a report. All of these will appear in a sequential process in the following stages. In this stage, limiting the scope of the study and inclusion criteria are the key decisions because they are the roadmaps of the study. The following three criteria are review protocols developed in order to limit the scope of the study:

1. The retrieved studies are conducted in EFL context.
2. The study period is from 2010 to 2021 (12 years).
3. The retrieved studies are conducted, focusing on critical thinking manifestation.

## **Step 3: Literature Search**

In the literature search, we used a purposive sample approach with keyword searching, forward searching, and backward searching, targeting the studies that focused on ELT textbook evaluation in EFL context. Databases used for literature search include Web of Science, ERIC, Elsevier, Scopus, Google Scholar, Science Direct, and PsychInfo. Keywords used in searching are English language textbook evaluation and critical thinking, critical thinking in language teaching materials, and critical thinking in ELT/EFL textbooks.

## **Step 4: Screening for Inclusion**

Step 4, the screening process is the preliminary screening stage of Step 5, which is the quality assessment. In this Step 4, the retrieved literature temporarily saved in a file was screened to determine whether or not to include it in the review. That appraisal was done by reviewing the abstracts of the studies and skimming through the entire article. Studies that did not follow the review protocol were excluded.

## **Step 5: Assessing Quality**

This step is a detailed evaluation of the content. The recruited studies were screened in more detail by reading the full texts of the studies. The decision to include the studies was made in reference to three criteria developed in the review protocol step. It is worth mentioning here that the inclusion of the studies was not restricted to peer-reviewed articles due to the little availability of the

studies and their role in EFL context. After excluding studies not relevant to the purpose and scope of the review and after removing two duplicate publications, a total of 41 empirical studies were obtained for data extraction, analysis, and interpretation.

### **Step 6 and Step 7: Extracting Data, Analysing and Synthesising**

Data extraction, analysing, and synthesising were performed, following Lune and Berg's (2017) three concurrent flows method: (1) data reduction, (2) data display, and (3) conclusion and verification. First, we developed coding themes or analytic categories based on the review protocol. Second, raw data were coded and transformed into interpretable themes. The raw data were first recorded in a Word file and then converted to an Excel sheet, organised by category to be analysed, and frequencies were counted. Data display is an inseparable process of data reduction. In that step, data were displayed in visualised forms for the analysis such as tables, figures, excel sheets, and tally sheets. From these visualisations, interrelated themes and patterns were discovered. Third, in the conclusion and verification step, the analytic conclusion was made through evaluation and decision making that occurred throughout the analysis. That analytic conclusion was verified by retracing the analytic steps of the data by the same researcher sometime after the first analysis had been done.

### **Step 8: Reporting**

The final stage is reporting the whole literature review work to the academic realm to disseminate the study such as this academic paper to be engaged with the concerned scholars in the field. This systematic literature review is reported following the PRISMA statement (Moher et al., 2009), a comprehensive checklist for reporting a systematic literature review.

## **Findings and Discussion**

This section presents the findings of the data derived from the systematic literature review that examines five aspects from the previous empirical studies such as (1) critical thinking frameworks used in the ELT textbook evaluation studies, (2) results of the primary studies on the percentage of critical thinking

tasks in textbooks, (3) language levels and language areas of ELT textbooks being examined, (4) validity and reliability of the studies, and (5) recommendations made by the studies.

### **Critical Thinking Frameworks Applied in the Studies**

The first central observation, as given in Table 1, is the application of critical thinking frameworks in the studies. There are eleven critical thinking frameworks applied by different numbers of studies: 15 studies applied Bloom's Taxonomy (Bloom et al., 1956) and 20 studies applied Bloom's Revised Taxonomy (Anderson et al., 2001) while the rest of the studies applied other critical thinking frameworks such as Facione's Critical Thinking Model (2011), Paul-Elder's Critical Thinking Model (Elder & Paul, 2012), Peterson's Critical Thinking Model (Peterson, 2008), Ilyas's Critical Thinking framework (2015), the survey method of questionnaires and interviews, and *Inference* and *Deduction* as essential sub-thinking skills. It should be noted that four studies are found to be applying other critical thinking frameworks while they were using Bloom's Taxonomy and Bloom's Revised Taxonomy in their analyses. For example, Akrong et al. (2021) used Bloom's Revised Taxonomy and Cummins's (1999) framework: cognitive, academic, and language dimensions. Shuyi and Renandya (2019) used Bloom's Revised Taxonomy and Webb's (2002) Domain of Knowledge Model, and Tabari and Tabari (2015) used Bloom's Taxonomy and Multiple Intelligence Model (Gardner, 1993, 1999). The actual number of reviewed studies is 41, but three of them used two critical thinking frameworks simultaneously. Thus, the total number of studies indicating the use of these all critical thinking frameworks is 44, as shown in Table 1. It should also be noted that there is an extreme imbalance in the frequency of the use of critical thinking frameworks. Each of the critical thinking frameworks other than Bloom's Taxonomy and Bloom's Revised Taxonomy was applied only once. The aim of the present literature review is not to discuss these less frequently applied frameworks. Therefore, in order to explore the central patterns and themes that emerge from the findings, we focused only on the studies that belong to the most frequently applied critical thinking frameworks, that is, Bloom's Taxonomy and Bloom's Revised Taxonomy. As a result, subsequent data extraction and analysis were conducted in studies using these two frameworks, the results of which are presented in the following section.

**Table 1***Critical Thinking Frameworks Applied as Evaluation Checklists*

Critical Thinking Frameworks	Frequency of Use
Bloom's Revised Taxonomy (2001)	20
Bloom's Taxonomy (1956)	15
Facione's Critical Thinking Model (2011)	1
Paul-Elder's Critical Thinking Model (2012)	1
Peterson's Critical Thinking Model (2008)	1
Ilyas's Critical Thinking Framework (2015)	1
Cummins's framework: cognitive, academic, and language dimensions (1999)	1
Webb's Domain of Knowledge Model (2002)	1
Gardner's Multiple Intelligence Model (1993; 1999)	1
Questionnaire and Interviews	1
Inference and Deduction	1
Total	44

**The Most Frequently Applied Critical Thinking Frameworks**

In seeking the critical thinking integration inside the textbooks, we limited our focus to the studies that used the two most common critical thinking frameworks, Bloom's Taxonomy and Bloom's Revised Taxonomy. It is not because the other critical thinking models and theories are less important and less credible, but because our study aims to draw a conclusive summary out of possible prominent themes and patterns found in the analyses. Therefore, we synthesised the results of 35 studies using these two taxonomies and summed the percentages obtained from each study. Another good reason to focus only on the two taxonomies is that they are well-established theoretical frameworks for curriculum development and learning goals in the education sector. Bloom's Taxonomy and Bloom's Revised Taxonomy were most frequently used critical thinking frameworks, with Bloom's Revised Taxonomy occupying 20 studies (49%) out of the total reviewed studies, Bloom's Taxonomy occupying 15 studies (36%) out of the total reviewed studies, resulting in a total of 49% of the studies. Figure 2 is the representation of the two taxonomies and overall percentages yielded after adding up the results of critical thinking skills in Bloom's Taxonomy and Bloom's Revised Taxonomy respectively. There are changes in the Revised Taxonomy from the original one; however, the clear classification between lower and higher order thinking skills, shown by a dotted line in Figure 2, remains the same in both taxonomies. Consequently, the studies

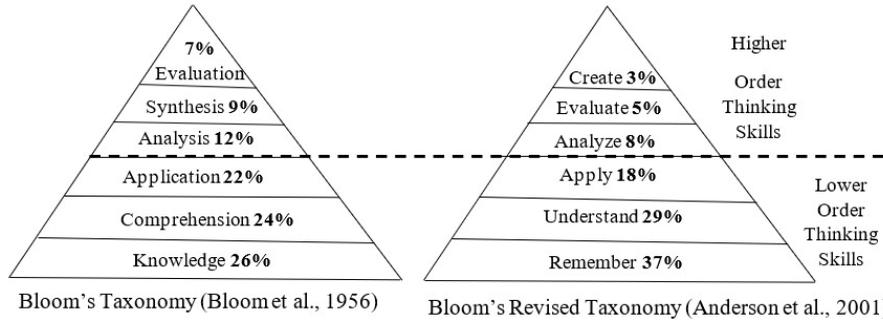
came up with the common categorisation of lower and higher order thinking skills in their findings. All 35 studies reached the common conclusion that the ratio of lower-order to higher-order thinking skills is unbalanced. Studies using Bloom's Taxonomy found that 28% of the content of the textbooks analysed activated higher-order thinking skills, with the remaining 72% being lower-order thinking skills. Studies using Bloom's Revised Taxonomy found that 16% of the content of the textbooks analysed activated higher-order thinking skills, with the remaining 84% being lower-order thinking skills. This situation is also emphasized in Krathwohl (2002), where less focus is placed on the higher order thinking skills of the six criteria, and the focus of educational practices and exercises is generally on the lower order thinking level such as knowledge and understanding.

When calculating the overall average percentage of critical thinking skills found within the textbooks, we had to omit some studies that did not analyse and present the integration of critical thinking activities by counting the number of times critical thinking activities were enhanced. For example, Birjandi and Alizadeh (2013) examined the integration of critical thinking skills on a Likert scale and thus we could not examine the average percentage from it. Therefore, this study, which evaluated textbooks using Bloom's Taxonomy, was omitted from the calculation of the overall percentage of all studies using Bloom's Taxonomy. Similarly, Jebbour (2019) did not provide a quantitative calculation of critical thinking activities in the textbooks they examined. The authors simply described which critical thinking skills were tapped into the textbooks and which were not. The reader cannot know what percentage of critical thinking activities were included from these studies. In addition, studies such as Razmjoo and Kazempourfard (2012) were not included in calculating averages of each thinking skill because the thinking skills are only presented cumulatively in two groups, the lower order thinking skills group and the higher order thinking skills group. Such cases occurred in two of the 15 studies using Bloom's Taxonomy and two of the 20 studies using Bloom's Revised Taxonomy. However, we assume that these studies were negligible in number compared to the overall number of studies and were unlikely to affect the overall percentage result. The average percentage of each thinking skill in both taxonomies was calculated from 31 studies (excluding four studies irrelevant to the overall calculation), and the results are presented in Figure 2.

For the information of the readers of this study, a brief historical background of these frequently applied taxonomies should be provided. The fundamental frameworks of the two taxonomies are based on the concept that the attainment of educational goals starts from simple to complex skills as in a hierarchical order of six learning skills in the pyramid structure. The first taxonomy, called Bloom's Taxonomy, was developed by Bloom and his colleagues, Engelhart, Furst, Hill, and Krathwohl from education and psychological domains.

**Figure 2**

*The Two Most Often Referenced Critical Thinking Frameworks: Bloom's Taxonomy and Bloom's Revised Taxonomy*



Their complete taxonomy has three major parts such as cognitive, affective, and psychomotor domains. The cognitive domain, which constitutes the six thinking levels, is most emphasised as critical thinking skills and most referred to in curriculum development to establish educational goals. Bloom's Taxonomy serves as a framework for educational objectives and as a standard assessment checklist for evaluating learner learning outcomes. Their six major categories of the cognitive domain start with an essential requirement in learning, *knowledge*. After acquiring the *knowledge*, learners are expected to *comprehend* the subject under study. The next improved level after *knowledge* and *comprehension* is *application* of the knowledge learned. These three basic levels are classified as lower order thinking skills. *Analysing*, *synthesising*, and *evaluating* the existing body of knowledge are three higher order thinking skills that learners are expected to attain as an ultimate learning goal.

In 2001, forty-five years after Bloom's Taxonomy, Anderson and colleagues modified the original taxonomy into a new thinking taxonomy called the Revised Taxonomy or Bloom's Revised Taxonomy. The primary distinction between the revised taxonomy and the original taxonomy is that the noun *Knowledge* is replaced by the verb *Remember*. Krathwohl (2002) explained why the knowledge of any subject matter the learners pursue should be targeted by the action verb that in fact will be the goal to achieve. For example, for an economics student, if the required content knowledge is the law of supply and demand, then the goal is to be able to remember that knowledge. In other words, the reason for modifying the *Knowledge* into *Remember* is to clarify that learners' acquired skill is not the knowledge itself, but the ability or effort to remember or to recall that knowledge.

There are two significant changes between the two taxonomies. The first change in revised taxonomy from the original taxonomy is converting the classification levels of nouns into verb categories. The second change is that

*Synthesis* skill from the original taxonomy is switched with evaluation level and renamed *Create* in the revised taxonomy. It is not surprising to find out in this review that the two taxonomies are the most frequently used frameworks. It is because of their prominent contribution to the classification of curricular objectives and learning programs since their development hitherto. More importantly, researchers' belief and application of theoretical frameworks is tied to their specialised disciplinary orientation as noted by Merriam and Tisdell (2016). The two taxonomies—which were originally developed for assessing curriculum developments and educational outcomes—become the lenses that textbook evaluation analysts used as critical thinking frameworks to evaluate the potential of the textbooks. However, none of the 15 studies using Bloom's Taxonomy or the 20 studies using the revised taxonomy mentioned any detailed explanation for their choice of one taxonomy over the other, or why the taxonomy they applied was particularly suited to their study compared to their counterpart. Such a lack of clarification on the appropriate choice of the taxonomic framework should be paid attention to in future studies. A critical comparison of the two taxonomies with an evidence-based discussion of teaching materials evaluation will be beneficial.

### Targeted Language Levels and Content of the Evaluated Textbooks

Knowing the target language levels of the textbooks allows future researchers to pay more attention to less emphasised areas and compare results across various target language levels. Table 2 shows the number of studies and the language levels of the textbooks being examined.

**Table 2**  
*Levels of Target Language in Textbooks Analysed*

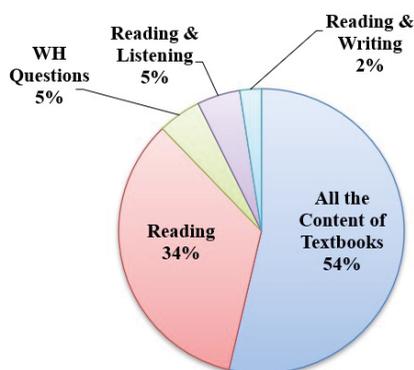
Levels of the Target Language	Junior High School & High School	University	Comparison of High School & University	Not Given	Others (Multi-level)	Total
Number of Studies	22	7	2	4	6	41
Percentages	53%	17%	5%	10%	15%	100%

As shown in Table 2, junior high school and high school levels receive the most attention among the studies. There are relatively fewer studies at the university level than at the high school level. Two textbook evaluation studies (Freatat & Smadi, 2014; Riazi & Mosalanejad, 2010) attempted to compare the rate of integration of critical thinking at the high school and college level

for the purpose of seeking transfer patterns and consistency of critical thinking activities embedded in the textbooks. Such studies contribute to the design and evaluation of curricula at different language levels. Four textbook evaluation studies (Birjandi & Alizadeh, 2013; Razmjoo & Kazempourfard, 2012; Sahragard & Alavi, 2016; Ulum, 2016) did not mention sufficient information for the level of textbooks. More complete information on the use of materials and methods would have been helpful for researchers of systematic literature review studies to calculate and integrate the actual existing literature.

**Figure 3**

*Units of Analysis in the textbook Evaluation Studies*



In our analysis, in addition to the level of textbooks, we examined the language domains of the textbooks being evaluated. Figure 3 shows the language domains of the textbooks analysed for critical thinking integration. There are a total of 125 textbooks distributed among 41 studies, and the different textbooks have different language foci and orientations. Therefore, the percentage of the language area for evaluation is influenced not only by the theoretical standpoint and choice of the researchers but simply by the fact that different textbooks have different language areas being focused on such as reading, writing, vocabulary, etc. To this end, it is safe to conclude by looking at Figure 3 that the majority of the studies (54%) indiscriminately analysed the entire content of the textbooks under study. Apart from that, reading received more attention (34%) for critical thinking integration analysis than other language skills. Despite a small percentage, it is noteworthy that two studies (Freahat & Smadi, 2014; Ighbaria, 2013) focused on “WH-questions” of the given texts with the theoretical standpoint that questions motivate learners to solve problems and activate learners’ critical thinking skills. Asking questions plays a central role in the reinforcement of critical thinking in language learning exercises. Questions asked in a teaching context can be classified as lower level ques-

tions where students need to gather and recall the information, and higher level questions where students need to apply such activities as analysing, synthesising, and evaluating (Nappi, 2017). Teachers' effective use of questions in language exercises can direct students to more strategic and critical thinking.

### **Validity and Reliability of the Studies**

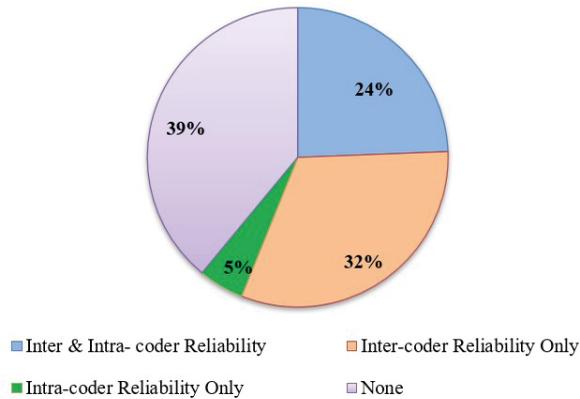
Methodologically valid and reliable research or *trustworthiness* of a study is a broad concept, referring to such factors as *credibility*, *transferability*, *dependability*, and *confirmability* (Shenton, 2004). It ultimately represents how a particular study is conducted methodologically and ethically, so that readers of the study can confidently apply its findings and recommendations to real world problems and can establish social policy in relevant contexts. In this study, we put our focus on two aspects of the coding processes, *validity* and *reliability*, which are critical aspects in the qualitative content analysis research design that the textbook analysts applied. At this point, we refer to Potter and Levine-Donnerstein's (1999) concept of *validity* and *reliability* in content analysis. They underscored the two-step process in content analysis: developing a coding scheme congruent with the theoretical framework of the study, and assessing coders' correct decision making according to certain guidelines and standards. Transparent and elaborative descriptions of the coding process of qualitative content analyses can show its methodological integrity and increase its replicability. This essential characteristic was not presented as favoured in all the reviewed studies. Most researchers did not provide transparent and thorough information about the development of the coding scheme and did not assess how the co-coders make decisions when categorizing critical thinking skills from the analysed textbook content.

In terms of validity, when we examined how many of the studies mentioned the validity in the analysis processes, we found only seven studies (17%) did so. Those seven studies reported that researchers consulted with the experts on their established operational definitions for each critical thinking skill standard and on the appropriateness of the choice of their measurement instrument or critical thinking framework used in their textbook analyses. For example, in, Assaly and Smadi (2015) and in Igharia (2013), the researchers constructed a list of critical thinking skills containing definitions and concepts of six critical thinking levels from Bloom's taxonomy and presented it to a committee of experts from their specialised field prior to analysis to validate the operational definition of the research instrument. Similarly, Abdelrahman (2014), by consulting a jury of seven experts, validated the suitability of using Bloom's revised taxonomy as a research indicator in assessing critical thinking integration in the analysed textbooks. It is important to have a mutually understood

operational definitions among evaluators on evaluative criteria of the critical thinking skills used to be assessed. This is underscored by Tomlinson and Masuhara (2004) who proposed a list of five questions in a materials evaluation checklist. One of these five questions was to check whether the evaluation criteria were reliable in the sense that other evaluators would define and interpret them in the same way. Although the authors of those studies noted the validity and appropriateness of their choice of research instruments, none of them elaborated on the process of consultation with experts or provided transparent explanations or arguments as to why they chose one particular critical thinking framework over other frameworks.

**Figure 4**

*Instances of Reliability Assessment*



Assessing the reliability of the content analysis is to look at the accuracy and consistency of the coding analysis in the analysis process. In other words, it means finding similar coding judgements about the same content between or among the coders. Figure 4 shows the number of studies that assessed the degree of consistency in their analyses. In reliability assessment, there are inter-rater reliability and intra-rater reliability. By inter-rater reliability, the same content is independently coded by the two or more coders and the findings are compared for agreement (Armstrong et al., 1997). One's own consistent and reliable coding practices are as important as consistent reliability among multiple coders. Therefore, sometime after the initial analysis, the same coder re-analyses the same content. This is called intra-coder reliability assessment. As shown in Figure 4, the majority of the studies, to be precise, 16 studies (39%) did not undergo any reliability assessment in the coding process. Ten studies (24%) applied both inter-coder and intra-coder reliability. Thirteen studies (32%) found the reliability agreement between the two coders. Two of the

studies (5%) sought only intra-coder agreement. It is also found that some of the studies that conducted the reliability assessment did not attempt to assess the full content of the analysis. For example, four studies in the intra-coder reliability assessment group acknowledged that only a small portion of the whole text was randomly selected to check the reliability agreement. In the inter-coder reliability assessment group, five studies acknowledged that only a few portions of the whole text were randomly selected and analysed by their co-coders.

These findings indicate that there is not enough interest or awareness among researchers in the reviewed studies regarding the validity and reliability of assessment methods. There is also little description of the detailed procedures of data handling and analysis methods, which would have been useful to mention for readers seeking to replicate the studies. For such an inadequate methodological description, as Elo and Kyngäs (2008) noted in their methodological guidelines for content analysis research method, researchers need to provide readers with a clear description of how the analysis is done. Overall, these methodological weaknesses reduce the potential value of the results and the reproducibility of the studies. Future studies should be improved in these major aspects of content analysis, validity, and reliability. For the deeper understanding of content analysis research method and its validity establishment and reliability assessment, textbook analysts are referred to Krippendorff (2019).

### **Recommendations Made by the Studies**

Of the three types of textbook evaluations presented in the literature review section of this study, pre-use, in-use, and post-use evaluations, all 41 empirical studies belong to the in-use evaluation phase because the textbooks analysed were already in use at the time of the analysis. None of the studies were in the post-use evaluation phase, as they did not evaluate the learning outcomes after the textbooks were used. Being in-use evaluation studies, their main recommendations were aimed at language teachers who are the primary users of language textbooks. The recommendations are primarily made for effective use of critical thinking enhanced activities in the class and incorporating higher order thinking skills in the textbooks. The following is a summary of the most common recommendations voiced by the textbook evaluation analysts.

- There should be a balance integration of both higher and lower order thinking skills.
- Teachers should add supplementary activities to reinforce learners' critical thinking activities.
- Teachers should be able to evaluate the textbooks they use in language classes.
- Course book writers and teachers should be aware of the importance of critical thinking activities in language teaching materials.

- There should be workshops and training courses to raise this awareness.
- More future studies should be done to investigate other textbooks used in language teaching programs.
- Textbooks should be analysed across the levels of the same series to see the relevancy and consistency of the thinking activities included.
- Teachers' guide books and students' workbooks are overlooked, which also should be evaluated.
- Decision makers in language programs should carefully select language textbooks that fully incorporate higher-order thinking activities.

To summarise the recommendations given, the goals are to include higher-order thinking skills in more language textbooks, to train teachers with the pedagogical knowledge to incorporate critical thinking into textbooks, and to enable teachers to evaluate textbooks from critical thinking perspective. Although these recommendations are reasonable, evidence-based recommendations as the result of post-use textbook evaluation studies will reveal actual classroom experiences more effectively than textual content analyses that do not involve student-teacher interaction and the assessments of learners' performances. Such kind of post-use evaluations tend to be labour-intensive and time-consuming because they are longitudinal studies. However, it is not impossible for language teachers and researchers to apply such post-use evaluation method. Ellis (1997) recommended seven feasible steps for language teachers to evaluate language teaching materials. Ellis's post-use evaluation methods for evaluating language textbooks indicated that language teachers play an autonomous role in this responsibility. Ultimately, it is essential that teachers know the potential contribution of the language materials they are using. However, we should note that since textbooks used in each language program are different, the practical application of textbook evaluation research is limited to specific context. Therefore, it is important to analyse the language textbooks applied in each context.

### **Limitations of the Study**

The three main limitations of this systematic literature review study are the peer review status of the reviewed studies, the limited scope of inclusion criteria in the literature search, and the lack of a full discussion of all critical thinking frameworks applied in the study. In order to incorporate many studies, the inclusion criteria for the literature research was not limited to peer-reviewed papers. However, the number of non-peer-reviewed papers in our analysis was very small compared to that of the peer-reviewed papers. Furthermore, we did not exclude non-peer-reviewed studies because we believe that lessons can be learned from the limitations and weaknesses of existing studies to improve

future research. During the literature review process, we found out that authors in the same EFL context tend to refer to each other's work in their studies. Thus, by pointing out the current status, strengths, and limitations of existing research in the EFL context regardless of being peer-reviewed or not, future researchers can recognise the contributions, limitations, and shortcomings of previous studies. For this reason, we did not exclude non-peer-reviewed papers. Of the critical thinking frameworks, we focused on the two most frequently applied ones by providing background information on these two frameworks. The purpose of this literature review was to examine frequently applied critical thinking frameworks, so we did not provide a thorough description of the remaining applied thinking frameworks. For detailed information, we refer the concerned reader to the work of educational psychologist Moseley et al. (2005), which is a systematic literature review on critical thinking frameworks developed by scholars in various fields over the 50 years prior to 2005. In their thorough review, they found that 41 thinking taxonomies are useful for developing critical thinking in students. These 41 thinking frameworks were then comprehensively described and evaluated by classifying them into three categories: (1) thinking frameworks related to instructional design, (2) thinking frameworks related to productive thinking, and (3) thinking frameworks related to cognitive structure and development. Of these, there are also 13 thinking frameworks related to instructional design for the purpose of curriculum planning and learning assessment. While a review of these critical thinking frameworks is beyond the scope of this study, we suggest that researchers in related fields consider the application of these critical thinking frameworks and conduct analyses comparing and contrasting them.

## **Conclusion and Recommendations**

The present systematic literature review study on ELT textbook evaluation within the period of 12 years delivers a comprehensive overview of the existing literature on the integration of critical thinking enhanced activities and tasks in English language teaching materials from a textbook analysis perspective. Findings of this review study can be scrutinised in relation to the five major units of analysis specified in the review protocol: (1) critical thinking frameworks applied in the studies, (2) the integration percentage of critical thinking, (3) levels and language areas focused on in the studies, (4) validity and reliability of the studies, and (5) recommendations made by the studies. These five major findings can be summarised as follows.

1. For the application of critical thinking frameworks, Bloom's Taxonomy and Bloom's Revised Taxonomy are the most commonly applied frameworks among ELT textbook evaluation studies.
2. As for the integration percentage of critical thinking, lower order thinking skills are overemphasised in the textbooks. So, language activities that foster higher-order thinking skills should be more incorporated into textbooks.
3. For the language levels and language areas being analysed, textbook evaluations are more frequent at junior high school and high school levels than at elementary and college levels.
4. In many studies, methodological validity and reliability assessment of textbook content analysis have not been carefully considered.
5. Reviewed studies mainly recommended that higher and lower order thinking skills should be integrated into balance inside the textbooks.

By reporting on the above findings, this study can benefit two groups of readers. On the one hand, this study can benefit pedagogues, curriculum developers, and stakeholders who are in a position to implement evidence-based policymaking in English language education. Another audience that would benefit from this study is individual researchers and teachers attempting to evaluate ELT materials. Based on the presented literature background, future researchers can establish a research design that is methodologically reliable and valid, especially to compensate for the weaknesses of previous studies.

Future studies should pay attention to the ethical reporting standards (for more detailed ethical reporting standard, see Publication Manual of the American Psychological Association, 2020). From the planning stage of the research project, judgement calls for inclusion and exclusion of what to analyse should be transparently developed so that the replicability of the studies can be enhanced. Data coding and analysis processes of textbook evaluations should be clearly reported because this information is important to the reader to determine the credibility of reported findings and research methods applied. As Nicholls (2003) emphasised in his "Methods of School Textbook Research," the basic elements that textbook researchers should focus on are the methods used to analyse textbooks and the used in the measurement criteria analysis process. We also recommend that an insightful analytical comparative study between the two educational taxonomies is needed so that textbook analysts can determine which theoretical framework is more appropriate for their research purposes. Evaluation studies of ELT textbooks at the college level should receive more attention since they are less prevalent than at the high school level. Examining a series of textbooks prescribed across different target language levels could reveal the pattern and sequencing of critical thinking activities, and could observe if the tasks and activities are developmentally appropriate to learners' cognitive levels. Content analyses of the reading texts should also focus on the topics and inner meaning of the texts because the degree and cognitive complex-

ity of critical thinking activities largely depend on the type of the text. Some texts require learners to possess only basic comprehension skills, while others require learners to possess deep critical thinking. So, the nature and the topic of the content should be included as one factor in ELT textbook evaluation. For that reason, more in-depth multi-layered analysis studies are required. For the last point, we report that two duplicate publications were detected when collecting studies for this systematic review. Such intentional self-plagiarism can jeopardise data computation and interpretation in systematic review studies. We suggest that future researchers avoid such self-plagiarism and pay attention to ethical reporting standards.

## References

- References marked with an asterisk indicate studies included in the systematic literature review.
- \*Abdelrahman, M. (2014). An analysis of the tenth grade English language textbooks questions in Jordan based on the revised edition of Bloom's taxonomy. *Journal of Education and Practice*, 5(18), 139–151.
- \*Adli, N., & Mahmoudi, A. (2017). Reading comprehension questions in EFL textbooks and learners' levels. *Theory and Practice in Language Studies*, 7(7), 590–595. <https://dx.doi.org/10.17507/tpls.07074>
- \*Akrong, B. S., Chaka, C., & van der Walt, C. (2021). A modified version of Bloom's taxonomy and critical language awareness: Evaluating EAL textbooks for junior high schools in Ghana through a critical literacy approach. *International Journal of Applied Linguistics and English Literature*, 10(4), 106–115.
- \*Alfadda, H., Fatima, M., Ghaffar, A., & Afzaal, M. (2020). Critical thinking perspective in ESL new English textbooks: A case study of Pakistan. *International Journal of Applied Linguistics and English Literature*, 9(2), 24–31.
- \*Alfaki, I. (2014). Sudan English language syllabus: Evaluating reading comprehension questions using Bloom's taxonomy. *International Journal of English Language Teaching*, 2(3), 53–74.
- American Psychological Association. (2020). *Publication manual of the American psychological association* (7th ed.). American Psychological Association.
- Alnofaie, H. (2013). A framework for implementing critical thinking as a language pedagogy in EFL preparatory programmes. *Thinking Skills and Creativity*, 10, 154–158. <http://dx.doi.org/10.1016/j.tsc.2013.09.002>
- Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P., Raths, J., & Wittrock, M. C. (2001). *Taxonomy of learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives* (2nd ed.). Addison and Wesley.
- Armstrong, D., Gosling, A., Weinman, J., & Marteau, T. (1997). The place of inter-rater reliability in qualitative research: An empirical study. *Sociology*, 31(3), 597–606.
- \*Askaripour, S. A. A. (2014). A textbook evaluation of new version of *Top Notch* English series. *English for Specific Purposes World*, 15(44), 1–30.
- \*Assaly, I. R., & Smadi, O. M. (2015). Using Bloom's taxonomy to evaluate the cognitive levels of master class textbook's questions. *English Language Teaching*, 8(5), 100–110.

- \*Azizi, M. (2012). Critical thinking in General English course books taught in Iranian Universities. *Journal of Language, Culture, and Translation*, 1(3), 15–36.
- Behar-Horenstein, L. S., & Niu, L. (2011). Teaching critical thinking skills in higher education: A review of the literature. *Journal of College Teaching & Learning*, 8(2), 25–42.
- \*Birjandi, P., & Alizadeh, I. (2013). Manifestation of critical thinking skills in the English textbooks employed by language institutes in Iran. *International Journal of Research Studies in Language Learning*, 2(1), 27–38.
- Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longman.
- Brown, H. D., & Lee, H. (2015). *Teaching by principles*. Pearson Education.
- Cooper, H. M. (1988). Organizing knowledge syntheses: A taxonomy of literature reviews. *Knowledge in Society*, 1(1), 104–126.
- Cottrell, S. (2017). *Critical thinking skills: Effective analysis, argument and reflection*. Bloomsbury Publishing.
- Cummins, J. (1999). BICS and CALP: Clarifying the distinction. Washington, DC: Opinion Papers. (ERIC Document Reproduction Service No. ED 438 551). Retrieved from <https://eric.ed.gov/?id=ED438551>
- \*Dabbagh, A., & Safaei, A. (2019). Comparative textbook evaluation: Representation of learning objectives in locally and internationally published ELT textbooks. *Issues in Language Teaching*, 8(1), 249–277.
- Dummett, P., & Hughes, J. (2019). *Critical thinking in ELT: A working model for the classroom*. National Geographic Learning.
- Dummett, P. (2016). *How to write critical thinking activities*. ELT Teacher 2 Writer.
- \*Es-Salhi, A., & Elfatih, M. (2019). Evaluating critical thinking skills in Moroccan EFL textbooks: *Gateway to English 2* as a case. *Higher Education of Social Science*, 17(1), 13–22. <http://doi.org/10.3968/11284>
- Elder, L., & Paul, R. (2012). Critical thinking: Competency standards essential to the cultivation of intellectual skills, Part 4. *Journal of Developmental Education*, 35(3), 30–31.
- Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT Journal*, 51(1), 36–42.
- Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of Advanced Nursing*, 62(1), 107–115.
- Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. *Topoi*, 37(1), 165–184. <https://doi.org/10.1007/s11245-016-9401-4>
- Facione, P. A. (2011). Critical thinking: What it is and why it counts. Academic Press. Retrieved from <https://courseware.e-education.psu.edu/downloads/geog882/Critical%20Thinking%20What%20it%20is%20and%20why%20it%20counts.pdf>
- Fahim, M., Barjesteh, H., & Vaseghi, R. (2012). Effects of critical thinking strategy training on male/female EFL learners' reading comprehension. *English Language Teaching*, 5(1), 140–145.
- \*Fitriani, D. H., & Kirana, D. P. (2022). Higher order thinking skill (hots) in English language textbook in senior high school. *International Journal of Research on English Teaching and Applied Linguistics*, 2(2), 55–66.
- \*Frehat, N. M., & Smadi, O. M. (2014). Lower-order and higher-order reading questions in secondary and university level EFL textbooks in Jordan. *Theory & Practice in Language Studies*, 4(9), 1804–1813. <https://doi.org/10.4304/tpls.4.9.1804-1813>
- Gardner, H. (1993). *Frames of mind: The theory of multiple intelligences*. Basic Books.

- Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.
- Golpour, F. (2014). Critical Thinking and EFL learners' performance on different writing modes. *Journal of Pan-Pacific Association of Applied Linguistics*, 18(1), 103–119.
- \*Gordani, Y. (2010). An analysis of English textbooks used at Iranian guidance schools in terms of Bloom's taxonomy. *Journal of Asia TEFL*, 7(2), 249–278.
- Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning*, 89, 69–74.
- Halpern, D. F. (2002). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New Directions for Teaching and Learning*, 1999(80), 69–74.
- \*Hei, J. X. (2017). Thinking skills in the 12th grade English course book in Timor Leste. *Beyond Words*, 5(1), 32–45.
- Hughes, J. (2014). Critical thinking in the language classroom. *Recanati (Italy): ELI*. Retrieved from: <http://www.elionline.com/eng/teachers-area/critical-thinking>
- Ilyas, H. P. (2015). *Critical thinking: Its representation in Indonesian ELT textbooks and education* [Unpublished doctoral thesis]. University of York.
- \*Igbaria, A. K. (2013). A content analysis of the WH-Questions in the EFL textbook of *Horizons*. *International Education Studies*, 6(7), 200–224. <https://doi.org/10.5539/ies.v6n7p200>
- \*Jebbour, M. (2019). Exploring the manifestation of critical thinking in the Moroccan textbooks of English: The case of *Ticket 2 English*. *JEES (Journal of English Educators Society)*, 4(1), 39–45.
- Jones, E., Hoffman, S., Moore, L., Ratcliff, G., Tibbetts, S. & Glick, B. (1995). *Essential Skills in Writing, Speech and Listening, and Critical Thinking for College Graduates: Perspectives of Faculty, Employers, and Policy Makers*. Project summary, U.S. Department of Education, OERI Contract No. R117GI0037. National Center for Post-secondary Teaching, Learning and Assessment. University Park, Pennsylvania State University Press.
- \*Khodadady, E., & Karami, M. (2017). An evaluation of textbooks designed for advanced English learners within a foreign context: A critical thinking perspective. *Porta LINGUARUM*, 28, 95–109.
- Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory into Practice*, 41(4), 212–218.
- Krippendorff, K. (2019). *Content analysis: An introduction to its methodology*. Sage.
- Lin, Y. (2018). *Developing critical thinking in EFL classes*. Springer.
- Littlejohn, A. (2011). The analysis of language teaching materials: Inside the Trojan Horse. In B. Tomlinson (Ed.), *Materials development in language teaching* (2nd ed., pp. 179–211). Cambridge University Press.
- Lune, H., & Berg, B. L. (2017). *Qualitative research methods for the social sciences*. Pearson.
- \*Margana, & Widyanoro, A. (2017). Developing English textbooks oriented to higher order thinking skills for students of vocational high schools in Yogyakarta. *Journal of Language Teaching and Research*, 8(1), 26–38. <http://dx.doi.org/10.17507/jltr.0801.04>
- McDonough, J., Shaw, C. & Masuhara, H. (2013). *Materials and Methods in ELT* (3rd ed.). John Wiley & Sons.
- McGrath, I. (2002). *Materials evaluation and design for language teaching*. Edinburgh University Press.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation*. Jossey-Bass.

- \*Mizbani, M., & Chalak, A. (2017). Analyzing reading and writing activities of Iranian EFL textbook *Prospect 3* based on Bloom's revised taxonomy. *Journal of Applied Linguistics and Language Research*, 4(2), 13–27.
- Moher D, Liberati A., Tetzlaff J., Altman D., the PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Annals of Internal Medicine*, 151(4), 264–270.
- Moore, T. (2013). Critical thinking: Seven definitions in search of a concept. *Studies in Higher Education*, 38(4), 506–522.
- Moore, B. & Parker, R. (2009). *Critical thinking* (9th ed.). McGraw-Hill Higher Education.
- \*Muchlis, M. (2015). An analysis of thinking order of reading comprehension questions in English textbook for young foresters of Forestry Vocational School of Samarinda. *Jurnal Nalar Pendidikan*, 3(1), 314–316.
- \*Nainggolan, B. R., & Wirza, Y. (2021). Evaluation of an English textbook: Does *English in Mind* promote critical thinking skills? *Advances in Social Science, Education and Humanities Research*, 546, 51–59.
- Nappi, J. S. (2017). The importance of questioning in developing critical thinking skills. *Delta Kappa Gamma Bulletin*, 84(1), 30–41.
- Nicholls, J. (2003). Methods in school textbook research. *International Journal of Historical Learning, Teaching and Research*, 3(2), 1–17.
- Nilson, L. B. (2021). *Infusing critical thinking into your course: A concrete, practical approach*. Stylus Publishing.
- \*Olimat, M. (2015). Analyzing action pack textbooks' questions according to revised Bloom taxonomy. *Journal of Education and Practice*, 6(28), 152–159.
- Peterson, S. (2008). *Master critical thinking for the SAT*. Peterson's.
- Potter, W. J., & Levine-Donnerstein, D. (1999). Rethinking validity and reliability in content analysis. *Journal of Applied Communication Research*, 27, 258–284.
- Puig, B., Blanco-Anaya, P., Bargiela, I. M., & Crujeiras-Pérez, B. (2019). A systematic review on critical thinking intervention studies in higher education across professional fields. *Studies in Higher Education*, 44(5), 860–869.
- \*Rahpeyma, A., & Khoshnood, A. (2015). The analysis of learning objectives in Iranian junior high school English textbooks based on Bloom's revised taxonomy. *International Journal of Education & Literacy Studies*, 3(2), 44–55. <https://doi.org/10.7575/aiac.ijels.v3n.2p.44>
- \*Raqqad, Y. M. A., & Ismail, H. H. (2018). Analyzing the reading questions of AP 12 textbook according to analyzing the reading questions of *apl2* textbook according to Bloom's taxonomy. *International Journal of Education, Psychology and Counselling*, 3(22), 84–94.
- \*Razmjoo, S. A., & Kazempourfard, E. (2012). On the representation of Bloom's revised taxonomy in *Interchange* course books. *The Journal of Teaching Language Skills (JTLS)*, 4(1), 171–204.
- \*Riazi, A. M., & Mosalanejad, N. (2010). Evaluation of learning objectives in Iranian high-school and pre-university English textbooks using Bloom's taxonomy. *TESL-EJ*, 13(4)1–16.
- \*Roohani, A., Taheri, F., & Poorzangeneh, M. (2013). Evaluating *Four Corners* textbooks in terms of cognitive processes using Bloom's revised taxonomy. *Research in Applied Linguistics*, 4(2), 51–67.
- \*Roohani, A. (2015). Analyzing cognitive processes and Multiple Intelligences in the nop-notch textbooks. *Journal of Modern Research in English Language Studies*, 2(3), 65–39.
- \*Sahragard, R., & Zahed Alavi, S. (2016). Investigating the predominant levels of learning objectives in general English books. *Journal of English Language Teaching and Learning*, 17(8), 93–114.

- Sanavi, V. R., & Tarighat, S. (2014). Critical thinking and speaking proficiency: A mixed-method study. *Theory and Practice in Language Studies*, 4(1), 79–87.
- Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal*, 42(4), 237–246.
- Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for Information*, 22(2), 63–75.
- Shirkhani, S., & Fahim, M. (2011). Enhancing critical thinking in foreign language learners. *Procedia-Social and Behavioral Sciences*, 29, 111–115. <https://doi.org/10.1016/j.sbspro.2011.11.214>
- \*Shuyi, N. S., & Renandya, W. A. (2019). An analysis of the cognitive rigour of questions used in secondary school English language textbooks in Singapore. *Asian Journal of English Language Studies (AJELS)*, 7, 169–188.
- \*Sirait, R. E. S. R. (2014). Comprehension levels of reading exercises in *Look Ahead English* coursebooks. *International Journal of English Language Education*, 2(2), 253–268. <http://dx.doi.org/10.5296/ijele.v2i2.6307>
- Solihati, N., & Hikmat, A. (2018). Critical thinking tasks manifested in Indonesian language textbooks for senior secondary students. *Sage Open*, 8(3), 1–8. <https://doi.org/10.1177/2158244018802164>
- \*Sucipto, S., & Cahyo, S. D. (2019). A content analysis of the reading activities in *Bright 2* an English textbook for junior high school students. *English Language Teaching Educational Journal*, 2(1), 13–21.
- \*Tabari, M. A., & Tabari, I. A. (2015). Links between Bloom's taxonomy and Gardener's Multiple Intelligences: The issue of textbook analysis. *Advances in Language and Literary Studies*, 6(1), 94–101. <https://doi.org/10.7575/aiac.all.v.6n.1p.94>
- \*Talebinezhad, M. R., & Matou, Z. (2012). EFL reading comprehension textbooks at university level: A critical thinking perspective. *Theory & Practice in Language Studies*, 2(8), 1746–1754. <https://doi.org/doi:10.4304/tpls.2.8.1746-1754>
- \*Tangsakul, P., Kijpoonphol, W., Linh, N. D., & Kimura, L. N. (2017). Using Bloom's revised taxonomy to analyze reading comprehension questions in *Team Up* in English 1–3 and grade 9 English o-net tests. *International Journal of Research-Grantaalayah*, 5(7), 31–41.
- \*Tayyeh, M. N. (2021). An analysis of reading comprehension questions in English textbook *English For Iraq* according to revised Bloom's taxonomy. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(7), 2868–2874.
- Thomson, C. (2011). Critical thinking across the curriculum: Process over output. *International Journal of Humanities and social science*, 1(9), 1–7.
- Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching*, 45(2), 143–179.
- Tomlinson, B., & H. Masuhara (2004). *Developing language course materials*. RELC Portfolio Series.
- \*Ulum, Ö. G. (2016). A descriptive content analysis of the extent of Bloom's taxonomy in the reading comprehension questions of the course book Q: Skills for success 4 reading and writing. *The Qualitative Report*, 21(9), 1674–1683.
- \*Wu, Y., & Pei, Z. (2018). An investigation of critical thinking manifested in the questions of EFL textbooks for tertiary-level English majors of China. *American Journal of Education and Learning*, 3(2), 72–84. <https://doi.org/10.20448/804.3.2.72.84>
- Webb, N. L. (2002). *Depth of knowledge levels for four content areas* [Unpublished manuscript]. Wisconsin Center for Education Research, University of Wisconsin Madison. Retrieved from <http://ossucurr.pbworks.com/w/file/fetch/49691156/Norm%20web%20dok%20by%20subject%20area.pdf>
- Williams, D. (1983). Developing criteria for textbook evaluation. *ELT Journal*, 37(3), 251–255.

- Xiao, Y., & Watson, M. (2017). Guidance on conducting a systematic literature review. *Journal of Planning Education and Research*, 39(1), 93–112. <https://doi.org/10.1177/0739456X17723971>
- Yang, Y. T. C., & Gamble, J. (2013). Effective and practical critical thinking-enhanced EFL instruction. *ELT Journal*, 67(4), 398–412.
- \*Zamani, G., & Rezvani, R. (2015). ‘HOTS’ in Iran’s Official Textbooks: Implications for Material Design and Student Learning. *Journal of Applied Linguistics and Language Research*, 2(5), 138–151.
- \*Zareian, G., Davoudi, M., Heshmatifar, Z., & Rahimi, J. (2015). An evaluation of questions in two ESP course books based on Bloom’s new taxonomy of cognitive learning domain. *International Journal of Education and Research*, 3(8), 313–326.

Thiri Soe

## **Untersuchung des kritischen Denkens in Englischlehrbüchern: Eine systematische Literaturübersicht über Studien zur Evaluation von Lehrbüchern**

### **Zusammenfassung**

Die zentrale Rolle des kritischen Denkens und dessen Integration in den Englischunterricht sind durch systematische Literaturübersichten und Metaanalysen in den Bereichen von pädagogischen Interventionen, Lehrmethoden und Bewertungsmethoden zusammengefasst. Es gibt allerdings kaum Belege für eine systematische Literaturübersicht über die Integration von kritischem Denken in Lehrmaterialien für den Englischunterricht, wie z. B. Lehrbücher. In der vorliegenden Arbeit wurden 41 empirische Studien zur Evaluation von Lehrbüchern aus den Jahren 2010-2021 im Hinblick auf kritisches Denken untersucht. Ziel der Arbeit ist es, einen Überblick über die Evaluation von Englischlehrbüchern in Bezug auf die Integration von kritischem Denken zu geben, die angewandten Rahmenkonzepte für kritisches Denken zu identifizieren und mögliche Forschungslücken aufzuzeigen. Hierfür wurde eine umfangreiche Recherche unter Anwendung der achsstufigen systematischen Literaturübersichtsmethode von Xiao und Watson (2017) durchgeführt. Basierend auf einer Sammlung von 41 empirischen Studien konzentriert sich die Arbeit auf folgende vier Aspekte: angewandte Rahmenkonzepte für kritisches Denken, Untersuchungsergebnisse, Empfehlungen sowie Sprachniveaus und Inhaltsbereiche der untersuchten Lehrbücher. Die Ergebnisse zeigen, dass (1) die Bloomsche Taxonomie und die revidierte Bloomsche Taxonomie die am häufigsten angewandten Rahmenkonzepte für kritisches Denken waren, (2) die Integration von Denkfähigkeiten höherer Ordnung in Lehrbüchern seltener war als die von Denkfähigkeiten niedrigerer Ordnung, (3) die Evaluation von Lehrbüchern häufiger in der Mittel- und Oberstufe erfolgte als in anderen Stufen und (4) methodische Überlegungen zur Reliabilität und Validität des Kodierungsprozesses bei der Inhaltsanalyse von Lehrbüchern wenig Beachtung fanden. Die Arbeit liefert einen zusammenfassenden Literaturüberblick über die Evaluation von Englischlehrbüchern mit Empfehlungen für methodische Verbesserungen in zukünftigen Studien.

*Schlüsselwörter:* systematische Literaturübersicht, kritisches Denken, Evaluation von Englischlehrbüchern, Denkfähigkeiten höherer Ordnung, Denkfähigkeiten niedrigerer Ordnung

## 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

### Format of headings

The following table demonstrates how to format headings in APA Style.

Level	Format
1	<b>Centered, Bold, Title Case Heading</b> Text begins as a new paragraph.
2	<b>Flush Left, Bold, Title Case Heading</b> Text begins as a new paragraph.
3	<b><i>Flush Left, Bold Italic, Title Case Heading</i></b> Text begins as a new paragraph.
4	<b>Indented, Bold, Title Case Heading, Ending With a Period.</b> Text begins on the same line and continues as a regular paragraph.
5	<b><i>Indented, Bold Italic, Title Case Heading, Ending With a Period.</i></b> Text begins on the same line and continues as a regular paragraph.

*Note.* In title case, most words are capitalized.

### 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).



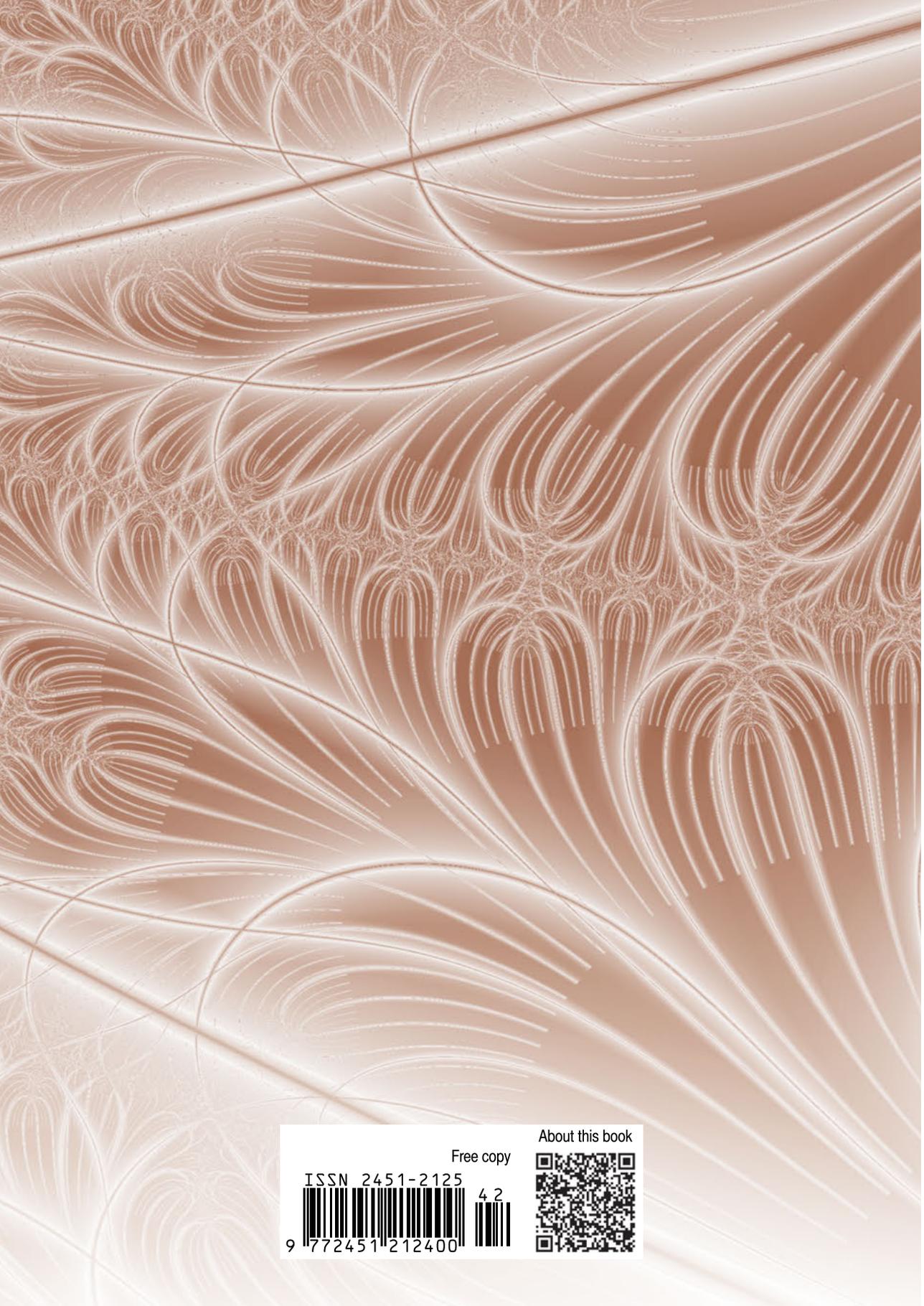
Copy editing: Gabriela Marszolek-Kalaga  
Proofreading: Joanna Zwierzyńska  
Typesetting: Tomasz Kielkowski  
Cover preparation for printing: Paulina Dubiel

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](http://www.wydawnictwo.us.edu.pl)  
e-mail: [wydawnictwo@us.edu.pl](mailto:wydawnictwo@us.edu.pl)

First impression. Printed sheets: 21,75. Publishing sheets: 26,0.



Free copy

ISSN 2451-2125



4 2

9 772451 212400

About this book

