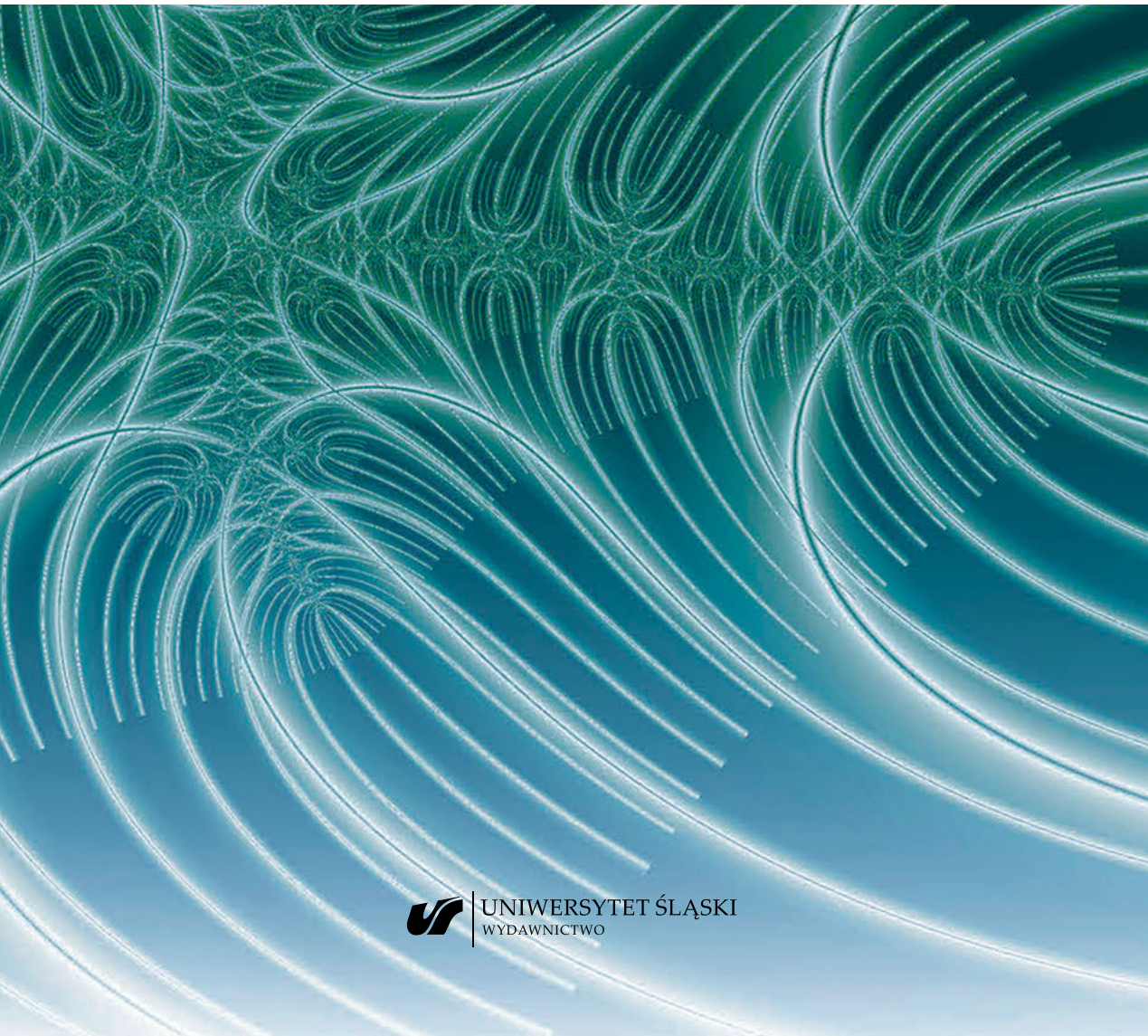
# Theory and Practice of Second Language Acquisition

Vol. 11 (1), 2025



UNIWERSYTET ŚLĄSKI
WYDAWNICTWO

# Theory and Practice of Second Language Acquisition

Vol. 11 (1), 2025

Wydawnictwo Uniwersytetu Śląskiego • Katowice 2025

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)
Danuta Gabryś-Barker (University of Silesia in Katowice)
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

**Bruno Costa, Jorge Pinto**
Who Cares If the Teacher Has a Cat? The Impact of Affect and Interaction upon Motivation in Online Portuguese L2 Classes during the COVID-19 Pandemic

**Jasrael D. Stokes**
The Impact of Foreign Language Anxiety on the Utterance Fluency of International Students in the UK: Does the Presence of an Audience Matter? . .

**Agata Wolanin**
"I Feel Like I'm a Different Person": Exploring Undergraduate Students' Imagined L2 Selves . . . . . .

**Katarzyna Morena**
Academics' Motivation to Learn Foreign Languages: The Case of English . . .

**Adelina Castelo**
Model for Using Music in Pronunciation Teaching (MOMUP): New Validation and Revision . . . . . .

**Andrew Schenck**
Gleaning Insights about Input-Based and Output-Based Form-Focused Instruction through Meta-Analysis of Korean EFL Learners . . . . . .

**Petra Langerová**
The Learning Styles of Technical Students and Their Role in Learning English as a Foreign Language . . . . . .

**Zdeňka Neumanová**
An Investigation of Complexity, Accuracy, and Fluency in the Speech of EFL Learners . . . . . .

**Eihab Abu-Rabiah**
The Productive Vocabulary Size of Second Language Learners upon Entry into Higher Education . . . . . .

**Meihua Liu**
Learning Style Preferences, Strategy Use and Chinese Undergraduate EFL Students' English Achievements . . . . . .

**Mohammad Hamad Al-khresheh**
The Role of Digital Storytelling in Jordanian School EFL Classrooms: A Qualitative Exploration of Teachers' Perceptions . . . . . .

## Reviews

Rachelle S. Savitz, Leslie D. Roberts, & Jason DeHart (Eds.) (2023), *Teaching Challenged and Challenging Topics in Diverse and Inclusive Literature: Addressing the Taboo in the English Classroom*—by Kornél Farkas .

Style Guide for the Authors . . . . . .

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/2
https://doi.org/10.31261/TAPSLA.18011



# Preface

As we begin the 11th volume of *Theory and Practice of Second Language Acquisition* (TAPSLA), it is with great pride that we reflect on the journal's journey and its well-established presence in the field of Second Language Acquisition research. Over the past decade, TAPSLA has evolved into a respected and thriving journal, recognized for its commitment to advancing both theoretical and practical insights into the complex processes of language learning and teaching.

From its inception, TAPSLA has served as a platform for scholars from across the globe, reflecting a truly international authorship. Each volume has contributed to a diverse and interdisciplinary body of knowledge, drawing together research from cognitive, sociocultural, pedagogical, and technological perspectives. These contributions have enriched the field of SLA, making TAPSLA a grounded and authoritative voice among journals dedicated to this area of research.

The contents of previous volumes, all available in open access via the University of Silesia in Katowice platform, have explored a wide range of topics—from the role of technology in language education, through the cognitive processes underlying language acquisition, to sociolinguistic factors shaping learners' experiences. Our mission has always been to bridge theory and practice, bringing together rigorous research with practical implications for educators and learners alike.

The present volume (11/1) continues this tradition. The papers featured in this issue delve into pressing topics relevant to contemporary language learning environments. These include the impact of **motivation and interaction in online language classes**, a theme that resonates deeply given the global shift towards digital education during the COVID-19 pandemic. The issue also addresses **foreign language anxiety**, a pervasive factor affecting learners' fluency, and explores the **imagined L2 selves** of students as they navigate their

linguistic identities. In addition, themes such as **pronunciation teaching**, **vocabulary acquisition**, and the **role of learning styles** are examined in various educational contexts. The volume also features **meta-analyses of instructional methods**, offering valuable insights for both researchers and practitioners.

As always, this issue reflects the diversity of approaches, languages, and contexts that characterize SLA research today. We are particularly pleased to see contributions from authors representing different educational systems and cultural perspectives, highlighting the universal and multifaceted nature of second language acquisition.

We extend our sincere gratitude to all the contributors, reviewers, and readers who continue to support TAPSLA. Your dedication has made the journal a vibrant and influential part of the global SLA community. We look forward to many more years of collaboration, innovation, and discovery in the field of second language research.

Sincerely,

*Editors-in-Chief*

*Jolanta Latkowska*
https://orcid.org/0000-0002-9478-7689

*Adam Wojtaszek*
https://orcid.org/0000-0003-0308-4337

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/24
https://doi.org/10.31261/TAPSLA.16026

CC BY SA

**Bruno Costa** https://orcid.org/0000-0002-3168-5108
School of Arts and Humanities, Centre of Linguistics
University of Lisbon, Portugal

**Jorge Pinto** https://orcid.org/0000-0002-5583-880X
School of Arts and Humanities, Centre of Linguistics
University of Lisbon, Portugal

# Who Cares If the Teacher Has a Cat? The Impact of Affect and Interaction upon Motivation in Online Portuguese L2 Classes during the COVID-19 Pandemic[1]

Abstract

This study examined the motivational impact of affect/interaction in emergency remote teaching (ERT) Portuguese L2 classes. Participants were two groups of students from 13 countries living in Portugal ($N$ = 16). A mixed-method paradigm was adopted, consisting of a closed questionnaire, class observations, and semi-structured interviews. Results showed a generally positive impact of the practices adopted by both teachers. Comparatively, the group whose teacher evidenced a stronger affect-based strategy presented better results in measures related to motivation to learn Portuguese (MLP), perception of the teacher/classes (PTC), perception of individual capacities (PIC) and test anxiety, regardless of learners being more critical of the lack of socialization and reticent about sharing personal information. However, the group whose teacher favored interactive tasks showed greater engagement and enjoyment/flow. Statistical tests revealed a significant positive correlation between PTC and MLP, and a negative correlation between these two and test anxiety. It is proposed that the combination of affect and interaction can compensate for the demotivating effect of lack of face-to-face contact in ERT L2 classes.

*Keywords*: emotions, motivation, interactions, engagement, ERT L2 teaching, COVID-19

[1] The authors acknowledge the financial support provided by the Portuguese National Science Foundation (FCT) to the Centre of Linguistics of the University of Lisbon (reference UIDB/00214/2020).

The unruly spread of the SARS-COVID-19 virus had an overwhelming impact on people's routines in every corner of the world. The short-notice closure of school buildings severely constrained education in most countries, forcing teachers to adapt their course plans, materials, and teaching strategies to a digital format with which many had been unfamiliar until then. Students were also challenged by a new virtual reality for which many were physically and emotionally unprepared. This new educational model has been referred to as emergency remote teaching (ERT), owing to the particularities that distinguish it from other online formats (Hodges et al., 2020). In such a context, L2 classes were no exception to the toll that this shift has taken on both teachers and students.

Lately, the role played by emotion/affect in learning has been brought to the foreground of educational research (Pekrun, 2006; Rodrigo-Ruiz, 2016) and studies on the acquisition of an L2 (Altarriba & Canary, 2004; Dewaele, 2011; Dewaele & MacIntyre, 2014, 2016; Fraschini & Tao, 2021). Research in this area has recently started to look into achievement emotions experienced by L2 learners in the classroom, such as foreign language enjoyment (FLE), foreign language anxiety (FLA), classroom test anxiety (CTA), foreign language boredom (FLB), and, more recently, flow.

FLE has been defined as a "complex emotion, capturing interacting dimensions of challenge and perceived ability that can reflect the human drive for success in the face of difficult tasks" (Dewaele & MacIntyre, 2016, p. 216). Horwitz and colleagues (1986, p. 128) defined FLA as a "distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning, arising from the uniqueness of the language learning process." CTA has been defined as "a type of performance anxiety stemming from a fear of failure in a test situation" (Zheng, 2010, p. 38). FLB refers to a state of low physical and cognitive arousal experienced in L2 classes, associated with some negative perceptions of the time not flowing, dragging on, and disengaging behaviors, like withdrawal and distraction (Li et al., 2021). Finally, flow is closely linked to the intensity and duration of FLE in an L2 classroom (Dewaele & MacIntyre, 2022), but excels it in the sense that it creates a distinguished state in which "a learner's thoughts, feelings, and behavior reflect effortless and harmonious coordination in challenging situations that offer both risk and reward" (Dewaele & MacIntyre, p. 20).

Teachers' emotions, attitudes, and classroom practices strongly impact the way learners participate in the educational process (Dörnyei, 2001; Arnold-Morgan & Fonseca-Mora, 2007; Dewaele, 2011). How they approach learners and class activities influences students' emotions, engagement, self-confidence, and motivation (Arnold-Morgan & Fonseca-Mora, 2007; Garret & Young, 2009; Dörnyei, 2009; Dewaele, 2011). The goal of this study is to examine the impact of two different strategies used by teachers in ERT classes conducted

with Portuguese L2 multicultural groups—one emphasizing learners' interactions and another underpinned by the sharing of affect/emotion—upon a cohort of achievement emotions (i.e., FLA, CTA, FLE, flow, FLB), as well as on other learner-centered factors (i.e., self-esteem, motivation, perception of the teachers and the classroom and engagement).

## Teachers' Impact of Different Variables in the Classroom

Several studies have analysed how teachers' emotions, attitudes, and strategies influence different variables in L2 classes. Teacher friendliness has been found to predict FLE (Dewaele, Magdalena, & Saito, 2019). While FLA is more linked to learners' internal variables, FLE is more predicted by teachers' characteristics. Thus, more positive attitudes towards the teacher are a strong predictor of FLE (Dewaele, 2023). Dewaele and MacIntyre (2022) also suggest that flow is more closely related to FLE than to FLA and that teachers' feedback can bring about flow, which can also be induced when they boost L2 enjoyment in class. Finally, Li and Dewaele (2020) highlight that learners' attitudes towards the teachers and the activities performed in class are two important predictors of FLB.

Recognizing that the teacher-learner relationship remains an under-examined factor in motivation, Moskowitz et al. (2022) investigated this variable more closely. The findings confirmed the impact of both the interpersonal teacher-learner relationships and teachers' behaviors and attitudes upon learners' motivation towards the L2. Particularly, participants revealed an appreciation for their teachers' taking interest in their personal lives and bonding with them, paralleled with criticism of teachers' distant and intimidating behavior, which has been reported as anxiety-provoking, underscoring the complex nature of emotions in L2 learning.

With the outbreak of COVID-19, a new context for the investigation of these variables has emerged. In the context of L2 studies, the emotional impact of ERT has been investigated by scholars like Resnik and Dewaele (2021), Dewaele, Albakistani, and Ahmed (2022a, 2022b), and Wang and Jiang (2022), to mention just a few. Two variables that appear to be central in most studies are the role played by teachers' and learners' emotions and by social interactions in this "disembodied" (Resnik & Dewaele, 2021, p. 21) context of learning. Findings have been somewhat inconclusive so far.

## Achievement Emotions in the Context of L2 ERT Classes during the COVID-19 Pandemic

Resnik and Dewaele (2021) studied tertiary-level EFL learners in ERT classes in Europe and reported a general decrease in both positive and negative emotions, proposing that ERT classes dull learners' emotions. Regarding FLE, a decline in the personal and social subscales was highlighted, whereas teacher appreciation remained more stable. The authors found a significant variation in the sources of FLE when comparing ERT to face-to-face classes. Participants referred to their teachers as being supportive, but increased boredom was reported owing to learners' social detachment and lower engagement, curbing the development of a sense of collectivity. The lower levels of anxiety were assumed to be a consequence of students interacting less and using the language less often. In general, levels of FLE outweighed those of FLA, and the moderate link between these two measures found in regular classes disappeared in ERT.

Dewaele, Albakistani, and Ahmed (2022a) carried out a mixed-method study with EFL Arab learners, investigating positive and negative views concerning both in-person and online classes. Positive views regarding in-person classes included the feeling of being more socially satisfying and more likely to generate positive emotions, while in ERT the flexibility and the feeling of confidence awarded by this format were underscored. The negative aspects of ERT classes consisted of technical issues in the use of the internet, lack of social interaction !and immediate feedback from teachers, absence of the pleasant atmosphere of in-person classes, and negative emotions related to physical separation. In face-to-face classes, the only negative aspect informed was that classes were more time-consuming. The authors concluded that both positive and negative emotions can co-occur in online classes, as is true of in-person sessions, and that the lower levels of FLA found in ERT could be a consequence of FLE neutralizing the deleterious effects of anxiety. In a different study, the authors found that learners in ERT classes experienced less flow than in regular classes and that the sources of flow differed considerably, being reduced to attitude towards teachers in ERT (Dewaele, Albakistani, & Ahmed, 2022b).

Resnik, Knechtelsdorfer, and Dewaele (2022) conducted a more granulated study on the measures of FLA in ERT classes with tertiary-level EFL students. Statistical analysis of the quantitative data revealed a considerable drop in levels of FLA. However, the qualitative data gathered from interviews showed that FLA is a complex phenomenon and that some aspects of online classes can be more anxiety-provoking than others. Not only did anxiety-provoking aspects of FLA in ERT outnumber those indicated in in-person classes, but also the

sources of anxiety varied significantly in the two contexts. In ERT classes, the use of technology, online exams, uncertainty about class/task requirements, turn-taking difficulties, shortage of feedback, and increased workload were among the main anxiety factors reported. The authors concluded that teachers need to ensure that learners are provided with plenty of interaction with teachers and peers, which is crucial for learners to get to know one another, build connections, and form social bonds, as well as for diminishing class embarrassment. FLB also increased in ERT classes, as participants reported experiencing distraction, disengagement, frustration, and social disconnection. Similarly, Li and Dewaele (2020) found that participants learning EFL in a university course in China in ERT mode experienced considerably more FLB than in face-to-face classes.

Looking more closely into FLE in ERT classes, Wang and Jiang's (2022) study with Chinese L2 learners revealed that participants experienced high levels of FLE in their classes. It was found that FLE private and FLE teachers presented higher measures than FLE atmosphere. FLE private was associated with students perceiving their progress in the L2, learning new things in class, as well as using Chinese in their classroom interactions. FLE teacher was linked to teachers' supportive behavior, as well as their diversified instructional approaches. Group activities engaging teachers, peers and learners positively predicted FLE, as participants referred to getting a feeling of belonging from such tasks. Teacher-related variables were found to contribute significantly to FLE.

Two studies took a qualitative-only approach to L2 learning in ERT. Elmas and Öztüfekçi (2021) examined L2 learning demotivation from the prism of Engeström's (1999, 2009) activity theory (AT) in a case study. The findings confirmed that demotivation was strongly linked to negative aspects of the institutionalized teaching/learning context, such as the methods used and the tasks assigned to participants. Wilson and Lengeling (2021) studied university students' reflective journal entries during the shift to virtual language classes. The challenges mentioned by learners related to technological obstacles, but also the lack of interaction. Emotions were described as a "rollercoaster," with highs and lows. Overall, learners' anxiety was surmounted by their motivation to learn the L2, and participants' reflections unveiled their resilience along the process.

The relationship between learners' perceived social support and their engagement in ERT L2 classes was tested by Luan et al. (2020). Learners' perceived social support predicted behavioral engagement, which, in turn, predicted cognitive, emotional, and social engagement. The findings confirmed that peer support plays a significant role in behavioral engagement. Likewise, Mihai et al. (2022) sought to investigate factors that determine learners' engagement in ERT L2 classes, targeting moderately engaged vs. highly engaged learners.

The study confirms that highly motivated students are less anxious and more engaged and that teaching methods play an important role, especially by promoting interaction, which, in turn, leads to a sense of community, belonging, friendliness, and safety.

# The Study

## Methodology

### *Objectives and Research Questions*

The purpose of this study is to check which of the two strategies used by two different teachers in their respective Portuguese L2 groups in ERT mode was more effective in terms of motivating their learners: one that is strongly based on affect or one that is mostly underpinned by tasks that explore classroom interaction. The following research questions have been formulated:

RQ1: How does an affect-based teaching strategy impact learners' motivation in ERT Portuguese L2 classes?

RQ2: How does an interaction-intensive teaching strategy impact learners' motivation in ERT Portuguese L2 classes?

RQ3: Which variables are impacted by each of these strategies?

### *Participants*

Participants were two groups of adult learners (eight per group) of Portuguese L2 (level B1.2) classes in the University of Lisbon, from 13 different countries. The teachers were two female, European Portuguese L1 speakers. Classes were held on Zoom, twice a week, during the whole semester.

### *Study Design and Instruments*

To answer our research questions, a mixed-method approach was adopted, combining qualitative and quantitative methods. The data analyzed herein were collected through the following instruments: (a) a closed questionnaire consisting of a small version of the AMTB (Gardner & MacIntyre, 1993), to measure participants' attitudes and motivation to learn Portuguese at the beginning and at the end of the course; (b) a class observation grid (Appendix A) adapted from Guilloteaux and Dörnyei's (2008) scheme to measure participants' behavioral engagement in online classes (attention, participation, and volunteering for

teacher-fronted activity); (c) a class observation journal where relevant information about the class dynamics, especially learners' behaviors, attitudes, interactions, teachers' attitudes, emotions and strategies and the reactions they elicited from the students, and the general classroom atmosphere, was registered; and (c) interviews with students, focusing on affective and motivational aspects of the classes. A thorough content analysis was then carried out on these data, and themes were identified both inductively and deductively. A final triangulation of all data was conducted.

For the purposes of the present study, four main variables in the questionnaire were emphasized, namely: (i) learners' perceptions of individual capacities (PIC), which is a measure combining self-esteem and language anxiety; (ii) test anxiety (CTA); (iii) attitude and intensity of motivation to learn Portuguese (MLP); and (iv) perceptions of the teacher and the classes (PTC). The items developed to measure PIC, which was not originally included in the AMTB, have been formulated based on Zheng (2010) and Deci and Ryan (1995). Statistical tests were run on the quantitative data, using the software SPSS. As the samples were rather small for parametric tests, we first decided to run a paired samples Wilcoxon test on each category, comparing the results from the first and the second questionnaire. Spearman's rank coefficient was calculated to verify the statistical dependence between the ranking of these four variables.

## Findings

### *The AMTB Questionnaire*

To analyze the results of the questionnaire, the 32 items were rearranged according to the eight categories they belong to and, following Zheng (2010), a numeric scale was designed for the answers, so that "I totally disagree" = 1, "I partly disagree" = 2, "I partly agree" = 3, and "I totally agree" = 4. The items that correspond to the categories PIC and CTA (self-esteem, language anxiety, and test anxiety) received inverted values, as the statements were written positively (e.g., "Before I take a Portuguese test, I feel confident and at ease").

First and foremost, it must be said that both Group 1 and Group 2 revealed highly motivated and rather similar profiles at the end of the course, displaying the same order in the ranking of all four variables (Figures 1 and 2).

**Figure 1**

*Ranking of Means at the End of the Course in Group 1 (N = 8)*

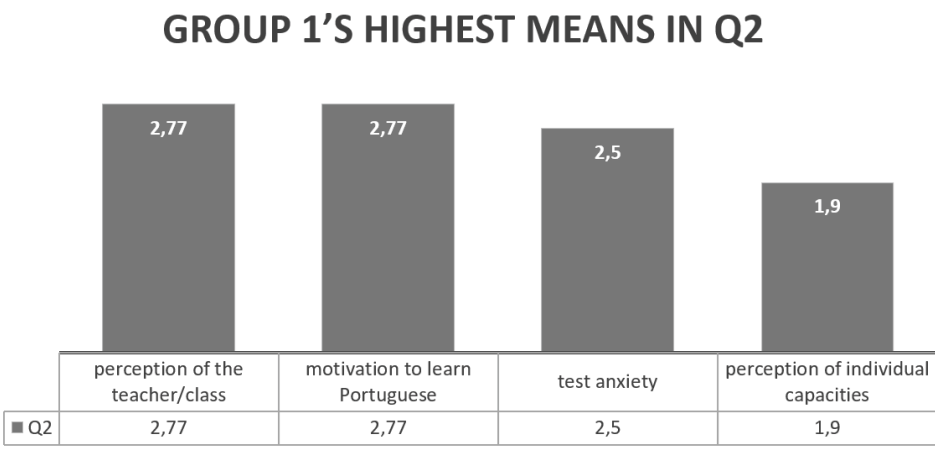
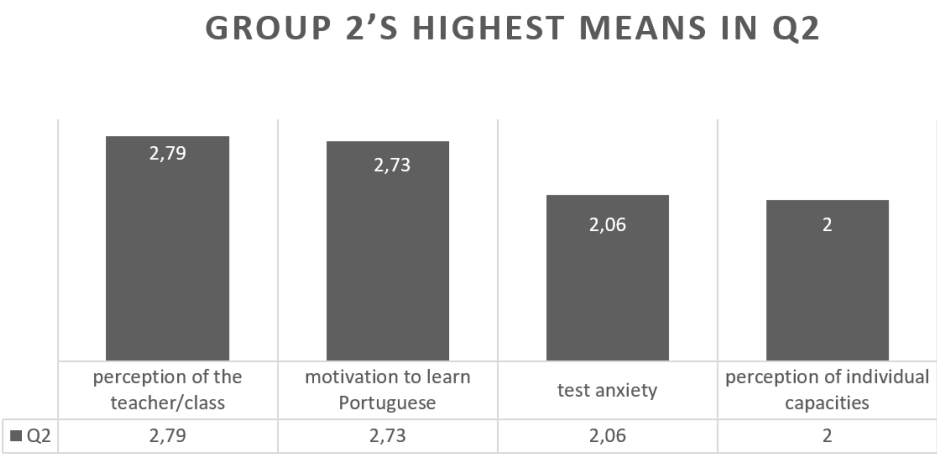



**Figure 2**

*Ranking of Means at the End of the Course in Group 2 (N = 8)*




The comparison of the mean in Questionnaire 1 (Q1) and in Questionnaire 2 (Q2) revealed that some variables behaved differently in Group 1 and in Group 2: in the former, there was a decrease in MLP (–0.22) and PTC (–0.3), and an increase in CTA (+0.21) and PIC (+0.21), denoting a drop in participants' self-esteem and a surge in their language/test anxiety. In the latter, there was a decrease in PTC (–0.37) and MLP (–0.16), as well as in PIC (–0.16) and CTA (–0.13), revealing, conversely, an increase in learners' self-esteem and less FL anxiety (see Table 1 and Figure 3).

**Table 1**

*Group 1's and Group 2's Variation of Means from the Beginning to the End of the Course*

| | GROUP 1 | | GROUP 2 | |
|---|---|---|---|---|
| Variable | Increase | Decrease | Increase | Decrease |
| PERCEPTION OF INDIVIDUAL CAPACITIES (PIC)* | +0.16 | | | –0.09 |
| COGNITIVE TEST ANXIETY (CTA)* | +0.21 | | | –0.13 |
| MOTIVATION TO LEARN PORTUGUESE (MLP) | | –0.22 | | –0.16 |
| PERCEPTION OF THE TEACHER AND THE CLASSES (PTC) | | –0.3 | | –0.37 |

*Note.* *These measures are reversed, so an increase represents a negative result in learners' self-esteem, FL classroom anxiety and test anxiety, whereas a decrease represents a positive result in these variables.

Spearman's rank coefficient results showed both in Group 1 and in Group 2 a significant positive correlation between PIC and CTA (Corr = 0.554, $\alpha < 0.05$) and a non-significant negative correlation between PIC and PTC (Corr = 0.489, $\alpha > 0.05$) and MLP (Corr = –0.505, $\alpha > 0.05$). A significant negative correlation was found between CTA and PTC (Corr = –0.635, $\alpha < 0.05$) and MLP (Corr = –0.571, $\alpha < 0.05$). Finally, a significant positive correlation between PTC and MLP was found (corr = 0.591, $\alpha < 0.05$).

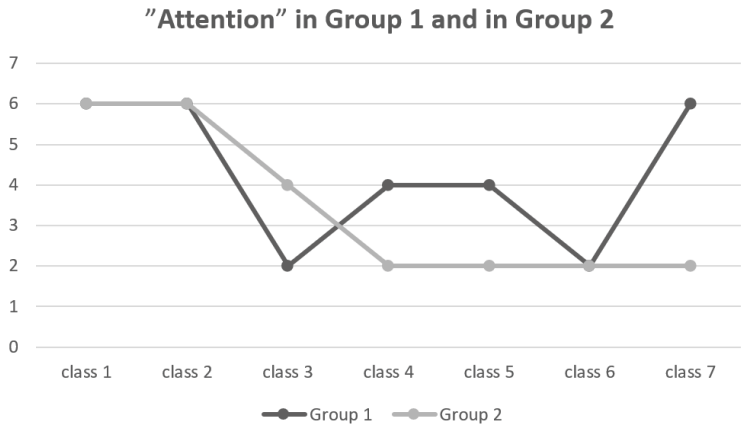
### *The Class Observation Grid*

The analysis of the learners' attention, participation, and volunteering shows the ebb and flow of behavioral engagement in both groups along the classes observed throughout the school term. Figures 4, 5, and 6 reveal that, as a general pattern, learners in Group 1 demonstrated greater behavioral engagement than those in Group 2. Each variable, however, presented a specific behavior that deserves an analysis of its own, as follows.

Starting with the learners' attention (Figure 3), it was observed that both groups start the course at a significantly high level (6), which, from class 2 onwards, drops brusquely, reaching its lowest value (2) in class 3 (Group 1) and class 4 (Group 2). From this point on, it stabilizes at this level in Group 2 and, in Group 1, increases again, to stabilize at 4 from class 5 to class 6. After a drop to level 2 in class 5, there is a peak to level 6 in class 7, where it remains. Hence, in terms of attention, Group 1 displays a more dynamic pattern, whereas Group 2 seems to have a constant downward tendency.

**Figure 3**

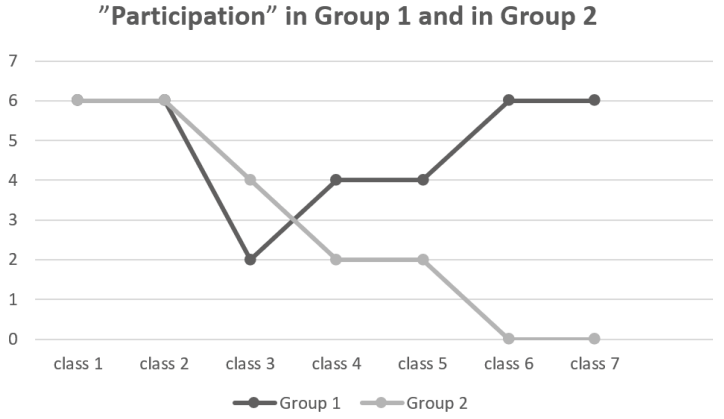
*Variation of "Attention" in Group 1 and in Group 2 along Class Observations*




Concerning the learners' participation in class (Figure 4), once again a more dynamic pattern in Group 1 combined with an uncurbed downward tendency in Group 2 is observed. Both groups start at a high level of participation at 6, then experience a fall from the second class onwards, landing at 2 in Group 1 in Class 3 and Group 2 in Class 4. Then, it goes back up in Group 1, stabilizing at 4 in classes 4 and 5, and reaching a new peak at 6, where it remains from class 6 onwards. In Group 2, however, after stabilizing at 2 in classes 4 and 5, it drops brusquely to 0, where it remains.

**Figure 4**

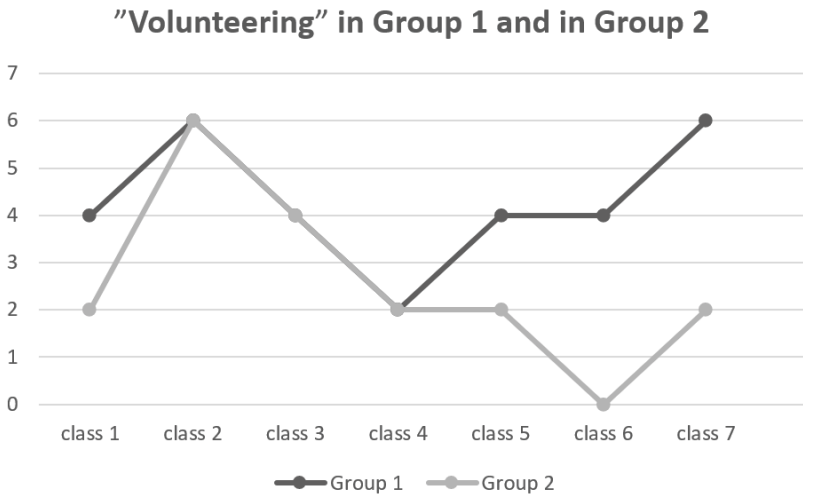
*Variation of "Participation" in Group 1 and in Group 2 along Class Observations*



Finally, regarding the learners' behavior of volunteering readily for teacher-fronted activities (Figure 5), the overall pattern observed in the two groups seems equally dynamic, although a greater downward tendency can still be noticed in Group 2 and a prevailing upward pattern is seen in Group 1. Thus, in Group 1, there is a clear upward tendency, with students gradually volunteering more in the classes, after a short period of less spontaneous participation, while in Group 2 this behavior continuously deteriorates, only to go back to the initial level in the last observed class. Hence, Group 2 starts and ends at the same point in terms of volunteering, which is a rather low level when compared to Group 1. The latter not only starts at a higher level but ends at its peak.

**Figure 5**

*Variation of "Volunteering" in Group 1 and in Group 2 along Class Observations*




In the next section, some qualitative data collected during the class observations shall be described and analyzed, attempting to explain the results found both in the questionnaire and in the class observation grid.

### *Class Observation Journal*

The content analysis of the journal used during the class observations has yielded two main categories, that is "teacher" and "learners." In the first category, the following subcategories were identified: "relationship with learners," "promoting learners' self-esteem/motivation," and "strategies, activities, materials, and topics." In the second category, learners' reactions to these subcategories were registered.

### *Teachers*

***Relationship with Learners.*** Compared to T1, T2 showed a much more affectionate attitude and relationship with her group, remarkably during the initial and final moments of the class, while greeting them and saying goodbye warmly. She was quite comfortable sharing with her group private information and made references to learners' real experiences shared in class. When she heard a meow coming from one of the computer cameras, she said, empathetically: "Oh, I had not noticed that someone else had a cat at home, too!" She also generally expressed tolerance and respect, fostering learners' autonomy and individuality. Occasionally, however, her answers to students' questions or remarks were less warm or welcoming, displaying an expression of impatience or frustration, especially when something in class did not play out as she had expected due to technological constraints.

In general, T1 only greeted students after they had greeted her, as they joined the virtual room for class. By and large, positive non-verbal communicative behaviors were predominantly observed (smiling, expressing a welcoming and friendly attitude, and laughing in class). However, occasionally she let on a relatively dull expression, showing her tiredness by laying her head upon her hand while teaching. At other times, she seemed upset by wrinkling her brow when a student asked questions.

***Promoting Learners' Self-esteem/Motivation.*** Predominantly, positive encouragement to learners' self-esteem by T1 was observed, among other things in the form of the delivery of positive feedback, remarking, for example: "It has never been so easy to assign an oral task to a group! It's been really great! Thank you!" Yet, occasional negative feedback was noticed, as was the case when T1 remarked to a student, who was taking too long to finish the assignment: "This task is so easy!" In terms of motivating learners, there was a great number of incentives for integrative motivation, illustrated by T1 bringing up cultural aspects of the L2. Motivating learners through cooperation was also a recurrent strategy, above all by encouraging their interaction in oral debates and classroom discussions.

T2 also frequently encouraged learners' self-esteem, mainly through careful positive individual comments on their performance after a presentation, something which was not seen in T1. However, T2 did not eschew negative comments. When learners' performance in an assignment was poor, for example, she told them that next time they should listen more carefully and more times to the audio. As for fostering learners' motivation, great emphasis was given to encouraging integrative motivation, by a comprehensive account of historical and cultural events in Portugal.

***Strategies, Activities, Materials, and Topics.*** T1 favored a strategy based on oral debates and classroom discussions, presenting daily expressions and other language structures in a contextualized way. The materials used in class were taken originally from real sources, such as blogs and newspaper articles. The topics covered issues that were mostly related to the routines and experiences of the learners, but also to the Portuguese culture.

T2's strategy was strongly based on having learners deliver individual oral presentations having texts or films previously assigned as prompts. Great emphasis was given to Portuguese culture and history. There were some successful experiences with the use of technology in class, but sometimes it went sideways, especially concerning the use of some materials adapted to be used online, as the teacher herself was relatively unacquainted with the digital environment. Another commonly observed practice was to discuss the course plan with the group (sometimes for several minutes), giving them the chance to collectively decide on the course plan and deadlines for the assignments.

### *Learners*

***Learners' Reactions.*** A great deal of interaction was observed between learners and T1 and among learners themselves, most of which happened during oral tasks that were based on a topic extracted from a previously read text. A considerable amount of personal information and emotion was shared during these discussions, most of which involved issues related to learners' identity. Group 1's multiculturality was highly explored, yielding lively debates about customs and cultural specificities. Some episodes of frustration and relative distress were also identified, nonetheless.

In Group 2, a significant amount of interaction was seen between the students and T2 in varied moments of the class, but considerably less communication was seen among the students themselves. Learners shared a great deal of personal information with the teacher, though. Still, occasionally, learners were unresponsive to T2's encouragement for them to volunteer their answers. There was also a particular episode in which a sheer atmosphere of boredom and lack of interest was noticed. This happened in a class in which T2 took up an excessive amount of time to discuss the course plan with the students. T2 and some students were seen on the verge of a fallout, as exemplified by student A204's remarks that she had already answered ten emails, asking T2 if they could go back to class.

***Interviews.*** The interviews focused on an array of topics related to learners' perceptions of the strategies and attitudes of the teachers, as well as on certain elements located in the classroom context. For the purposes of this study,

two main topics stand out: "perceptions of the class," and "perceptions of the relationship with the teacher." The former includes aspects such as activities, participation, and interaction, among others, and the latter comprises perceptions of the teacher's attitudes and the sharing of personal information and emotions in class.

***Perceptions about the Class.*** Participants in Group 1 were particularly fond of the oral debates, cited as the most motivating activity and an opportunity for learners to get to know each other better, getting around the constraints imposed by the ERT mode: "I had motivation when we delivered presentations and discussed, we had a class in which we presented our region or something about our country" (A106). Participants considered topics related to the culture of the L2 and their own cultural identities as particularly motivating, but the fact that not all students had the chance to speak proportionally combined with the amount of time used in these discussions earned severe criticism from some respondents.

Group 2's answers revealed a different profile, as learners reported a particular appreciation for activities that focused on grammar and vocabulary. The oral presentations were generally referred to as demotivating, as they lasted for too long, and "each person speaks individually and then it is too long for them to be able to speak again" (A206). Student A104 claimed that what she found least motivating was "those oral presentations in general, especially in front of a computer, I think they are super selfish because each student minds their own business and disregards what is going on," highlighting that these tasks could have been more motivating if they had engaged the class as a group, exploring some tools that the digital platforms provide for interaction, such as "breakout rooms."

In fact, almost all participants highlighted the lack of interaction as something negative. Student A207, for instance, stressed that what he found least positive was the context because, as he admitted: "We didn't notice many debates, we mostly just read." Student A204 added that "it was a pity not getting to know the other students better, not having had more discussions with them." In this group, the most enjoyable aspect of the course was often related not to the practice, but to the outcome of the tasks, when students realized their progress: "The strongest points of the course were the writing activities and the oral presentations. I benefited a lot from these practices. I feel that I have improved my writing skills and I feel much more confident in speaking Portuguese" (A201).

***Perceptions about the Teachers.*** When asked whether they notice when their teacher greets them before they do, participants in Group 1 provided threefold answers: one-third claimed that they do not notice and do not care if they do;

one-third informed that they do notice, but do not really care; and another one-third stated that they do notice, and that they would have appreciated it that T1 had greeted them before. As for smiling, the unanimity of the interviewees informed that they notice when teachers smile, rating it as an important aspect of teacher-learner communication. A negative aspect regarding T1's attitude was that she sometimes judged students' opinions when she should limit herself to teaching the language: "Unfortunately, whenever we spoke, she judged our way of thinking, and did not stick to correcting our Portuguese […] the class is not about my opinion!" (A102). Another negative point attributed to T1 was some impatience when learners asked her questions.

In Group 2, most of the interviewees reported that they notice when teachers smile, and a close minority claimed that they do not notice or care about that. When it comes to greetings, most of Group 2's students claimed that they notice and feel good about teachers greeting them first. An aspect that was highly criticized was T2's lack of expertise with the technology, resulting in poorly adapted materials for ERT classes. Likewise, T2's frequent complaints when something went sideways were stressed by some students. As for the fact that T2 adopted a strategy of discussing the program with the group, it was remarked that "she spent too much time discussing the activities and the 'why' of the tasks" (A206). A positive aspect that was reported was the way she presented some topics to the group. Thus, student A205 remarked: "What I enjoyed the most in the course was the class when she told us the history of the April 25th Revolution in Portugal […] we barely spoke in that class, but I think it was a good class because I noticed she was well, she was emotional and willing to share."

Concerning the exchange of personal information between teachers and learners in class, Group 1 showed a relatively more open stance than Group 2 in both directions. Most students acknowledged the importance of teachers sharing information about their lives, as this ushers in a friendly environment and contributes to topics of mutual interest. As for the importance of students sharing personal information with their teachers, most respondents stated that, if it is restricted to basic information and is relevant for the class, it can be helpful, highlighting, again, its contribution to the choice of topics. Others, still, acknowledged that this could help to set a connection between pupils and teachers.

Group 2's responses marked a less acknowledgeable position regarding the importance of teachers sharing personal information with their students and vice-versa. Student A204 was assertive in this respect: "For me, that has nothing to do with the class!", adding that "teachers should be, remain, 'teachers'… whether or not they have kids, or cats, who cares?". As for sharing their personal information with teachers, the bulk of the students restricted its relevance to the topic of the class and general information about them.

Regarding the exchange of emotions, a more equitable pattern was found between both Groups, in both directions. Hence, in Group 1, most respondents were at least partly receptive to emotional exchanges, while others highlighted the importance of being professional: "It's a job, and we all have our moments, but I think that when you get to work, all that's happened before should be left at the door, both for students and for teachers" (A106). In Group 2, the human side of teachers was stressed: "Yes, I want the teacher to feel comfortable, to know that she can share. They're people too, so, to an extent, it is (ok for them to share)" (A206). A red flag was raised to sharing negative emotions, and in Group 2 there was an explicit reference to T2's frustrations when something went wrong.

## Discussion

### *The Impact of Emotion: Motivation to Learn Portuguese (MLP), Perceptions of the Teacher and Classes (PTC), and Perceptions of Individual Capacities (PIC)*

As observed in the results of the AMBT questionnaire, both groups exhibited a highly motivational result at the end of the course. It is suggested that, to a large extent, this is because both teachers shared a lot of emotions and engaged in real communication with their learners. What is more: both positive and negative emotions were shared, with an overall positive impact in the end. This contrasts with data presented in studies like Resnik and Dewaele (2021), where ERT classes were reported as showing significantly less emotion than traditional classes.

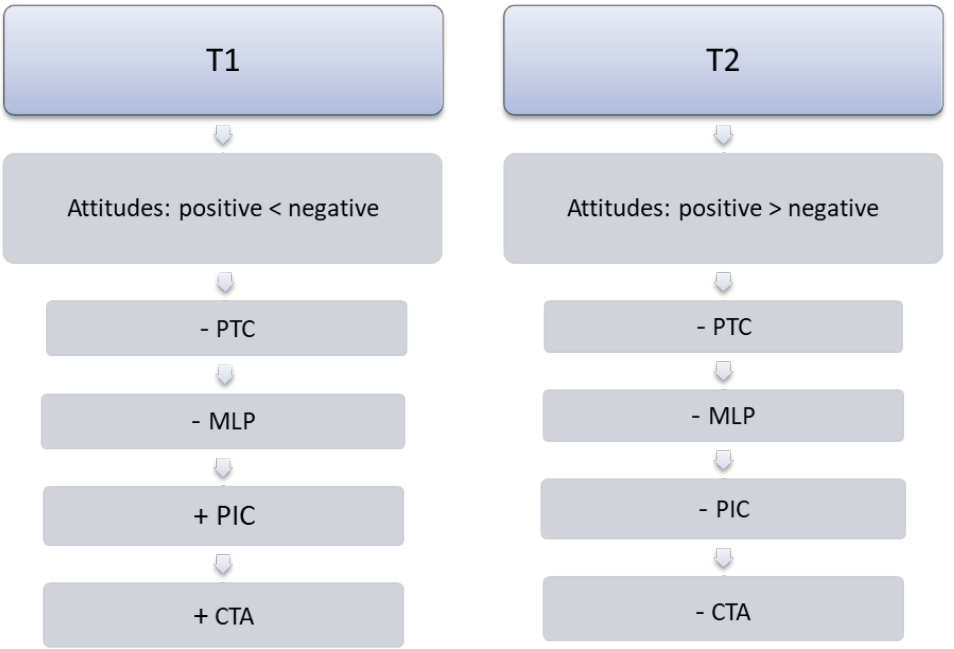
In the present study, both teachers acted as facilitators (Rogers, 1983), which does not imply a fear of expressing one's feelings and emotions (good or bad) but prescribes an awareness of how to communicate them to learners (Rogers, 1983, in Zimring, 1994, p. 4). That is exactly what the data show: they generally expressed positive emotions, but did not avert expressing tiredness, difficulties, or frustrations in a genuine and transparent way. This meets the conclusion advanced by Dewaele and MacIntyre (2014, 2016), that is, it seems more advisable to find the correct ratio between positive and negative emotions than to concentrate on eliminating the latter from the classroom.

Participants in both groups reported less motivation (MLP) and less appreciation for their respective teachers and classes (PTC) by the end of the course. As seen, despite T1's and T2's generally friendly and positive approach to students in class, they both exhibited some negative attitudes and experienced some hardships which explain the decrease in the measures of MLP and PTC. The way T1 responded to some of the students' questions was a bit distant and

sometimes revealed some impatience, as the interviews confirm, with students reporting feeling judged on their opinions rather than on the language. Her overall impersonal attitude in the early moments of the class is another aspect that did not go unnoticed: all students in T1 reported noticing and considering it important that teachers smile in class. As for T2, it was noticed that the amount of time spent on discussing the course plan and the recurrent problems faced with the online format took a definite toll on learners. The statistical tests confirm the correlation between PTC and MLP in both groups, giving further empirical validity to the data.

**Figure 6**

*The Impact of Affect in T1's and T2's Attitudes upon PTC, MLP, PIC and CTA*




The better results exhibited by Group 2 (↑PIC) are assumed to stem from the more affectionate attitude T2 maintained in class, balancing the overall negative perception students had of the class and exerting a positive effect on learners' self-esteem, classroom anxiety, and test anxiety. Conversely, T1's less affectionate attitude is assumed to account for the lower measures Group 1 exhibited in PTC, MLP, PIC, and test anxiety. This conclusion finds support in Rubio-Alcalá (2017), who claims that teachers can reduce learners' self-esteem when they express an attitude of intolerance, do not give them enough time to answer a question, and are incautious of their feedback and corrections,

among other things. Figure 6 below shows the impact of the different attitudes adopted by T1 and T2 in their respective groups.

Hence, the findings reported in previous studies in face-to-face contexts showing that teacher variables are more closely linked to enjoyment than to anxiety (Dewaele & Ferrer, 2022) are not entirely licensed in the present study. At least to some extent, learners' self-esteem and anxiety were affected by teachers' different attitudes in each group. The qualitative data support this conclusion, which is further corroborated by the statistical tests revealing a significant negative correlation between test anxiety (CTA) and PTC and MLP. Zare and Risiati (2012) inform of an association between test anxiety and the two variables in PIC (self-esteem and classroom anxiety), endorsing the conclusions from the present results.

### *The Impact of Interaction: Engagement, FLE, Flow, and FLB*

As seen, a great deal of interaction was reported in both groups. In Group 1, both a 'teacher-student' and a 'peer-peer' type were identified, based on debates which led to greater interaction, while in Group 2 the individual presentations led to mostly 'teacher-student' interaction. Group 1 excelled against Group 2 in all components of learners' behavioral engagement (attention, participation, and volunteering). Thus, it is proposed that promoting debates in the classroom yields greater engagement than individual presentations. This is also suggested to lead to more episodes of flow, as was observed in a class in Group 1 in which learners failed to notice that T1 had lost her internet connection, continuing the debate without her. In Group 2, conversely, an opposite pattern was identified, that is, while one student was delivering their presentation, other students would mind their own business, answering emails and the like.

Participants' FLE during their interactions also seemed to predict their engagement, confirming the findings of studies focusing on traditional L2 classes (Aubrey, King, & Almukhaild, 2020; Phung, 2017) and ERT classes (Mihai et al., 2022). Aubrey, King, and Almukhaild (2020) found greater engagement ensued in tasks that represented diminished anxiety and increased enjoyment, such as those that yielded exciting discussions. This seems to be exactly what happened in Group 1, where students were interactive and collaborative, illustrating what Phung (2017) defined as social engagement. Also confirmed was the impact of social support on engagement reported by Luan et al. (2020), leading to a sense of community, belonging, friendliness, and safety (Mihai et al., 2022). That was specifically stressed in some of the answers found in this study: in Group 1, such interactions were recalled with joy, pointing to FLE social; in Group 2, the lack of it was referred to with lament, with students experiencing mostly FLE private, associated with a feeling of making progress in the L2.

It seems that teachers' impact on learners' enjoyment (FLE teacher) was incidental, through the choice of topics that seemed personally relevant and that validated learners' experiences and the support needed for engaging in such tasks. Participants appreciated the sense of engaging in meaningful communication (Phung, 2017), sharing and hearing about each other's cultures and the customs of their countries, confirming that, with multicultural groups, the topic "culture" seems naturally suitable (Aubrey, Kim, & Almukhaild, 2020; Dewaele & MacIntyre, 2022).

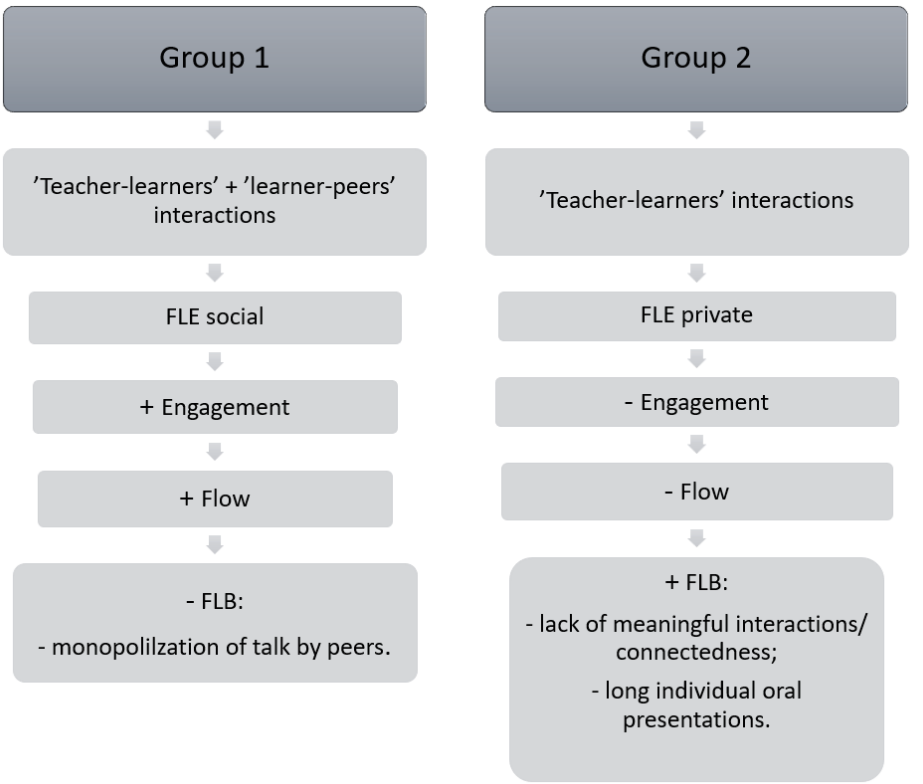
In Group 2, some students specifically criticized the lack of social contact in distal classes, which was not verified in Group 1, where plenty of interaction was afforded. This echoes the findings of Elmas and Öztüfekçi (2021), who concluded that demotivation in ERT classes is strongly linked to the methods used and the tasks assigned to participants. Similarly, Wilson and Lengeling (2021) found that learners' disengagement in class can be associated with technological obstacles and the lack of interaction. Dewaele, Albakistani, and Ahmed (2022a) reported that physical distance also curbed the chance to build meaningful relationships with classmates and the teacher, as seen in Group 2. Yet, the authors highlight that ERT classes could provide opportunities for relationship-building. We propose that the processes experienced in Group 1 illustrate the latter, making up for the lack of physical contact in ERT classes. Figure 7 illustrates the processes ensuing from the different interaction patterns in the two groups.

In short, the analysis of the patterns of interaction and their impact on each group has shown that engagement, FLE, flow, and FLB are closely linked to task and strategy choices. In classes where learners have a chance to interact and receive social support from teachers and peers, more engagement, FLE social, and the experience of flow are expected to ensue. Conversely, mostly FLE private, less engagement, and more FLB are predicted in environments that lack 'peer-peer' interaction.

Some pedagogical implications for online teaching are that teachers invest in tasks that encourage learners to embark on personally meaningful and emotionally engaging interactions. They should also provide enough social and academic support. Lack of interaction prevents the development of social cohesion and teacher-learner connection. Finally, it is proposed that students who do not engage as much in meaningful exchanges with their teachers and peers may become less sensitive to the benefits of relatedness, and, consequently, unable to account for the emotional and cognitive value of simple things, such as the discovery of a particular appreciation for cats shared by their teachers.

**Figure 7**

*The Impact of the Pattern of Interaction upon FLE, Engagement, Flow and FLB in Group 1 and in Group 2*




## Conclusions

In short, the findings in this study can be thus synthesized: the more affectionate way that T2 related to her group yielded better results in terms of motivation to learn Portuguese, self-esteem, anxiety, and test anxiety. The interactions-intensive practice set in train by T1, in turn, resulted in more engagement, FLE social, and occasional experiences of flow. Boredom was predicted by activities that did not allow for interaction to unfold, as was the case with the long individual oral presentations in Group 2.

Pedagogically, it is worth keeping in mind that any conclusion regarding what predicts learners' emotional responses in L2 classes must be interpreted with a grain of salt, regardless of the format of the classes. It seems more advis-

able that, instead of focusing on one strategy or another, teachers try to combine different practices in their classes, to account for the different emotional profiles of their groups. Martins, Silva, and Pinto (2022) make a statement for a new approach to language teaching known as a "post-method" (Kumaravadivelu, 1996, 2006), consisting of multiple approaches that complement each other, attending to all dimensions of the learner. We believe that the combination of genuine and meaningful interactions (e.g., plentiful pair work and group work using breakout rooms and whole class interactions in engaging themes) with emotional and affective support from teachers (e.g., delivering positive and individualized feedback, discussing the challenges experienced throughout the course) is one way of leading online classes on the right track.

## Limitations of the Study

These findings have encountered some limitations, however: a first shortcoming has to do with the fact that the qualitative data used for analyzing FLE and FLA could have been complemented with validated quantitative instruments, such as the scales used in Dewaele and MacIntyre (2014). Another limitation relates to the small number of participants in this study, which prevents a broad generalization of these findings, though the diversified qualitative and quantitative data analyzed do provide strong validity to the conclusions advanced herein.

## References

Altarriba, J., & Canary, T. M. (2004). The influence of emotional arousal on affective priming in monolingual and bilingual speakers. *Journal of Multilingual and Multicultural Development, 25*(2,3), 248–265.

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

Arnold-Morgan, J., & Fonseca-Mora, M. (2007). Affect in teacher talk. In B. Tomlinson (Ed.), *Language acquisition* (pp. 107–121). Continuum.

Aubrey, S., King, J., & Almukhaild, H. (2020). Language learner engagement during speaking tasks: A longitudinal study. *RELC Journal, 53*(3), 1–15.

Deci, E., & Ryan, R. (1995). Human autonomy: The basis for the self-esteem. In M. Kernis (Ed.), *Efficacy, agency and self-esteem* (pp. 31–49). Plenum Press.

Dewaele, J.-M. (2011). Reflections on the emotional and psychological aspects of foreign language learning and use. *Anglistik: International Journal of English Studies, 20*(1), 23–42.

Dewaele, J.-M. (2023). Psychology of language learning: Personality, emotion and motivation. In L. Wei, Z. Hua, & J. Simpson (Eds.), *Routledge handbook of applied linguistics* (pp. 374–385). Routledge.

Dewaele, J.-M., Albakistani, A., & Ahmed, I. (2022a). Levels of foreign language enjoyment, anxiety and boredom in emergency remote teaching and 'in-person' classes. *The Language Learning Journal*, *52*(1), 117–130. https://doi.org/10.1080/09571736.2022.2110607

Dewaele, J.-M., Albakistani, A., & Ahmed, I. (2022b). Is flow possible in the emergency remote teaching foreign language classroom?. *Education Sciences, 12*(7), 444. https://doi.org/10.3390/educsci12070444

Dewaele, J.-M., & Ferrer, M. (2022). Three entangled foreign language learner emotions: Anxiety, enjoyment and boredom. *Elia: Estudios de Lingüística Inglesa Aplicada*, *22*, 237–250.

Dewaele, J.-M., & MacIntyre, P. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies of Second Language Learning and Teaching, 4*(2), 237–274.

Dewaele, J.-M., & MacIntyre, P. (2016). Foreign language enjoyment and foreign language classroom anxiety. The right and left feet of FL learning? In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215–236). Multilingual Matters.

Dewaele, J.-M., & MacIntyre, P. (2022). You can't start a fire without a spark. Enjoyment and the emergence of flow in foreign language classrooms. *Applied Linguistics Review*. https://doi.org/10.1515/applirev-2021-0123

Dewaele, J.-M., Magdalena, A., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal, 103*(2), 412–427.

Dörnyei, Z. (Ed.). (2001). *Motivational strategies in the language classroom*. Cambridge University Press.

Dörnyei, Z. (2009). The psychology of Second Language Acquisition. Oxford University Press.

Elmas, E., & Öztüfekçi, A. (2021). L2 demotivation in online classes during Covid-19: From an Activity Theory perspective. *International Journal of Education, 9*(3), 72–78.

Fraschini, N., & Tao, Y. (2021). Emotions in online language learning: Exploratory findings from an *ab initio* Korean course. *Journal of Multilingual and Multicultural Development, 45*(5), 1305–1323. https://doi.org/10.1080/01434632.2021.1968875

Garret, P., & Young, R. (2009). Theorizing affect in foreign language learning: An analysis of one learner's responses to a communicative Portuguese course. *The Modern Language Journal, 93*(2), 209–226.

Guilloteax, M., & Dörnyei, Z. (2008). Motivating language learners: A class-oriented investigation of the effects of motivational strategies on student motivation. *TESOL Quarterly, 42*(1), 55–77.

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, M. (March 27, 2020). The difference between emergency remote teaching and online teaching. *Educause Review*. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning

Horwitz, E., Horwitz, M., & Cope, J. (1986). Foreign Language Classroom Anxiety. *The Modern Language Journal, 70*(2), 125–132.

Li, C., & Dewaele, J.-M. (2020). The predictive effects of trait emotional intelligence and online learning achievement perceptions on foreign language class boredom among Chinese university students. *Foreign Languages and their Teaching, 5*, 33–44.

Li, C., Dewaele, J.-M., & Hu, Y. (2021). Foreign Language Learning Boredom: conceptualization and Measurement. *Applied Linguistics Review, 11*, 485–511.

Luan, L., Hong, J-C., Cao, M., Dong, Y., & Hou, X. (2020). Exploring the role of online EFL learners' perceived social support in their learning engagement: a structural equation model. *Interactive Learning Environments, 31*(3), 1703–1714.

Martins, C., Silva, F., & Pinto, J. (2022). O que ainda não sabemos e precisávamos de saber para o ensino do Português Língua Não-Materna. *Revista da Associação Portuguesa de Linguística*, *8*, 9–10, 301–313.

Mihai, M., Albert, C., Mihai, V., & Dumitras, D. (2022). Emotional and social engagement in the English language classroom for higher education students in the COVID-19 online context. *Sustainability, 14*, 4527. https://doi.org/10.3390/su14084527

Moskowitz, S., Dewaele, J.-M., & Resnik, P. (2022). Beyond the boundaries of the self: applying Relational Theory towards an understanding of teacher-student relationship as a driver of motivation in foreign language learning. *Journal of the Psychology of Language Learning 4*(2), 1–17.

Pekrun, R. (2006). The Control-Value Theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychology Review, 18*, 315–341.

Phung, L. (2017). Task preference, affective response, and engagement in L2 use in a US university context. *Language Teaching Research, 21*(6), 751–766.

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, 14*(3), 473–501.

Resnik, P., Knechtelsdorfer, E., & Dewaele, J.-M. (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.

Rogers, C. (1983). *Freedom to learn for the 80s*. Charles E. Merrill Publishing.

Rodrigo-Ruiz, D. (2016). Effect of teachers' emotions on their students: some evidence. *Journal of Education & Social Policy, 3*(4), 73–79.

Rubio-Alcalá, F. (2017). The links between self-esteem and language anxiety and implications for the classroom. In C. Gkonou, M. Daubney, J.-M. Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 198–216). Multilingual Matters.

Wang, Q., & Jiang, Y. (2022). A positive psychology perspective on positive emotion and foreign language enjoyment among Chinese as a Second Language learners attending virtual online classes in the emergency remote teaching context amid the Covid-19 pandemic. *Frontiers in Psychology, 12*, 798650.

Wilson, A., & Lengeling, M. (2021). Language learning in the time of Covid-19: ELT students' narrated experiences in guided reflective journals. *Íkala: Revista de Lenguaje y Cultura, 26*(3), 571–585.

Zare, P., & Risiati, M. (2012). The relationship between language learning anxiety, self-esteem and academic level among Iranian EFL learners. *Pertanika Journal of Social Sciences & Humanities, 20*(1), 219–225.

Zheng, Y. (2010). *Chinese University Students' Motivation, Anxiety, Global Awareness, Linguistic Confidence, and English Test Performance: A causal and correlational investigation* (Doctoral Dissertation). Available from ProQuest dissertation and theses database (Document ID 28389413).

Zimiring, F. (1994). Carl Rogers (1902–1987). *Prospects: The Quarterly Review of Comparative Education, 24*(3/4), 411–22.

Appendix A

# Extract from the Classroom Observation Grid for Assessing Learners' Motivated Behavior in Online Classes

| Variable | Description | 1–2 times | 3–4 times | 5 or + times |
|---|---|---|---|---|
| Attention | Students appear to be paying attention. They are not displaying any inattentive or disruptive behavior: they keep their cameras on during most of the class; they do not turn their heads to gaze out on their surroundings or to speak to someone near them; their eyes are staring at the computer screen as the teacher or their colleagues speak or as materials are being screenshared; they offer appropriate non-verbal responses, by nodding their heads, for example. | | | |
| Participation | Students are actively taking part in classroom interaction or working on assigned activity | | | |
| Volunteering for teacher-fronted activity | At least one third of the students are volunteering without the teacher having to coax them in any way. | | | |

Source: adapted from the *MOLT Classroom Observation Scheme* (Guilloteaux & Dörnyei, 2008, p. 62).

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/30
https://doi.org/10.31261/TAPSLA.16033

CC BY SA

**Jasrael D. Stokes**
University of Bristol, UK
https://orcid.org/0009-0004-1331-070X

# The Impact of Foreign Language Anxiety on the Utterance Fluency of International Students in the UK: Does the Presence of an Audience Matter?

Abstract

Foreign language anxiety (FLA) can have a detrimental impact on language production, which can have dire consequences for students completing a degree in a language other than their mother tongue. Studies have shown speakers to cope with FLA by using time-gaining mechanisms, manifesting in utterance disfluencies. Research exploring the mechanisms in which FLA impacts oral fluency measures is limited particularly among learners completing a degree abroad immersed in the target language, where FLA can impact life both inside and outside of the classroom. The current study investigated the impact of FLA on the English utterance fluency of international students undertaking a degree in the UK and examined how the presence of native-speaker listeners affected this relationship. After the completion of questionnaires regarding FLA, the participants took part in an online Zoom call and completed a monologue-style speaking task. They were randomly allocated to either a control condition or an experimental condition, which had four native-speaker listeners in the audience. Speech samples for the speaking task were then analyzed for utterance fluency variables. The results showed that the participants scoring higher on FLA produced a significantly higher percentage of utterance disfluencies while speaking. Contrary to expectations, the participants allocated to the control condition were no less anxious than participants in the experimental condition, indicating possibly high FLA merely due to the requirement of speaking. The findings of this study provide important implications for educators in understanding how FLA may impact students' utterance fluency when required to speak in class.

*Keywords*: foreign language anxiety, utterance fluency, international students, study abroad

Foreign language anxiety is a widely studied emotion in the second language acquisition literature. Converging research shows lower levels of FLA to be strongly associated with higher levels of second language performance as defined by course grades (Bekleyen, 2009; MacIntyre & Gardner, 1994b) essay writing (Woodrow, 2011), and listening tests (Zhang, 2013). The majority of FLA research has focused on students learning a foreign language in a classroom in their home country (MacIntyre & Gardner, 1991; Zhang, 2013), or students studying abroad with the purpose of learning the language of the host country (Dewey, Belnap, & Steffen, 2018; Liu, 2018; Thompson & Lee, 2014), but there is limited research examining the impact of FLA on students studying abroad long term undertaking a degree in their target language.

For international students from non-English speaking backgrounds undertaking a degree in an English-speaking country, English is not only used as the medium of study but is also necessary for the students to be able to express themselves, understand and cooperate with people on a daily basis, effectively navigate personal problems, and be self-determining on their own behalf (Sawir et al., 2012; Tran & Pham, 2016). International students with higher English abilities are more likely to have higher levels of academic achievement (Mamiseishvilli, 2012; Martirosyan, Hwang, & Wanjohi 2015; Neumann, Padden, & McDonough, 2019), lower levels of homesickness (Poyrazli & Lopez, 2007), and a greater sense of wellbeing (Basow & Gaugler, 2017). For international students who must communicate using their target language both in and outside of the classroom for an extended length of time, FLA may greatly impact their experience abroad and influence the degree to which they are able to successfully navigate not only their academics, but also everyday life. This study examines the impact of FLA on the English oral fluency of international students undertaking a degree in the UK in order to elucidate the effects it may have on students' ability to communicate.

## Literature Review

Foreign language anxiety (FLA) is a situation-specific anxiety referring to worry, nervousness, and apprehension when communicating in one's second language (Horwitz, Horwitz, & Cope, 1986). FLA has been shown to have deleterious effects on language communication, as the speaker focuses on reasons for being anxious and is distracted from the main task at hand (Kormos, 2006). International students who experience FLA may use avoidance to cope with anxiety and segregate themselves to socialize and form groups with fellow nationals, which limits opportunities to practice the language (Brown, 2008;

Sawir et al., 2012). Despite meeting university language requirements, international students often report high levels of anxiety about their English communication abilities while studying in the UK (Brown, 2008; Sawir et al., 2012). As individuals experiencing FLA are more likely to have lower grades (Bekleyen, 2009; Ganschow & Sparks, 1996), take longer to complete tests (MacIntyre & Gardner, 1994b), and avoid attempting personal or difficult messages in their second language (Horwitz et al., 1986), international students undertaking a degree may struggle when completing coursework, exams, dissertations, or conducting meetings and discussions with classmates or supervisors.

As fear of negative evaluation is one of the main components of FLA (Horwitz et al., 1986), and students typically feel the most anxiety when speaking with strangers or speaking in front of others (Bekleyen, 2009; Birney et al., 2020; Dewaele, 2007), international students are likely to experience high levels of anxiety when required to partake in classroom activities and speak in front of classmates. Furthermore, research has suggested that non-native speakers tend to experience anxiety and fear of being ridiculed particularly by native speakers, even in an online environment (Lin, 2022; Russel, 2020). International students have also reported that feeling anxiety while speaking to locals leads them to pause mid-sentence, having to rephrase, or ask others to repeat sentences (Brown, 2008). International students undertaking a degree in the UK will not only have to use English in front of other learners, but also in front of native English speakers both in and outside of the classroom, increasing the likelihood of experiencing high levels of FLA in a variety of situations, and in turn experience barriers to oral communication and navigation of daily life.

Anxiety can impede one's ability to clearly communicate their ideas, as task irrelevant thoughts and ruminations about reasons for anxiety compete with thoughts relevant to the task demands (Kormos, 2006; Trebtis, 2014). When producing speech, the speaker is required to simultaneously plan their message and map their thoughts onto the appropriate language. This may be relatively automatic in one's mother tongue but require more effort in a non-native language (Kormos, 2006). Mutual communication requires quick listening comprehension and a relatively immediate response, so there are potentially high demands on international students who are asked to speak, answer questions, provide opinions, or participate in discussions in class when English is not their native language. These contexts may place students in a situation under pressure with limited opportunity to allocate extra time for comprehension and response. Strategies such as allocating more time to the task, reviewing material, and articulatory rehearsal, though effective and practical in reading or writing contexts (Bekleyen, 2009; Eysenck et al., 2007), may not be available in situations such as attending lectures, participating in discussions during seminars, or communication in daily life. Instead, anxiety may lead to the individual resorting to time-gaining mechanisms during speech such as pausing or repeating

a message (Castillejo, 2019; Gots, 2013). Furthermore, in Castillo's (2019) study, anxious participants paused even more frequently than low-proficiency learners during an oral exam, emphasizing the fact that anxiety may disadvantage even high-proficiency learners' ability to speak fluently. This highlights the potential that FLA may have on the speakers' ability to smoothly communicate their intended message regardless of their actual language knowledge.

In the same light, international students who are required to use English in the classroom and speak in front of the class may struggle due to their experiences with FLA. For example, when required to discuss classroom material and provide answers or opinions to questions in front of others, a student with high FLA juggles thoughts related to their reasons for anxiety and these thoughts may impede their ability to focus their attention on listening to others or preparing for a response, hindering their ability to smoothly communicate. These students eventually may decide to avoid participating and communicating in English all together (Brown, 2008; Sawir et al., 2012). For those who do persevere and participate, these struggles may manifest in the form of utterance disfluencies during speech. Utterance fluency refers to oral features that can be measured and focuses on how fluid one's speech is (Skehan, 2003; Tavakoli, 2005). This includes the extent to which the speech is interrupted, as well as the number of self-corrections and repetitions in speech. In addition to pausing (Castillejo, 2019), international students may have difficulty with word-finding while distracted by task-irrelevant thoughts, and therefore show a higher number of self-repairs when speaking (Zuniga & Simard, 2022). Utterance disfluencies such as these may act as coping mechanisms for speakers to bide time until they are able to retrieve the correct grammar or vocabulary required to complete their message.

Thus far, much of the FLA literature focuses on general outcome variables such as grades or tests (Bekleyen, 2009; Cheng, Horwitz, & Shallert, 1999; MacIntyre & Gardner, 1994b; Linck & Weiss, 2015; Liu, 2018; Saito, Horwitz, & Garza, 1999; Verhagen & Leseman, 2016), and there is limited research examining the impact of FLA on utterance fluency (Bielak, 2022, Castillejo, 2019, 2021), particularly among individuals who are immersed in the target language rather than inside the language classroom. Examining measurable fluency outcomes may provide a clearer picture as to how the effects of FLA manifest in speech. Furthermore, examining utterance fluencies in an ecologically valid context similar to situations international students typically experience may elucidate the way FLA impacts oral fluency in real life communication.

In addition, in order to understand how FLA impacts students' oral fluency in a real-life setting, it is important that the methodology used to investigate the impact of FLA is similar to what students would experience in an educational environment. Much of the previous research on FLA has used cross-sectional (Bekleyen, 2009; Bielak, 2022; Castillejo, 2019; Liu, 2018) or reflective inter-

view (Brown, 2008; Gregersen, Meza, & MacIntyre, 2014; Sawir et al., 2012) methods, which although provide a strong rationale for establishing the negative effects of FLA, they do not reveal the distinct ways FLA can impact students on a day-to-day basis. Some studies have, however, used experimental methods, such as the study by Rai et al. (2001), which showed how FLA can negatively impact reading comprehension by leading participants to require more time to process material. Furthermore, other studies have attempted to use ecologically valid settings to investigate how students' experience FLA while doing a presentation in the classroom (Gregersen, Meza, & MacIntyre, 2014).

Gregersen, Meza, and MacIntyre's (2014) study investigated the impact of anxiety on presentation performances among students learning Spanish. These performances were videorecorded, and the participants were able to rewatch the videos and indicate their levels of anxiety as the presentation went on, highlighting how students experience FLA during typical classroom assignments. The results showed that those with high FLA continued to be anxious throughout the entire presentation, while those with low FLA had decreasing anxiety over time, suggesting that those with high FLA struggle to cope with their anxiety during performance. However, it is unclear how this struggle influenced actual speech and utterance fluency during the presentation, and whether their performance had indeed been negatively affected. Considering the results of the aforementioned study combined with research strongly suggesting that FLA impacts performance in terms of grades and tests (Bekleyen, 2009; Ganschow & Sparks, 1996; MacIntyre & Gardner, 1994b), because international students are also required to communicate while living abroad, understanding how oral performance is affected by FLA and the negative impacts it can have on student classroom participation is vital.

The current study intends to combine both measurable fluency outcomes with ecologically valid settings to investigate how FLA impacts international students' ability to communicate their ideas within a classroom or small group setting, specifically when speaking in front of others. The study will use experimental methods to investigate the impact of FLA on utterance fluency in an online seminar style situation. As oral communication in English is unavoidable for international students undertaking a degree and living in the UK, examining the impact of FLA on utterance fluency variables may shed light on the impact FLA has on real-life communication events and how it may impact students' experience participating in classroom activities abroad.

## Current Study

The aim of this study was to investigate the effect of FLA on international students' English oral fluency when they are required to speak in front of others as is often the case in a classroom situation. The study used experimental methods to examine the impact of international students' FLA on utterance fluency variables when speaking in front of native speakers through the use of an audience in an online Zoom call. As the COVID-19 pandemic led to a widespread increase of technology and online education (Burns, 2020; Moorhouse, 2023), and considering that students report experiencing FLA and fear of negative evaluation among native speakers (Lin, 2022) even in an online environment (Russel, 2020), speaking to others through Zoom can be considered an ecologically valid method of examining oral communication. Being aware of the impact of the COVID-19 pandemic and the possible ways it may have affected the international student participants, this study will not be an accurate reflection of typical international students studying in higher education but provide a unique understanding of the enduring impact the pandemic and societal situation has had on the international students experiencing it. In an attempt to fill a gap in research by examining the impact of FLA on the oral fluency of international students undertaking a degree abroad, the following research questions were raised:

1. How does foreign language anxiety affect international students' utterance fluency during a speaking task?
2. How does the presence of native English speakers impact international student anxiety and utterance fluency during a speaking task?

Regarding the first research question, it was expected that the participants with higher levels of anxiety would be correlated with a higher percentage of utterance disfluencies during speech, suggesting that anxiety leads to difficulty focusing on the task at hand and the use of time-gaining mechanisms. As for the second research question, it was expected that the participants required to speak in front of native English speakers would experience higher levels of anxiety, and in turn produce a higher percentage of utterance disfluencies in their speech compared to the participants who did not have a native English speaker audience.

# Methods

In order to address the research questions, this project took an experimental between-participants design to examine differences in international students' English utterance fluency while partaking in an oral communication task among while either allocated to an experimental group, which contained an audience of four confederate listeners, or a control group without an audience. The experiment took place entirely online. The participants were recruited through online social media platforms and university mailing lists. They were offered a £10 voucher for participation. The participants for this study were international students currently attending higher education at a university in the UK. They were also required to have learned English as a second or additional language and not have had previous education in another English-speaking country.

A total of 72 undergraduate ($N$ = 10) and postgraduate ($N$ = 62) international students from multiple universities across the United Kingdom participated in this study. The participants were from 34 different countries, including China ($N$ = 14), Hong Kong ($N$ = 6), India ($N$ = 5) Pakistan ($N$ = 4), Italy ($N$ = 3), Iran ($N$ = 3), Chile ($N$ = 3), Malaysia ($N$ = 2), Germany ($N$ = 2), Romania ($N$ = 2), Indonesia ($N$ = 2), Bangladesh ($N$ = 2), Peru ($N$ = 2), Colombia ($N$ = 2), Brazil ($N$ = 2), Denmark ($N$ = 1), Russia ($N$ = 1), Uruguay ($N$ = 1), Costa Rica ($N$ = 1), Japan ($N$ = 1), Thailand ($N$ = 1), Slovakia ($N$ = 1), Kenya ($N$ = 1), *N*orway ($N$ = 1), Ethiopia ($N$ = 1), Lithuania ($N$ = 1), Taiwan ($N$ = 1), Poland ($N$ = 1), Greece ($N$ = 1), Mexico ($N$ = 1), Saudi Arabia ($N$ = 1), Spain ($N$ = 1), and Egypt ($N$ = 1). Among these participants, there were 25 different native language backgrounds, including Chinese ($N$ = 22), Spanish ($N$ = 11), Hindi ($N$ = 5), Urdu ($N$ = 4) Italian ($N$ = 3), Romanian ($N$ = 2), German ($N$ = 2), Persian ($N$ = 2), Indonesian ($N$ = 2), Arabic ($N$ = 2), Bengali ($N$ = 2), Portuguese ($N$ = 2), Japanese ($N$ = 1), *N*orwegian ($N$ = 1), Farsi ($N$ = 1), Russian ($N$ = 1), Greek ($N$ = 1), Lithuanian ($N$ = 1), Malay ($N$ = 1), Thai ($N$ = 1), Danish ($N$ = 1), Slovak ($N$ = 1), Afaan Oromo ($N$ = 1), Polish ($N$ = 1), and Kiswahiki ($N$ = 1). A full account of the demographic information about the participants can be found in Table 1 and Table 2.

**Table 1**

*Descriptive Statistics*

| Variable | Min | Max | Mean | SD |
|---|---|---|---|---|
| Age | 19 | 43 | 27.29 | 4.80 |
| Age of acquisition | 2 | 29 | 7.64 | 4.26 |
| Self-rated English proficiency | 2 | 7 | 4.68 | 1.46 |

**Table 2**

*Descriptive Frequencies*

| Variable | Frequency | Percent |
|---|---|---|
| Gender | | |
| Female | 52 | 72.2 |
| Male | 20 | 27.8 |
| Year in Uni | | |
| Undergrad 1st | 3 | 4.2 |
| Undergrad 2nd | 3 | 4.2 |
| Undergrad 3rd | 2 | 2.8 |
| Undergrad 4th | 2 | 2.8 |
| Master's | 34 | 47.2 |
| PhD 1st | 12 | 16.7 |
| PhD 2nd | 9 | 12.5 |
| PhD 3rd | 4 | 5.6 |
| PhD 4th | 3 | 4.2 |
| Length in UK | | |
| 0–6 months | 12 | 16.7 |
| 7–11 months | 29 | 40.3 |
| 1–2 years | 18 | 25.0 |
| 2–3 years | 3 | 4.2 |
| 3–4 years | 8 | 11.1 |
| 5+ years | 2 | 2.8 |
| Percentage of English used per day | | |
| 0% | 0 | 0 |
| 10% | 5 | 6.7 |
| 20% | 5 | 6.7 |
| 30% | 5 | 6.7 |

*Table 2 continued*

| Variable | Frequency | Percent |
|---|---|---|
| 40% | 5 | 6.7 |
| 50% | 16 | 21.3 |
| 60% | 7 | 9.3 |
| 70% | 9 | 12.0 |
| 80% | 8 | 10.7 |
| 90% | 11 | 14.7 |
| 100% | 1 | 1.3 |

## Measures

### *General Demographics*

A general demographics questionnaire (See Appendix A) asked the participants about their current age, gender, first language, age at which they began learning English, name of university currently attending, year in university, length of time living in the UK, whether they had a history of speech disorders, and self-reported English proficiency. The respondents were asked to rate their English reading, writing, speaking, and listening proficiency on a seven-point Likert scale ranging from "poor" to "native like," a scale used in other studies investigating variables related to English proficiency (Kim & Cha, 2017; Thompson & Lee, 2014).

### *Anxiety*

There were three separate measures used to address the participants' anxiety. The first two were questionnaires, one focusing on foreign language anxiety within the classroom, and the second one focused on foreign language anxiety outside of the classroom in day-to-day life. As research has shown that those communicating in a foreign language may experience anxiety both in and out of the classroom (Brown, 2008; Sawir et al., 2012), both measures were used in combination to get a grasp of international students' experience of foreign language anxiety. The final measure of anxiety was used as a manipulation check to make sure the experimental condition of the study did indeed induce anxiety compared to the control condition.

***Foreign Language Classroom Anxiety Scale.*** The Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al. 1986; See Appendix B) was used to assess the participants' foreign language anxiety in the classroom and adapted to suit international students partaking in classes with English as a medium rather than learning English. The questions on the FLCAS are answered on a five-point Likert scale ranging from "strongly disagree" to "strongly agree." There are 33 items in total, with a minimum of 33 and maximum of 231. The questions ask the respondents about their feelings of nervousness, worry, and anxiety in a classroom where they are using English, or situations where the use of English is necessary.

***Foreign Language Anxiety in Formal Contexts Scale.*** To investigate foreign language anxiety in situations where international students must communicate in English outside of the classroom, the Foreign Language Anxiety in Formal Contexts Scale (FLAFS) was used (Gargalianou et al., 2016; See Appendix C). These items reflect anxiety used in general contexts outside of the language classroom and can be applicable to multiple contexts where the foreign language must be used, including business settings, or communicating with friends. There are ten items in the scale, with two items measuring each the degree of anxiety, extent of understanding, fear of making mistakes, feeling of competence, and divergence from general communication apprehension. The items are on a 7-point Likert scale ranging from strongly disagree to strongly agree. The minimum score is ten while the maximum score is 70.

***Anxiety Check.*** To check the effectiveness of the experimental condition of this study, a single item measure of anxiety was used. This item was presented after the participants were provided the instructions for the speaking task and directly before they were asked to partake in the task. They were then asked, "How anxious do you feel about the speaking task?" The participants were instructed to indicate their current level of anxiety on a scale from 1 (not at all) to 100 (extremely). The purpose of using a quick, one question anxiety check was to avoid any interference, alteration of anxiety levels, or distraction from to lengthy questionnaires, as manipulation checks which distract the participant from the experimental manipulation may affect the results of the experiment (Hauser, Ellsworth, & Gonzales, 2018).

### *Speaking Task*

In an attempt to enhance the ecological validity of this study, the participants completing this task were asked a question in order to elicit free speech in the form of a monologue under conditions similar to situations in which students are asked to provide an answer to a question in a seminar.

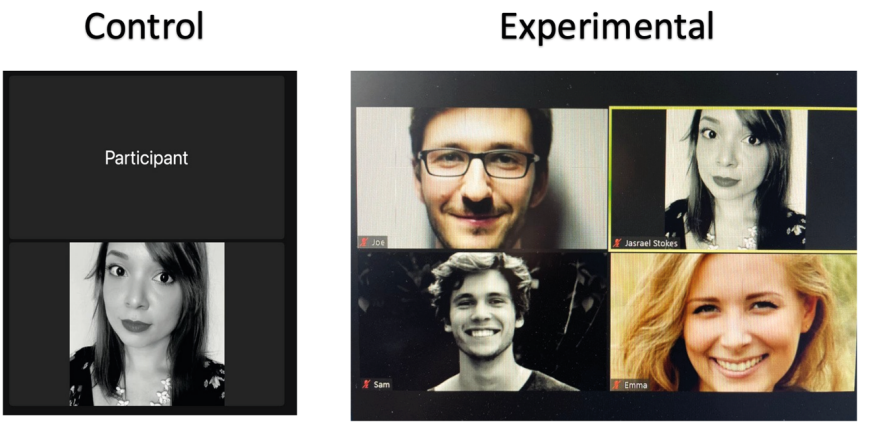
Monologues are useful as they avoid the variability of the interactions and the influence that the interlocutor introduces to the task (Segalowitz, 2016). Therefore, to avoid confounding variables from an interlocutor, interview style speaking tasks used by some researchers analyzing fluency (Cadena-Aguilar, Ortega-Cuellar, & Cadena-Aguilar, 2019; Gagne, French, & Hummell, 2022; Georgiadou & Roehr-Brakin, 2017) were avoided. Instead, the respondents were asked a question without visual prompts requiring spontaneous speech, as may be the case while participating in seminar activities. The question, "From the perspective of an international student, what advice would you give to future international students planning to study in the UK for a degree in higher education in your department?" was chosen as it was relevant to the university experience and was something that each international student could answer regardless of their background. During the time this study was conducted, students were familiar with online video platforms and the use of breakout rooms for participating in small group discussions due to the COVID-19 pandemic (Beason-Abmayr, Caprette, & Gopalan, 2021; Burns, 2020; Moorhouse, 2023).

***Control Condition.*** In the control condition, the participants were merely provided with the question they were required to answer. The experimenter was the only other person in the Zoom call when the participant performed the speaking task. This condition was intended to allow the participants to perform the task naturally without any additional anxiety-inducing stressors.

***Experimental Condition.*** The participants allocated to the experimental condition were told that there would be an audience listening to their response as they completed the speaking task in the Zoom call. For those in this condition, in addition to the researcher, there were three fake Zoom profiles created as the audience. The experimenter used multiple devices to log in to these three accounts and join the Zoom call for the speaking task. These fake listeners had profile photos taken from free stock image websites, and names were created for them that reflected common names of British native English speakers, such as Emma, Joe, and Sam (See Figure 1). This was intended to induce anxiety among the participants, as speaking to or performing in front of others, especially with strangers or native speakers, has been said to be the most anxiety provoking aspect of second language communication (Dewaele, 2007; Bekleyen, 2009; Lin, 2022).

**Figure 1**

*Control and Experimental Conditions from the Participants' view on Zoom*




## Design and Procedure

The study was conducted entirely online through Qualtrics (Qualtrics, Provo, UT; https://www.qualtrics.com). It was conducted in two parts, which included the series of questionnaires, including the demographic and anxiety questionnaires, and an online speaking task through Zoom. The participants first completed the questionnaires. After completion, they were directed to a link which opened a calendar page where they were able to book a time to partake in the second part of the study, the speaking task. Ten minutes before their scheduled time, the participants received a link to a Qualtrics page explaining what would happen during the speaking task. The respondents were allocated to either the control or experimental group and read about whether they had an audience before completing anxiety check measure. Upon completion, they were directly sent to the Zoom call.

Once the participants entered the Zoom call, they were asked to keep their cameras off, but microphones on in order to protect their anonymity. Furthermore, the lack of camera use reflected typical behavior of students during the pandemic and would avoid additional anxieties due to concern about physical location or appearance, or weak internet connection (Castelli & Sarvary, 2020). When the participants were in the Zoom call, they were then reminded of the question they were required to answer. As oral performance may be affected by the amount of planning allocated (O'Loughlin, 1995), the respondents were given two minutes to plan their answer before they were required to speak (Tavakoli, 2011; Trebtis, 2016). Planning and thinking time

were also expected to reflect real-life situations during seminars. The content of the seminar and overview of what will be discussed is often provided before the actual seminars take place, giving students time to consider the subject matter before partaking in any activities. The two-minute planning time was provided as a substitute for this. When providing their answer, the participants were asked to speak for at least two minutes. Once the preparation time was over, the researcher then asked permission to begin the recording. After the recording started, the participants were asked to provide their response. This recording was then downloaded and transcribed for analysis. Once the participants completed the speaking task, they were debriefed and able to end the call and exit the experiment.

### Tools for Transcription

The recordings of the narrative tasks were transcribed through Codes for the Human Analysis of Transcripts (CHAT) and analyzed through Computerized Language Analysis (CLAN) both of which are free downloadable software programs used for language analyses developed by (MacWinney, 2000). These softwares were used to assess the participants' speech sample for utterance fluency measures, including breakdown and repair fluencies. Breakdown fluencies indicate a lack of delay in speech and include filled or silent pauses. Repair fluencies include non-communicative words such as self-repetitions and self-corrections. These disfluencies have been shown to be valid indicators of assessing language fluency and have been widely used in fluency literature (Bosker et al., 2012; Castillejo, 2019; Cucciarini, Strik, & Boves, 2000; Kessler, 2010; Segalowitz, 2010, 2016; Tavakoli & Skehan, 2005; Zuniga & Simard, 2022).

## Results

Quantitative analyses were conducted through IBM SPSS 29 to explore the demographic variables in the participant population and examine the relationships among the variables of interest in this study. First, an exploration of the demographic variables was conducted using correlation analyses. Then, differences in demographic variables and anxiety between participants in terms of their language background were examined using one-way ANOVA. For the main analyses, correlation analyses were conducted to examine relationships among the variables of interest, and t-tests were used to compare differences between the experimental and control groups.

## Exploration of Participant Demographic Variables

Before beginning the main analysis, to investigate the general relationship between international student anxiety and their experience using English as an additional language, an analysis of the relationship between the FLAFS, FLCAS, and demographic variables was conducted. First, correlations between both anxiety measures and participant age, age of acquisition, percentage of English use per day, self-rated English proficiency, year in university, length living in the UK, and the anxiety check were explored. Correlations between participant age, age of acquisition, and the anxiety check were conducted using Pearson's correlation. However, as percentage of English use per day, self-rated English proficiency, year in university, and length living in the UK were ordinal variables, Spearman's Rho was a more appropriate analysis for conducting correlations for these variables.

### *Demographic Variables and the Foreign Language Anxiety in Formal Contexts Scale*

A Pearson's correlation showed the FLAFS to be significantly correlated with participant age of acquisition ($r = .324$, $p = .006$), and the anxiety check measure ($r = .534$, $p < .001$). However, the FLAFS had no significant relationship with participant age ($r = .084$, $p = .485$). A Spearman's Rho correlation conducted for the ordinal variables showed a significant correlation with percentage of English use per day (rho = –.301, $p = .010$), self-rated English proficiency (rho = –.628, $p < .001$). There was no significant correlation between the FLAFS and year in university (rho = .143, $p = .231$), nor length the participant has lived in the UK (rho = –.038, $p = .754$). Those participants who scored higher in anxiety were more likely to have learned English later in life, use English less often per day, and rate themselves as having lower English proficiency. Those with high anxiety were also more likely to have higher anxiety about completing the speaking task as indicated on the anxiety check measure.

### *Demographic Variables and the Foreign Language Classroom Anxiety Scale*

A Pearson's correlation indicated a significant correlation between the FLCAS and participant age of acquisition ($r = .373$, $p = .001$), and the anxiety check measure ($r = .627$, $p < .001$). The relationship between the FLCAS and participant age was non-significant ($r = .130$, $p = .275$). A Spearman's Rho correlation showed a significant relationship between percentage of English used per day (rho = –.394, $p < .001$), and self-rated English proficiency (rho = –.719, $p < .001$). The FLCAS was not significantly correlated with the participants' year in university (rho = .223, $p = .060$) nor the length they have lived in the

UK (rho = –.005, $p$ = .967). Similar to the results of the FLCAS, those with high anxiety were more likely to have learned English later in life, use English less often per day, rate themselves as having lower English proficiency, and have higher anxiety before completing the speaking task.

### *Demographic Variables and the Utterance Fluency*

A Pearson's correlation showed no significant correlations between the participants' percentage of utterance disfluencies in their speech samples, and participant age of acquisition ($r$ = .218, $p$ = .066), age ($r$ = .145, $p$ = .244), and the anxiety check measure ($r$ = .173, $p$ = .170). A Spearman's Rho correlation showed no significant relationships with length living in the UK (rho = .045, $p$ = .704), but did, however, show significant correlations with year in university (rho = .288, $p$ = .014), percentage of English used per day (rho = –.342, $p$ = .003), and self-rated English proficiency (rho = –.440, $p$ < .001). This indicates that the participants current age, age which they began learning English, the length they lived in the UK, nor their indication of how anxious they felt before the task has no relationship to the percentage of disfluencies during the speaking task. On the other hand, those participants who spoke a higher percentage of English per day, and those who self-rated their English proficiency higher had fewer disfluencies in speech. In addition, the respondents who were further along in their university career also had a higher percentage of disfluencies during the speaking task.

### *Foreign Language Anxiety and Utterance Disfluency by Participant Nationality*

Considering the wide variety of nationalities and language backgrounds reported by the participants in this study, an exploration of how these differences in background related to scores on the FLAFS, FLCAS, and utterance fluency outcomes were examined using a series of one-way ANOVAs. The language groups were separated into four groups by continent: East Asia ($N$ = 25), Europe ($N$ = 16), and Central and South America ($N$ = 11) West Asia and Africa ($N$ = 20) (See Table 3 for a list of the countries in each group).

**Table 3**

*Participant Nationality by Continent*

| East Asia | Europe | Central & South America | West Asia & Africa |
|---|---|---|---|
| Chinese | Italy | Chile | India |
| Hong Kong | Germany | Peru | Pakistan |
| Malaysia | Romania | Colombia | Iran |
| Indonesia | Denmark | Brazil | Bangladesh |
| Japan | Russia | Uruguay | Kenya |
| Thailand | Slovakia | Costa Rica | Ethiopia |
| Taiwan | Slovenia | Mexico | Bengali |
| | Norway | | Farsi |
| | Lithuania | | Saudi Arabia |
| | Poland | | Egypt |
| | Spain | | |

Due to the group sizes being unequal, a Welch's ANOVA was conducted to examine differences in means between the participants from East Asia, Europe, Central and South America, and West Asia and Africa. There were no significant effects found between the participants' regional nationality and their scores on the FLAFS ($F(2,69) = 1.957$, $p = .149$), FLCAS ($F(2,69) = 1.163$, $p = .319$), nor was there a significant effect found in terms of regional nationality and percentage of utterance disfluencies during the speaking task ($F(2,69) = .108$, $p = .897$). This indicates that regardless of where the participants came from, there was no difference in their foreign language anxiety scores nor in their performance on the speaking task.

## Research Question 1: How does foreign language anxiety affect international students' utterance fluency during a speaking task?

To answer the first research question examining the impact of FLA on utterance fluency, correlational analyses were conducted between the anxiety questionnaires and the utterance fluency variables examined within the international students' speech samples to determine whether general feelings of anxiety affected utterance fluency during speech. Due to the relatively short speech sample per participant, utterance fluency measures were combined into the percentage of total utterance disfluencies per speech sample, which included word repetitions, phrase repetitions, word revisions, phrase revisions, word fragments, and filled pauses.

The correlational analysis shows that participant scores on the FLCAS were correlated with total percentage of utterance disfluencies ($r = .260$, $p = .029$) in the speech sample during the speaking task, indicating that those with higher classroom anxiety scores were more likely to have a higher number of utter-

ance disfluencies while speaking. Similarly, participant scores on the FLAFS were also correlated with total percentage of utterance disfluencies ($r$ = .255, $p$ = .033) indicating that those with high FLA in general contexts were also more likely to have a higher percentage of total utterance disfluencies. Descriptives of both the FLCAS and FLAFS can be found in Table 4. Furthermore, those who experienced high FLA in the classroom as measured by the FLCAS also experienced high FLA outside of the classroom, as measured by the FLAFS ($r$ = .921, $p$ < .001), indicating that FLA is indeed not limited to the language classroom but pervades throughout the participants' daily lives. In addition, both the FLCAS ($r$ = .525, $p$ < .001), and the FLAFS ($r$ = .620, $p$ < .001), were correlated with the anxiety check measure, indicating that those with typically high FLA are more likely to find participating in an oral speaking task more anxiety provoking.

**Table 4**

*Descriptive Statistics of the Scores on the FLCAS and FLAFS*

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| FLCAS | 39.6 | 13.1 | 39 | 145 |
| FLAFS | 91.8 | 26.8 | 13 | 68 |

**Research Question 2: How does the presence of native English speakers impact international student anxiety and utterance fluency during a speaking task?**

To address the second research question, a series of correlational analyses and t-tests were conducted between the variables of interest. First, an examination of whether participant scores on the anxiety check measure was related to the total percentage of utterance disfluencies in the speech samples was conducted. A correlation analysis showed a non-significant relationship between the anxiety check measure and the total percentage of utterance fluencies ($r$ = .202, $p$ = .093), indicating that the anxiety the participants felt going into the task did not relate to the utterance disfluencies they produced during the speaking task. Next, an independent sample t-test was conducted between the anxiety check and the experimental conditions to investigate whether the participants speaking in front of an audience were more anxious about the speaking task than those without an audience. The t-test showed no difference in the anxiety check ($t$ (69) = .059, $p$ = .477), between the control and the experimental conditions (see Table 5 for descriptives of both conditions) indicating that there was no difference in anxiety during the speaking task regardless of whether there was a native-speaker audience.

**Table 5**

*Descriptive Statistics of the Scores on the Anxiety Check Measure Separated by Experimental Condition*

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| Control | 42.0 | 27.5 | 0 | 82 |
| Experimental | 41.6 | 29.8 | 0 | 91 |

A second independent samples t-test was conducted between the total percentage of utterance disfluencies and the experimental condition to investigate whether those speaking in front of a native-speaker audience produced more disfluencies in their speech sample. The analysis showed no difference in the percentage of utterance disfluencies ($t$ (68) = .212, $p$ = .137), between the control and the experimental conditions, indicating that the presence of an audience had no impact on the percentage of utterance disfluencies in participants' speech samples. Descriptives of the total percentage of utterance disfluency scores among both conditions can be found in Table 6.

**Table 6**

*Descriptive Statistics of the Total Percentage of Utterance Disfluencies by Experimental Condition*

| Variable | Mean | SD | Min | Max |
|---|---|---|---|---|
| Control | 7.8 | 4.3 | 1.95 | 18.5 |
| Experimental | 9.6 | 8.2 | 2.0 | 49 |

## Discussion

The aim of the current study was to examine the effects of FLA on English utterance fluency among international students in the UK who have learned English as an additional language. Research has shown that FLA is strongly associated with second language proficiency (Ardasheva et al., 2018; Bekleyen, 2009; Woodrow, 2011; Zhang, 2013), especially in the oral domain (Bielak, 2022; Castillejo, 2019). The first research question attempted to examine how FLA affects international students' utterance fluency during a speaking task in order to identify the real-life effects of FLA on the way learners speak. It was expected that the participants who scored higher on the FLA measures would

have more utterance disfluencies in their speech sample during the speaking task. This hypothesis was supported. Both scores on the FLCAS and FLAFS were correlated with a higher percentage of utterance disfluencies during speech, suggesting that FLA impedes the fluency of the speaker. This is in line with other research that has found a relationship between anxiety and oral fluency (Bielak, 2022; Castillejo, 2019; Trebtis, 2014). As anxiety impedes retrieval, this can lead to "freezing up" moments, where the speaker requires more time to select the appropriate vocabulary due to the ease of access to target language knowledge being reduced (Zheng, 2008). These moments in turn lead anxious individuals to cope by using time-gaining mechanisms which may impede the fluency of their speech (Castillejo, 2019; Gotz, 2013). The results of the study suggest that international students undertaking a degree in the UK who experience high FLA may also produce more utterance disfluencies in their speech when partaking in classroom discussions and activities, as well as during life outside of the classroom. This may make it difficult for students to clearly communicate their ideas, be understood by others, fully engage in discussions, and navigate activities required in daily life (Brown, 2008; Isaacs & Trofimovich, 2012; Sawir et al., 2012).

The second research question attempted to further examine the relationship between FLA and utterance fluency and investigate how this relationship is affected by the presence of native speaker listeners. As research has shown speaking with strangers and in front of others is particularly anxiety-provoking (Bekleyen, 2009; Dewaele, 2007), it was expected that those in the experimental condition with native speakers in the audience would experience higher anxiety and therefore more disfluencies in their speech. Surprisingly, this hypothesis was not supported. Although those with higher anxiety scores on the FLFAS and FLCAS were more likely to feel anxious about the speaking task, the anxiety check for the speaking task had no relationship to the experimental condition nor the total percentage of utterance disfluencies. This was unexpected considering the results of previous literature reporting feelings of pressure and high anxiety among learners when they are required to speak in front of native speakers (Lin, 2022; Sato, 2007). It is possible that although research has found learners to be anxious in online situations (Russel, 2020) the lack of camera use in the current study may not have made the native speaker listeners salient enough to affect the participants' anxiety and lead to differences in fluency.

Interestingly, despite random allocation to the experimental conditions, the descriptive statistics showed that there were participants in the control condition who scored particularly highly on the anxiety check measure, indicating that some participants may have felt extremely anxious about the mere prospect of the requirement to speak. However, the anxiety check measure was not related to participants' total percentage of utterance disfluencies during the speaking task, meaning that the way the respondents felt going into the task

did not necessarily affect the outcome. It is possible that for some participants, high anxiety was only felt at the beginning of the task, but waned as time went by, similar to the results of the study by MacIntyre and Gardner (1994a). In their study examining induced anxiety through the use of video cameras, the highest level of anxiety was reported by participants immediately after the introduction of the camera, and it was at this point when they had reduced performance. Furthermore, in a similar study by Gregersen, MacIntyre, and Meza (2014), participants were videorecorded while giving a speech and were asked to review the video and rate their moment-to-moment anxiety throughout. The results indicated that although all participants felt high anxiety at some point, those who were typically low in anxiety had a reduction in anxiety over the course of their speech, while conversely, those typically high in anxiety experienced an increase in anxiety over time. Likewise, in the current study, although all participants may have felt anxious upon having to complete the speaking task on Zoom, those typically high in FLA may have experienced increases over time, while those with typically lower FLA may have had reduced anxiety as they continued speaking. Therefore, for international students who struggle with typically high FLA, feelings of anxiety during their interactions with others may not decrease as the interaction continues and they may suffer from increased anxiety as they continue speaking, which may further impede utterance fluency. On the other hand, international students with typically low FLA may experience high anxiety at the beginning of a conversation but are able to cope with the anxiety as the conversation goes on.

## Limitations and Future Directions

There are a several limitations to the current study. The first limitation is regarding the anxiety check measure used as a manipulation check between the experiment and control conditions. Hauser, Ellsworth, and Gonzalez (2018) suggest that manipulation checks may change the experience of the participant and have an effect on the independent variable. Manipulation checks may cause participants to be wary, distract participants, or lead them to guess what the researcher is expecting. The timing of the manipulation check in this study may not have directly related to the anxiety the participants actually felt in the anxiety task. It is possible that despite the respondents reporting anxiety upon entering the speaking task, the effect of anxiety may have been temporary and did not last throughout the task. It may have been beneficial to have another anxiety check after the speaking task to see whether the participants' feelings of anxiety held through to the end of the experiment, or a series of anxiety

checks throughout. As manipulation checks may interfere with participant focus on the main task and interfere with the actual experimental manipulation (Hauser, Ellsworth, & Gonzales, 2018), rather than directly asking participants how they are feeling during the task, future research may also benefit from implementing alternative methods of assessing anxiety throughout the task, such as the use of reflective interviews or physiological measures. Physiological measures would allow for the possibility to track participants' biological indices of anxiety throughout the task, and reflective interviews may provide information regarding whether they subjectively felt anxiety along with their reasons for experiencing anxiety from the initiation of speech until the end.

Another limitation is related to the experimental condition itself. Although research has shown that interacting with native English speakers via technology may cause heightened anxiety among language learners (Kessler, 2010; Lee, 2004), the constraints of the experiment may have prevented this. In the current study, video cameras were not used for the confederate accounts nor by the participants. This may have given them the opportunity to cope with their anxiety, as they were not actually seeing and interacting with the native English speakers. Despite the lack of camera use being typical for students who attended online classes during the pandemic (Castelli & Sarvary, 2020) the anonymity provided by the lack of camera use may have been beneficial to those who struggle with anxiety. Studies have reported that some students with high language anxiety choose to enroll in online courses due to the ability to secure anonymity (Pichette, 2009). Students with high FLA may be better able to cope during online engagement, as they have more time to formulate responses in the L2 compared to face-to-face contexts (Garcia-Castro & O'Reilly, 2022). Furthermore, studies comparing the severity of social anxiety in online and real-life interactions have found decreased social anxiety in the online environment (Yen, Chen, & Wang, 2012). In the current study, as the participants were asked to answer a question with no verbal interruptions, responses, or follow-up questions from the researcher or confederates, the participants in this study may not have felt pressure to generate their response, especially considering they had the option of using up to two minutes to prepare their answer.

In addition to not being identified through video cameras, the respondents also did not have any identifiable information connecting them to the speaking task and had their profile names changed to a participant ID number, which may have further given them the security of anonymity (Shepherd & Edelmann, 2005; Weidman et al., 2012). This sense of anonymity may have taken away from the participants fear of being negatively evaluated by others, a major component of foreign language anxiety (Horwitz et al., 1986; Lin, 2022).

## Conclusion

Despite these limitations, the results of this study provide noteworthy insights into the effects of FLA on the utterance fluency of international students who have participated in online education in the UK during the pandemic and provide a valuable contribution to the literature. This study fills a gap in literature by being one of the few studies to use experimental methods to examine how different online settings affect international students' ability to smoothly communicate in English, and to measure how the effects of FLA manifest during communication. Students who struggle with FLA may have difficulty clearly expressing their ideas, and this study provided evidence that FLA may not only affect students in general terms such as course grades (Bekleyen, 2009; Cheng, Horwitz, & Shallert, 1999; MacIntyre & Gardner, 1994b) but may also affect individuals attempting to communicate their ideas verbally, even in an online setting. The study also suggests that the presence of native speakers may not always exacerbate anxiety as found in previous literature (Lin, 2022; Russel, 2020). It is possible that if the salience of the individuals listening are reduced or unable to be seen by the speaker, they may not negatively affect the speakers' ability to communicate. This indicates that online settings where students are not required to use cameras may be beneficial for international students using English as an additional language and may lead to a higher likelihood of them feeling less pressure and better able to communicate smoothly.

The exploration of the background variables of the participants in this study suggests that FLA may negatively affect international students regardless of their nationality background. These negative effects may become particularly apparent when a student has been called upon to answer a question, has been asked to partake in a group discussion, or present in front of others. It may require patience on the listener side as the speaker requires extra time to retrieve necessary vocabulary and grammar for their message. This highlights the need for consideration on the part of educators that any delays an anxious international student makes when required to provide an answer or add their opinions to a discussion may not necessarily be due to a lack of knowledge, but rather a delay in being able to retrieve the language needed to respond. This may be especially true during oral exams, or assessments where a student is required to orally communicate in a second language, especially in the presence of others. Considering that in the current study, students who used a higher percentage of English per day had lower FLA and fewer disfluencies during the speaking task, it may be beneficial for educators to encourage students to use English more often and provide opportunities for students to practice their language abilities in low pressure situations where they do not feel judged.

The results of the current study show that FLA can indeed affect international student communication in an online setting, but it is possible that these negative effects may be more pronounced in a physical classroom. Providing students with sufficient time to speak and creating an environment with a lack of pressure to respond rapidly may help international students to cope with their anxious thoughts and fully participate in classroom activities.

## References

Ardasheva, Y., Carbonneau, K. J., Roo, A. K., & Wang, Z. (2018). Relationships among prior learning, anxiety, self-efficacy, and science vocabulary learning of middle school students with varied English language proficiency. *Learning and Individual Differences, 61*, 21–30. https://doi.org/10.1016/j.lindif.2017.11.008

Basow, S. A., & Gaugler, T. (2017). Predicting adjustment of U.S. college students studying abroad: Beyond the multicultural personality. *International Journal of Intercultural Relations, 56*, 39–51. https://doi.org/10.1016/j.ijintrel.2016.12.001

Beason-Abmayr, B., Caprette, D. R., & Gopalan, C. (2021). Flipped teaching eased the transition from face-to-face teaching to online instruction during the COVID-19 pandemic. *Advances in Physiological Education, 45*, 384–389. https://doi.org/10.1152/advan.00248.2020

Bekleyen, N. (2009). Helping teachers become better English students: Causes, effects, and coping strategies for foreign language listening anxiety. *System, 37*(4), 664–675. https://doi.org/10.1016/j.system.2009.09.010

Bielak, J. (2022). To what extent are foreign language anxiety and foreign language enjoyment related to L2 fluency? An investigation of task-specific emotions and breakdown and speed fluency in an oral task. *Language Teaching Research, 29*(3), 1–31. https://doi.org/10.1177/13621688221079319

Birney, M. E., Rabinovich, A., Morton, T. A., Heath, H., & Ashcroft, A. (2020). When speaking English is not enough: The consequences of language-based stigma for non-native speakers. *Journal of Language and Social Psychology, 39*(1), 67–86. https://doi.org/10.1177/0261927X1988390

Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T., & de Jong, N. H. (2012). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing, 30*(2), 159–175. https://doi.org/10.1177/0265532212455394

Brown, L. (2008). Language and anxiety: An ethnographic study of international postgraduate students. *Evaluation & Research in Education, 21*(2), 75–95. https://doi.org/10.1080/09500790802152167

Burns, R. (2020). A COVID-19 panacea in digital technologies? Challenges for democracy and higher education. *Dialogues in Human Geography, 10*(2), 1–4. https://doi.org/10.1177/2043820620930832

Cadena-Aguilar, R. F., Ortega-Cuellar, J. H., & Cadena-Aguilar A. (2019). An approach to foster oral fluency of English as a foreign language in adolescents. *Profile: Issues in Teachers' Professional Development, 21*(2), 29–44. https://doi.org/10.15446/profile.v21n2.71364

Castelli F. R., & Sarvary, M. A. (2020). Why do students not turn on their video cameras during online classes and an equitable and inclusive plan to encourage them to do so. *Academic Proactive in Ecology and Evolution, 11*, 3565–3576. https://doi.org/10.1002/ece3.7123

Castillejo, S. P. (2019). The role of foreign language anxiety on L2 utterance fluency during a final exam. *Language Testing, 36*(3), 327–345. https://doi.org/10.1177/0265532218777783

Castillejo, S. P. (2021). Prior processing, foreign language classroom anxiety, and L2 fluency. *International Review of Applied Linguistics in Language Teaching, 61*(2), 519–544. https://doi.org/10.1515/iral-2021-0091

Cheng, Y., Horwitz, R. K., & Shallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417–466. https://doi.org/10.1111/0023-8333.00095

Cucciarini, C., Strik, H., & Boves, L. (2000). Quantitative assessment of second language learners' fluency: Comparisons between read and spontaneous speech. *Journal of the Acoustical Society of America, 111*(6), 2862–2873. https://doi.org/10.1121/1.1471894

Dewaele, J. (2007). The effect of multilingualism, sociobiographical, and situational factors on communicative anxiety and foreign language anxiety of mature learners. *International Journal of Bilingualism, 11*(4), 391–409. https://doi.org/10.1177/13670069070110040301

Dewey, D. P., Belnap, R. K., & Steffen, P. (2018). Anxiety: Stress, foreign language classroom anxiety, and enjoyment during study abroad in Amman, Jordan. *Annual Review of Applied Linguistics, 38*, 140–161. https://doi.org/10.1017/S0267190518000107

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353. https://doi.org/10.1037/1528-3542.7.2.336

Gagne, N., French, L. M., & Hummell, K. M. (2022). Investigating the contribution of L1 fluency, L2 initial fluency, working memory, and phonological memory to L2 fluency development. *Language Teaching Research, 29*(2), 1–24. https://doi.org/10.1177/13621688221076418

Ganschow, L., & Sparks, R. (1996). Anxiety about foreign language learning among high school women. *The Modern Language Journal*, *80*(2), 199–212. https://doi.org/10.2307/328636

Garcia-Castro, V., & O'Reilly, J. (2022). Foreign language anxiety and online engagement during the COVID-19 pandemic: A comparison between EMI and FMI university students. *English Teaching & Learning, 46*, 273–291. https://doi.org/10.1007/s42321-022-00120-x

Gargalianou, V., Muehlfeld, K., Urbig, D., & van Witteloostuijn, A. (2016). Foreign language anxiety in professional contexts. *Schmalenbach Business Review*, *17*(2), 195–223. https://doi.org/10.1007/s41464-016-0007-6

Georgiadou, E., & Roehr-Brakin, K. (2017). Investigating executive working memory and phonological short-term memory in relation to fluency and self-repair behavior in L2 speech. *Journal of Psycholinguistic Research, 46*, 877–895. https://doi.org/10.1007/s10936-016-9463-x

Gots, S. (2013). *Fluency in native and non-native speech.* John Benjamins.

Gregersen, T., Macintyre, P. D., & Meza, M. D. (2014). The motion of emotion: Idiodynamic case studies of learners' foreign language anxiety. *The Modern Language Journal*, *98*(2), 574–588. https://doi.org/10.1111/j.1540-4781.2014.12084.x

Hauser, D. J., Ellsworth P. C., & Gonzales, R. (2018). Are manipulation checks necessary. *Frontiers in Psychology, 9*(988). https://doi.org/10.3389/fpsyg.2018.00998

Horwitz, E. K. Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125–132. https://doi.org/10.2307/327317

Isaacs, T., & Trofimovich, P. (2012). Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition, 34*, 475–505. https://doi.org/10.1017/S0272263112000150

Kessler, G. (2010). Fluency and anxiety in self-access speaking tasks: The influence of environment. *Computer Assisted Language Learning, 23*(4), 361–375. https://doi.org/10.1080/09588221.2010.512551

Kim, H. I., & Cha, K. A. (2017). Effects of experience abroad and language proficiency on self-efficacy beliefs in language learning. *Psychological Reports*, *120*(4), 670–694. https://doi.org/10.1177/0033294117697088

Kormos, J. (2006). *Speech production and second language acquisition.* Lawrence Erlbaum Associates Publishers.

Linck, J. A., & Weiss, D. J. (2015). Can working memory and inhibitory control predict second language learning in the classroom? *SAGE Open*, *5*(4). https://doi.org/10.1177/2158244015607352

Lin, Q. (2022). Anxiety and self-efficacy in Chinese international students' L3 French learning with L2 English and L3 French. *Frontiers in Psychology, 13*, 1–11. https://doi.org/10.3389/fpsyg.2022.998536

Liu, M. (2018). A study of Chinese University learners' anxiety in German language classrooms in at-home and study-abroad contexts. *Cogent Education*, *5*(1). https://doi.org/10.1080/2331186x.2018.1475917

Martirosyan, N. M., Hwang, E., & Wanjohi, W. (2015). Impact of English proficiency on academic performance of international students. *Journal of International Students, 5*(1), 60–71. https://doi.org/10.32674/jis.v5i1.443

MacIntyre, P. D., & Gardner, R. C. (1991). Language anxiety: Its relationship to other anxieties and to processing in native and seconds languages. *Language Learning, 41*(4), 513–534. https://doi.org/10.1111/j.1467-1770.1991.tb00691.x

MacIntyre, P. D., & Gardner, R. C. (1994a). The effects of induced anxiety on three stages of cognitive processing in computerized vocabulary learning. *Studies in Second Language Acquisition, 16*, 1–17. https://doi.org/10.1017/S0272263100012560

MacIntyre, P. D., & Gardner, R. C. (1994b). The subtle effects of language anxiety on cognitive processing in the second language. *Language Learning, 44*(2), 283–305. http://doi.org/10.1111/j.1467-1770.1994.tb01103.x

MacWinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk.* 3rdr ed. Lawrence Erlbaum Associates.

Mameseishvilli, K. (2012). International student persistence in U.S. postsecondary institutions. *Higher Education, 64*, 1–17. https://doi.org/10.1007/s10734-011-9477-0

Moorhouse, B. L. (2023). Teachers' digital technology use after a period of online teaching. *ELT Journal*, *77*(4), 1–13. https://doi.org/10.1093/elt/ccac050

Neumann, H. Padden, N., & McDonough, K. (2019). Beyond English language prodiciency scores: Understnding the academic performance of international undergraduate students during the first year of study. *Higher Education Research & Development*, *38*(2), 324–338. https://doi.org/10.1080/07294360.2018.1522621

O'Loughlin, K. (1995). Lexical density in candidate output on direct and semi-direct versions of an oral proficiency test. *Language Testing, 12*(2), 217–237. https://doi.org/10.1177/026553229501200202

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.x

Poyrazli, S., & Lopez, M. D. (2007). An exploratory study of perceived discrimination and homesickness: a comparison of international students and American students. *The Journal of Psychology*, *141*(3), 263–280. https://doi.org/10.3200/JRLP.141.3.263-280

Russell, V. (2020). Language anxiety and the online learner. *Foreign Language Annals, 53*(3), 338–352. https://doi.org/10.1111/flan.12461

Saito, Y., Horwitz, E. J., & Garza, T. J. (1999). Foreign language reding anxiety. *Modern Language Journal, 83*, 202–218. http://dx.doi.org/10.1111/0026-7902.00016

Sato, M. (2007). Social relationships in conversational interaction: Comparison of learner-learner and learner-NS dyads. *JALT Journal, 29*(2), 183–208. https://doi.org/10.37546/JALTJJ29.2-2

Sawir, E., Marginson, S., Forbes-Mewett, H., Nyland, C., & Ramia, G. (2012). International student security and English language proficiency. *Journal of Studies in International Education*, *16*(5), 434–454. https://doi.org/10.1177/1028315311435418

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Segalowitz, N. (2016). Second language fluency and its underlying cognitive and social determinants. *International Review of Applied Linguistics in Language Teaching, 54*(2), 79–95. https://doi.org/10.1515/iral-2016-9991

Shepherd, R., & Edelmann, R. (2005). Reasons for internet use and social anxiety. *Personality and Individual Differences, 39*(5), 949–958. https://doi.org/10.1016/j.paid.2005.04.001

Skehan, P. (2003). Task-based instruction. *Language Teaching*, *36*, 1–14. https://doi.org/10.1017/S026144480200188X

Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning, 6*(1), 37–72. https://doi.org/10.1111/j.1467-9922.2011.00642.x

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). Benjamins.

Thompson, A. S., & Lee, J. (2014). The impact of experience abroad and language proficiency on language learning anxiety. *TESOL Quarterly*, *48*(2), 252–274. https://doi.org/10.1002/tesq.125

Tran, L. T., & Pham, L. (2016). International students' transnational mobility: Intercultural connectedness with domestic and international peers, institutions and the wider community. *Compare: A Journal of Comparative and International Education, 46*(4), 560–581. https://doi.org/10.1080/03057925.2015.1057479

Trebits, A. (2016). Sources of individual differences in L2 narrative production: The contribution of input, processing, and output anxiety. *Applied Linguistics, 37*(2), 155–174. https://doi.org/10.1093/applin/amu006

Verhagen, J., & Leseman, P. (2016). How do verbal short-term memory and working memory relate to the acquisition of vocabulary and grammar? A comparison between first and second language learners. *Journal of Experimental Child Psychology*, *141*, 65–82. https://doi.org/10.1016/j.jecp.2015.06.015

Weidman, A. C., Fernandez, K. C., Levinson, C. A., Augustine, A. A., Larsen, R. J., & Rodebaugh, T. L. (2012). Compensatory internet use among individuals higher in social anxiety and its implications for well-being. *Personality and Individual Differences, 53*(3), 191–195. https://doi.org/10.1016/j.paid.2012.03.003

Woodrow, L. (2011). College English writing affect: Self-efficacy and anxiety. *System, 39*(4), 510–522. https://doi.org/10.1016/j.system.2011.10.017

Yen, H., Chen, C., & Wang, P. (2012). Social anxiety in online and real-life interaction and their associated factors. *Cyberpsychology, Behavior, and Social Networking 15*(1), 7–12. https://doi.org/10.1089/cyber.2011.0015

Zhang, X. (2013). Foreign language listening anxiety and listening performance: Conceptualizations and causal relationships. *System, 41*(1), 164–177. https://doi.org/10.1016/j.system.2013.01.004

Zuniga, M., & Simard, D. (2022). Exploring the intricate relationship between foreign language anxiety, attention and self-repairs during L2 speech production. *System, 105*, 1–10. https://doi.org/10.1016/j.system.2022.102732

Appendix A

# Demographics Questionnaire

1. Gender
2. □ Male □ Female □ Prefer not to say □ Other
3. Age
4. Nationality
5. What is your current year in university?
6. □ First year □ Second year □ Third year □ Fourth year □ Master's PhD
7. How old were you when you first began learning English?
8. What language(s) did you speak at home growing up?
9. How long total have you lived in an English-speaking country?
10. What percentage of time do you use English on a daily basis?
11. □ 0% □ 10% □ 20% □ 30% □ 40% □ 50% □ 60% □ 70% □ 80% □ 90% □ 100%
12. Please rate your English proficiency on these four aspects.

| | Very poor | Poor | Ok | Good | Very good | Excellent | Native-like |
|---|---|---|---|---|---|---|---|
| Reading | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Writing | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Speaking | ○ | ○ | ○ | ○ | ○ | ○ | ○ |
| Listening | ○ | ○ | ○ | ○ | ○ | ○ | ○ |

Appendix B

## Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz, Horwitz & Cope, 1986)

Answer the following questions related to your experience and feelings of anxiety in foreign language communication on a 7-point Likert scale ranging from "strongly agree" to "strongly disagree."

1. I never feel quite sure of myself when I am speaking English in class.
2. I don't worry about making mistakes while speaking English in class.
3. I tremble when I know that I'm going to be called on and have to speak in English.
4. It frightens me when I don't understand what the teacher is saying in English.
5. It wouldn't bother me at all to take more classes in English.
6. During classes in which the subject is taught in English, I find myself thinking about things that have nothing to do with the unit.
7. I keep thinking that the other students are better at English than I am.
8. I am usually at ease when I have to use English for tests in class.
9. I start to panic when I have to speak English in class without preparation.
10. I worry about the consequences of failing my classes due to my English.
11. I don't understand why some people get so upset over having to use English in class.
12. In classes where I must use English, I can get so nervous I forget things I know.
13. It embarrasses me to volunteer answers using English in class.
14. I would not be nervous speaking English with native speakers.
15. I get upset when I don't understand what the teacher is correcting.
16. Even if I am well prepared for a class taught in English, I feel anxious about it.
17. I often feel like not going to my classes where the subject is taught in English.
18. I feel confident when I speak English in class.
19. I am afraid that lecturers are ready to correct every English mistake I make.
20. I can feel my heart pounding when I'm going to be called on in class and be forced to use English.
21. The more I study for a test in which I will have to use English, the more confused I get.
22. I don't feel pressure to prepare very well for classes taught in English.

*Appendix B continued*

23. I always feel that the other students speak English better than I do.
24. I feel very self-conscious about speaking English in front of other students.
25. Classes that are taught in English move so quickly I worry about getting left behind.
26. I feel more tense and nervous in classes taught in English than in classes taught in my native language.
27. I get nervous and confused when I am speaking English in class.
28. When I'm on my way to classes taught in English, I feel very sure and relaxed.
29. I get nervous when I don't understand every word the language teacher says in English.
30. I feel overwhelmed by the number of rules you have to learn to speak English.
31. I am afraid that the other students will laugh at me when I speak English.
32. I would probably feel comfortable around native speakers of English.
33. I get nervous when the language teacher asks questions that I haven't prepared for in advance.

Appendix C

# Adapted Measure of Foreign Language Anxiety in Formal Contexts Scale (FLA-FS) – English Version

To answer the following questions, imagine you are participating in an important meeting/discussion that takes place in English. To communicate with the rest of the participants, you have to use English. Now, please evaluate the following items on a scale from 1 = "strongly disagree" to 7 = "strongly agree."

1. I feel overwhelmed by the number of rules you have to learn to speak English.
2. I can feel my heart pounding when I'm going to be called on in a meeting in English.
3. I am afraid that many people will laugh at me when I speak English.
4. I get nervous and confused when I am speaking English.
5. I get nervous when I don't understand every word people who have power over me say to me in English.
6. I get nervous when persons who have power on me ask questions in English which I haven't prepared in advance.
7. When interacting in English, I can get so nervous I forget things I know.
8. I am afraid that people above me are ready to correct every mistake I make when speaking English.
9. I don't worry about making mistakes when I interact in English.
10. I keep thinking that many other people are better in English than I am.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/21
https://doi.org/10.31261/TAPSLA.15909

CC BY SA

**Agata Wolanin**
https://orcid.org/0000-0002-8513-2154
University of the National Education Commission
Krakow, Poland

# "I Feel Like I'm a Different Person": Exploring Undergraduate Students' Imagined L2 Selves

Abstract

Identity has a prominent place in language education. It can be manifested by the imagined L2 self, understood as a realization of how L2 learners imagine, perceive and talk about their experience of being an L2 user. It might be argued that undergraduate students of foreign languages should exhibit more complex language identities, as their imagined L2 selves are shaped by their rich language repertoires. The major aim of the study was to examine undergraduate students' imagined L1 and L2 selves, by investigating two dimensions of their complex language identity, namely: how they view these languages and how they feel using these languages. A total of 200 students (88 English major, 67 German major and 45 Swedish major students) completed a questionnaire designed to explore their specific perception of the languages they speak and the way they feel using them. The results offer an interesting insight into the complex language identity profiles of the participants and their imagined selves. Overall, the students seem to show high language awareness; they are sensitive to the differences in language systems, including the sounds and the pragmatics of a given language. The students also exhibit very positive attitudes towards their imagined L2 selves. It is, thus, suggested that language teachers provide a platform for the students to express their complex identities by incorporating language journals, or language biographies, in order to allow the students to explore their identities in more depth. It can also be argued that the teachers could capitalize on the affective dimension of the students' L2 identity, by fostering the students' intrinsic motivation and supporting their investment in learning.

*Keywords*: language identity, imagined identity, imagined L2 self, language education

Reflecting on the words of Bonny Norton (1997), who observed that "every time language learners speak, they are not only exchanging information with their interlocutors; they are also constantly organizing and reorganizing a sense

of who they are and how they relate to the social world" (p. 410), it can be argued that language learners should have a more complex understanding of who they are and how they identify themselves as language users. The exploration of identity has a long tradition in language education (see, e.g., Norton, 2006; Douglas Fir Group, 2016), with self-concept playing an important role in a variety of processes that are essential in language education research, such as motivation, investment in learning, agency, autonomy, self-esteem and self-efficacy, to name only a few (e.g., Huang & Benson, 2013; Darvin & Norton, 2015; De Costa & Norton, 2016). A substantial proportion of language identity research focuses on language learners and their teachers, yet relatively few studies have been conducted that investigate the identities of university students majoring in foreign languages. Assuming that language users indeed think and feel differently when expressing themselves in different languages, the language identities of students majoring in foreign languages is an intriguing area of research.

Thus, university students who chose foreign languages as their major constitute an interesting group of L2 learners. Studying a language major could mean much more than the mere study of a language system itself; philology programmes in Poland, which typically involve 3–5 years of study, offer a full immersion in the language, literature and culture of the subject. It might be assumed that the university students would exhibit high levels of language and cultural awareness, which might consequently result in complex and dynamic language identities. This aspect of identity research has yet to be fully addressed, and the study presented in this paper aims at bridging this gap.

The main objective of the text is to, firstly, briefly discuss the existing research on language identity in general terms, and the concept of an imagined L2 self in particular, and secondly, to explore the language identities of students majoring in three different languages: English, German and Swedish. The way their imagined L1/L2 self is perceived and manifested is the focal point of the study.

## The Concept of Identity in Language Education – Literature Review

Identity may be defined as a reflection of "how people understand their relationship to the world, how that relationship is constructed across time and space, and how people understand their possibilities for the future" (Norton, 1997, p. 410). In other words, identity represents the struggle of making sense of who we are with relation to each other (Darvin & Norton, 2014). In the context of language education, the concept of identity is deeply embedded

within the sociocultural perspective and, more specifically the Social Identity Theory (Tajfel & Turner, 1986). Within this framework social identity is formed by a group of individuals who share many common traits and qualities in the cognitive, behavioural, and affective domains. Thus, it is assumed that identity is a sociocultural construct that allows for creating and defining "the individual's place in society" (Tajfel & Turner, 1986, p. 283). The subjectivity of identity within this relation is bidirectional – a person can be both "subject *of* […] and subject *to* a set of relationships" (Norton, 2013, p. 4). Thus, a constant tension is observed between who one sees oneself as and how one is seen by others, as identity is both relational and comparative (Tajfel & Turner, 1986). Additionally, within the framework of the post-structuralist and transmodern paradigm, identity is seen as dynamic, diverse, context-dependent, complex and sometimes also contradictory (Darvin & Norton, 2014, p. 57).

Identity often stands in close proximity to the concept of *self*, which also emphasizes the relational dimension with the other (for an overview of different types of self, see Komorowska, 2019). It is, however, often argued that there is a substantial difference between the two, as identity embraces the continuity of who we are across time and space, and *self* represents a temporary, embodied sense of who we are and who we want to be (Komorowska, 2019, p. 17). It could be, therefore, argued that a particular expression of self is a manifestation of the all-embracing identity, which might be in dialogue with different types and dimensions of self.

The importance of imagination in discussing language identity and its role in identity formation has been strongly emphasized by Pavlenko and Blackledge (2004). Also, the concept of *imagined identity*, as defined by Norton (2013), closely corresponds with the concept of *imagined community* (Anderson, 1991), with the latter operating within the symbolic dimension of group identity and group cohesiveness. The impact of imagined communities on the learning trajectories of L2 learners has been recently explored in more depth (Norton & Pavlenko, 2019). From this perspective, *an imagined L2 self* can be understood as a realization of being part of an imagined community of L2 users, and this manifests itself in how language learners imagine, perceive and talk about themselves as L2 users.

Much emphasis is placed on the symbolic power of language use inside and outside an L2 classroom (for an overview see Kramsch, 2021). As Claire Kramsch (2021) correctly points out, "language educators have an additional responsibility to teach something about the symbolic power of language as discourse – how it works, how it affects people, how they can harness it to represent themselves and the reality that surrounds them, to act upon it, and to create future possible selves" (p. 201). The language learning objectives have, thus, changed and shifted, from the mere mastery of the language system, towards a more complex understanding of how much power is gained by learning and using a given language.

This interrelationship between language, power, identity and education has also impacted the way the L2 learners' willingness to learn is understood and defined. Ushioda (2011) critically examines the relevance of integrative motivation, with reference to the complexity and ambiguity of what constitutes the target language community. Assuming that language communities are indeed imagined communities, the core motivation to learn a language is not to communicate with representatives of a given country or region – the learner's motivation is propelled by the inner need to become part of an imagined, global community and is measured by the level of *investment* in which a student wishes to engage (McKinney & Norton, 2008). In fact, L2 learning has become more deterritorialised and more context dependent, where learners "invest in learning because they know that they will acquire a wider range of symbolic and material resources, and these social and economic gains in turn enhance the range of identities they can claim in a particular community" (Darvin & Norton, 2014, p. 57).

In this sense, imagined identity is very much dependent on the complex and dynamic intricacies of the imagined language communities. From the perspective of L2 education, it seems important to seek the answer to the question of how the process of language learning is affected by the liminal and dialogic nature of the imagined communities (Norton, 2013, p. 8). It is, thus, of interest to explore L2 learners' manifestation of their imagined identity – their *imagined L2 self*, that is how they see, perceive and imagine themselves as L2 users.

There is a substantial body of research on language identity, which predominantly revolves around three key themes: focusing on L2 teachers' identities (see, e.g., Duff & Uchida, 1997; Varghese et al., 2005; Menard-Warwick, 2008; Johnson & Golombek, 2011; Gabryś-Barker, 2012; Mercer et al., 2016), exploring the identities of bi/multilingual minority language speakers (see, e.g., Khilkhanova & Khilkhanov, 2004; King & Ganuza, 2005; Mazak, 2012; Dołowy-Rybińska, 2016), and investigating migrant children's transnational identities (see, e.g., Darvin & Norton, 2014; Evans & Liu, 2018). When it comes to the language identities of university students in their academic contexts, the vast majority of researchers investigate the writer's identity formation and expression specifically in the process of producing academic texts (see, e.g., Canagarajah, 2004, 2015; Lehman, 2014; Hryniuk, 2018; Furman & Aleksandrzak, 2023). However, there are relatively few research studies on L2 learners' imagined identities in formal settings, particularly those of L2 students majoring in foreign languages, with the notable exception of Yamamoto (2017) and Gabryś-Barker (2019).

Yamamoto (2017) adopted an ethnographic approach and designed a narrative inquiry involving a series of interviews, spanning the course of two years, which looked for possible dimensions of the L2 imagined identity. The author presents a case study of one participant, a Japanese student majoring in English,

who described in detail the complex and dynamic process of identity formation. The study shows that the participant's imagined L2 self evolved during interactions with other members of the imagined community; but more importantly, the process also affected her L1 identity (Japanese) – the participant adopted a more critical and reserved attitude towards her L1. Additionally, her level of investment changed and appeared to be more socially oriented (Yamamoto, 2017).

Another study that aimed at investigating university students' L2 identity was designed by Gabryś-Barker (2019), who invited 28 multilingual English major students to answer questions and reflect on their language identity. All the participants were pre-service teachers and all spoke at least three languages: Polish (L1), English (L2), and German (L3). The results demonstrate that the participants saw their mother tongue as the most emotionally loaded and internalized; their second language identity was predominantly connected to personal and intellectual growth; finally, their third language was often seen as a challenge and their L3 identity as still growing and evolving, without any clearly defined qualities.

These two studies offer an interesting insight into how language identities are manifested by university students. Nevertheless, there seems to be a need for more data that might support the tendencies observed. The research project introduced and discussed below aims at further exploring the L2 imagined selves of undergraduate students majoring in foreign languages and bridging the research gap.

## Research Design

The main research assumption that underpins the present research project was that the students majoring in foreign languages, due to the fact that they are immersed in the language and culture of their subject, can be considered multilingual or even translingual, as they fluently move in and out of different codes and meanings. They are also expected to be more sensitive to and aware of the subtleties of language use. In this sense there is a need to navigate a more complex language identity.

The major aim of the study was to learn more about the undergraduate students' imagined L1 and L2 selves, by investigating two dimensions of their complex language identity, that is, the participants' perspective on and their emotional response to the languages they speak. In other words, the main objective was to explore how the students view the languages and how they feel expressing themselves in those languages, as well as how they imagine themselves as L1 and L2 users. Secondly, as the study was addressed to university students

majoring in three different languages, a secondary aim would be to look for possible differences in the way they consider a particular language.

This study aims at answering the following research questions:

*RQ1: How do the university students imagine themselves when using their L1 and L2s?*

*RQ1a: How do the undergraduate students perceive their L1 and L2s?*

*RQ1b: How do they feel when speaking those languages?*

*RQ2: What are the major differences between the imagined selves described by the English, German and Swedish major students?*

Purposive sampling was used for the purposes of the study (see Rallis & Rossman, 2009). The participants had to meet specific criteria to fill in the questionnaire: they had to be full-time undergraduate students of foreign language programmes at Polish universities, majoring in either English, German or Swedish. A link to a questionnaire was sent to three major universities in Poland. The participation in the study was voluntary. In the end, a total of 200 undergraduate students of foreign languages participated in the project, with 88 students majoring in English, 67 students majoring in German and 45 students majoring in Swedish. The participants were approximately the same age: the average age was 22 in the English major group, 21 in the German major group, and 20 in the Swedish group. In terms of the participants' language repertoire, Table 1 illustrates the number of languages spoken by the students. As can be seen, the group of Swedish major students seem to exhibit rich language repertoires, with the vast majority of students reported speaking three foreign languages or more—in some cases speaking six languages.

**Table 1**

*The Participants' Linguistic Repertoire*

| | ENG major students (*N* = 88) | GER major students (*N* = 67) | SWE major students (*N* = 45) |
|---|---|---|---|
| One L2 | 33% (*N* = 29) | 12% (*N* = 8) | 0% (*N* = 0) |
| Two L2s | 47% (*N* = 41) | 63% (*N* = 42) | 18% (*N* = 8) |
| Three L2s or more | 20% (*N* = 18) | 25% (*N* = 17) | 82% (*N* = 37) |

All the participants were enrolled on 3-year BA programmes in foreign languages, programmes that are specific to language studies in Poland. A student majoring in a foreign language at a Polish university is obliged to attend courses that introduce elements of literature, linguistics, as well as the history and culture of the region, in addition to developing practical language skills. Most

of the courses are predominantly conducted in the target language. The foreign language students typically write their diploma papers in the target language, exploring different areas of the language. In this way, the intention is that the students are immersed in the language and culture of the language.

The data was collected by means of three almost identical questionnaires, designed for English major students (inquiring about Polish and English), for German major students (inquiring about Polish and German), and for Swedish major students (inquiring about Polish and Swedish). The instrument consisted of two sections: (a) close-ended questions targeting the students' beliefs about their language identity in general terms, together with selected demographic data (a total of nine questionnaire items), and (b) four open-ended questions focusing on the participants' perception of the languages they speak and the way they feel about them. The theoretical framework behind the design of the open-ended items was based on the premises outlined in Pavlenko and Blackledge (2004), who emphasized the subjectivity of the process of identity formation, which is grounded in "an attempt to self-name, to self-characterize, and to claim social spaces" (Pavlenko & Blackledge, 2004, p. 19). The design of the tool was informed by the research strategies put forward by Block (2010); hence, the open and subject-oriented nature of the open-response items offered the participants a platform to freely express their own ideas connected to their imagined selves (Brown, 2009). For the purposes of this paper, the analysis and findings will relate only to the students' answers to the open-ended items, which allowed for the collection of qualitative data.

Thus, a qualitative approach was adopted to consider and analyse the collected data using a thematic analysis. Based on the literature review (Pavlenko & Blackledge, 2004; Norton, 2013; Pavlenko, 2013; Darvin & Norton, 2014; Norton & Pavlenko, 2019), key identity markers were identified and organized into categories. As a result, a framework was designed, consisting of four distinctive thematic groups:

a) linguistic (i.e., those referring to the students' knowledge of language as a complex system);
b) pragmatic (i.e., those referring to language learning and language use);
c) affective (i.e., those referring to the students' attitudes and emotions);
d) symbolic (i.e., those referring to concepts such as group identity, community, power dynamics and status).

Before the data analysis was undertaken, the respondents' answers were organized into categories, informed by the language of their major and the questions to which they responded, and were later transferred into dedicated files. The students' answers were first categorised as positive, negative or neutral. If an answer was more complex or ambiguous, a note was made next to the item, and a separate category was created. The final list was subsequently organized according to their frequency of occurrence. Finally, each item was coded

and assigned to a given theme. The codes that emerged in the process of data analysis were later visually presented in the squared graphs, where the size of the squares corresponds with the frequency of appearance of a given code. Throughout the process of designing, collecting and analysing the data, the guidelines found in Brown (2009) were followed.

## Exploring the Imagined Selves of Undergraduate L2 Students in Poland—Results

### English Major Students' Perception of L1 and L2

When considering the perception of their L1 and L2, the English major students highlighted many specific details in the description of their L1 (Polish), including the sound and melody of the language (see Table 2). The vast majority of the students recognized its complexity and emphasized the difficulty of learning and speaking the language. In the affective category, positive adjectives dominated, emphasizing its beauty and poetic potential, although there were also instances of more critical comments and attitudes. With regard to the overall character of the codes, out of the total of 188 code occurrences, 51% were classified as negative, 27% as positive and 22% as neutral.

**Table 2**

*English Major Students' Perception of L1 (PL)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| complicated: 22<br>stiff/hermetic: 5<br>grammar: 2<br>diverse: 2<br>precise: 2<br>complex: 2<br>rules/exceptions: 2<br>codes occurring once: *descriptive*, *fast*, *flowery*, *gender-based*<br><br>codes relating to sound:<br>melodic/rhythmic: 5<br>crisp: 4<br>harsh (sound): 3<br>codes occurring once: *monotonous* | difficult/demanding/hard: 53<br>flexible/creative: 4<br>expressive: 2<br>codes occurring once: *inflexible*, *good to know* | beautiful: 10<br>poetic: 6<br>exceptional: 5<br>interesting: 5<br>emotional: 4<br>rich: 3<br>pride: 3<br>unusual: 2<br>unique: 2<br>conservative: 2<br>less interesting/boring: 2<br>serious: 2<br>distinguished/elegant: 2<br>codes occurring once: *important*, *rudimental*, *mundane*, *poor*, *inelegant*, *strange*, *mean*, *charmful*, *potential*, *square*, *sentimental* | national/nationality: 4<br>mother tongue: 4<br>codes occurring once: *identity*, *tradition*, *unpopular*, *less useful*, *natural*, *official* |

As regards the perception of the foreign language studied, the English major students were even more specific in expressing how they saw the language (see Table 3). The symbolic category dominated in terms of code occurrences—the students recognized the global status of the language and considered its positive influence. They perceived English to be an easy language to learn and communicate, which sounds nice and opens doors with respect to their future. They also appreciated its simplicity and their ability to express themselves easily in English. It could be even stated that in some cases the students expressed more positive attitudes toward the language, as compared to their L1—out of the total of 192 code occurrences, the general perception of English was undeniably more positive (47%) than negative (5%).

**Table 3**

*English Major Students' Perception of L2 (ENG)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| simple/less complicated: 6<br>complicated: 4<br>dynamic: 3<br>unlimited/free: 3<br>structured/organized: 2<br>diverse/diversified: 2<br>codes occurring once: *focused on I*, *unemotional*, *illogical*, *humour*, *impersonal*, *precise*, *adaptive*<br><br>codes relating to sound:<br>melodic: 8<br>pleasant: 6<br>sounds better/great/ nice: 5<br>rhythmic: 3<br>soft: 3<br>codes occurring once: *delicate* | easy/easier: 17<br>flexible/creative: 6<br>difficult/demanding: 5<br>communicative: 3<br>almost native: 3<br>straightforward/direct: 2<br>learning/knowledge: 2<br>codes occurring once: *scientific*, *intuitive* | beautiful: 6<br>interesting: 6<br>rich: 4<br>fluid: 4<br>dignified/distinguished/ elegant: 3<br>serious/official/ stiff: 2<br>codes occurring once: *fashionable*, *pragmatic*, *expressive*, *concrete*, *light*, *wonderful*, *friendly*, *overrated* | universal: 17<br>useful: 10<br>international/ Lingua Franca: 9<br>popular: 8<br>omnipresent: 6<br>necessary: 5<br>global: 4<br>opens the door/ gives possibilities/ future: 4<br>important: 2<br>codes occurring once: *known* |

## English Major Students' Imagined L1 and L2 Selves

Having collected and analysed the answers to the question about how they felt in a given language, two profiles of imagined selves could be identified. As it can be seen in Figure 1[1], the students felt predominantly positive when speaking their native language, which comes as little surprise. The results, however, are more interesting when juxtaposed with the profile of the students' imagined L2 self (see Figure 2).

[1] The size of the square corresponds with the number of code occurrences.

**Figure 1**

*English Major Students' Imagined L1 Self—A Profile*

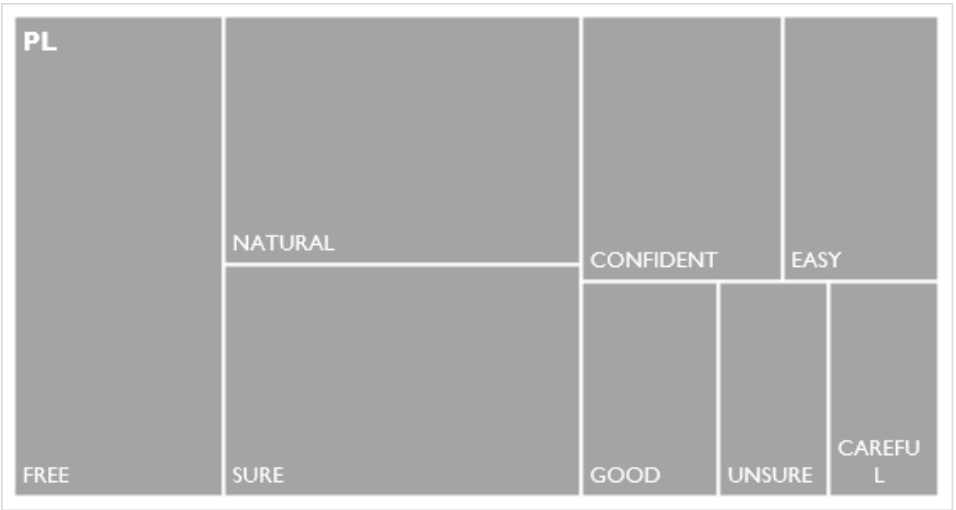
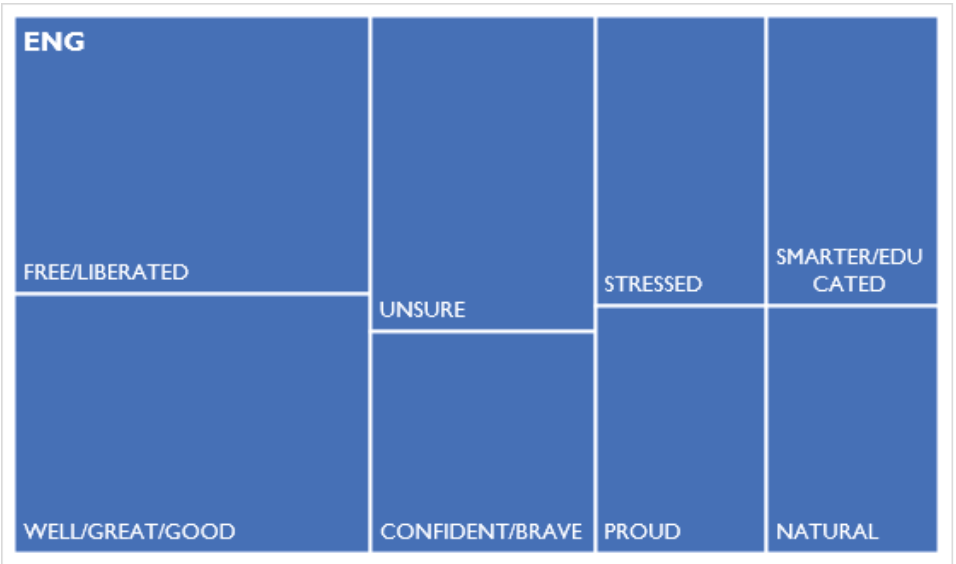



**Figure 2**

*English Major Students' Imagined L2 Self—A Profile*



In their imagined L2 selves, the English major students expressed themselves with the use of more positive and more emotionally loaded adjectives. As compared to the profile of their L1 imagined self, speaking English made them feel smarter and educated, confident and brave, as well as proud. They emphasized being fluent—not only free, but also liberated to express how they think and feel—as an important element of their L2 identity. In the analysis of their responses there were nine instances of positive comparative adjectives, such as *more*, *better*, *smarter*, *closer*, *funnier*. Their imagined L2 self reflected their positive attitude to the studied language and also corresponded with its symbolic dimension, that is, English being the language of possibilities and "open doors."

## German Major Students' Perception of L1 and L2

The German major students focused predominantly on the linguistic aspects of their L1 (see Table 4). They offered a very detailed analysis of the language in terms of its grammar, lexis and phonology. Similarly to the English major students, they agreed that Polish is complicated, difficult to master and use to communicate. Its lower status was also recognized by the participants. In general, the students were rather critical about their native language, as out of the total of 125 code occurrences, 38% were classified as negative, only 8% as positive and 54% as neutral.

**Table 4**

*German Major Students' Perception of L1 (PL)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| complicated: 28<br>exceptions: 4<br>simple: 2<br>diverse: 2<br>illogical: 2<br>codes occurring once: *wordy*, *complex*, *dynamic*, *conjunction*, *rich in vocab*, *cases*, *specific*, *unpredictable*<br><br>codes relating to sound:<br>crisp: 3<br>codes occurring once: *sounds nice*, *hard sounds*, *sound*, *loud* | difficult: 39<br>codes occurring once:<br>*easy* | interesting: 3<br>unique: 2<br>codes occurring once: *poetic*, *uninteresting*, *obvious*, *no sympathy*, *free*, *unimportant*, *problematic*, *conservative* | native: 9<br>less useful/useless: 2<br>home: 2<br>codes occurring once: *less known*, *culture*, *traditional*, *natural*, *developing*, *historically heavy* |

When considering the German major students' perception of the studied language, it could be observed that they were much more positive about German, as compared to Polish (see Table 5), with 22% positive, 61% neutral and 17% negative codes. Although they considered it a difficult language, they pointed to a number of positive aspects, admitting it was logical, well-organized, interesting and even useful. Similarly to their L1, the students concentrated on the linguistic features of the German language, going into much detail about its syntax. Interestingly, despite the fact that the German major students were asked about their L1 and L2 (i.e., German) only, they frequently compared German to English.

**Table 5**

*German Major Students' Perception of L2 (GER)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| logical: 15<br>order/orderly/organized: 11<br>concrete/precise: 6<br>simple: 6<br>dynamic: 5<br>rules: 3<br>schematic: 3<br>complex: 3<br>transparent: 3<br>specific: 3<br>codes occurring once: *direct*, *word formation*<br><br>codes relating to sound:<br>heavy (sound): 5<br>nice sound: 4<br>codes occurring once: *ugly*, *aggressive sound*, *less melodic* | difficult/demanding: 16<br>easy/*easier: 11<br>[*easier than English]<br>codes occurring once: *business language*, *intuitive*, *creative*, *learning*, *almost native*, *easier to make a mistake* | interesting: 10<br>pleasant/nice: 3<br>strong: 2<br>beautiful: 2<br>complicated: 2<br>codes occurring once: *original*, *characteristic*, *character*, *ugly, rich, wonderful*, *diverse* | useful: 4<br>opens possibilities/ future: 3<br>important: 2<br>codes occurring once: *unpopular*, *dominated by English*, *part of great*, *culture*, *traditional*, *foreign*, *recognizable* |

## German Major Students' Imagined L1 and L2 Selves

The imagined L1 selves of German major students are comparable to the English major students. They reported feeling free, natural, confident and able to express themselves well (see Figure 3). Their imagined L2 self, in contrast, could be described as a mixture of negative and positive feelings. On the one hand, the students admitted to feeling anxious, particularly about making mistakes and being accurate, unsure and limited. At the same time, however, they felt good, happy, wise and free (see Figure 4).

**Figure 3**

*German Major Students' Imagined L1 Self—A Profile*

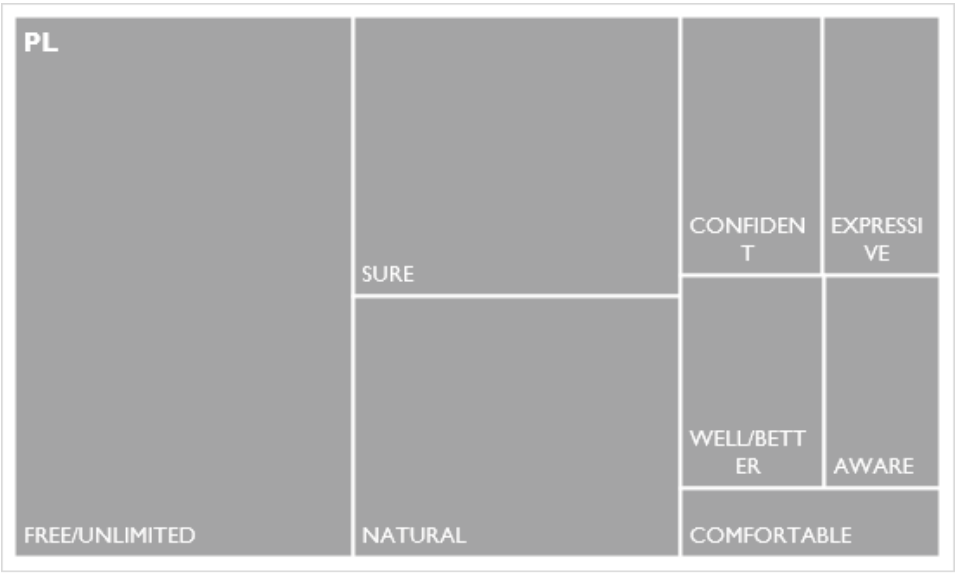
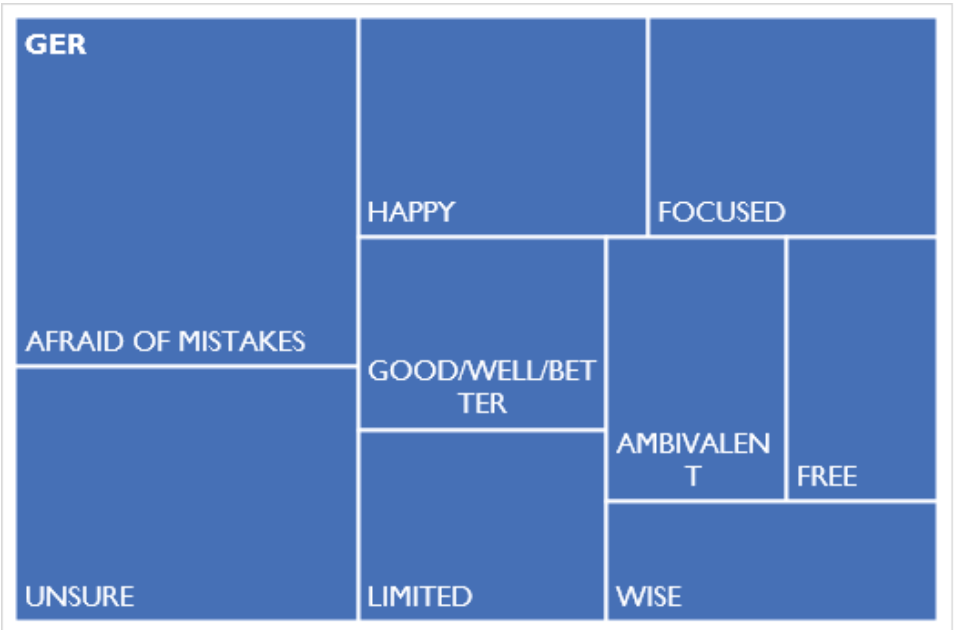



**Figure 4**

*German Major Students' Imagined L2 Self—A Profile*



What should be noted here is that while analysing their responses, there were a number of instances where the students expressed having mixed feelings when speaking German. For example, they reported to have felt "different, but good," or "less sure, but well." Six such instances were coded collectively as a form of ambivalence.

## Swedish Major Students' Perception of L1 and L2

The Swedish major students described their L1 as complicated and difficult, yet beautiful and interesting. What should be highlighted is that they considered speaking and knowing Polish to be beneficial and conducive to learning other languages. Overall, the characteristics of the total number of 88 code occurrences were spread almost evenly, with 34% of the codes labelled as positive, 35% neutral and 31% negative.

**Table 6**

*Swedish Major Students' Perception of L1 (PL)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| complicated: 13<br>rich: 6<br>complex: 5<br>irregular: 2<br>codes occurring once: *non-inclusive*, *descriptive*, *synthetic*, *diverse*, *exceptions*<br><br>codes relating to sound:<br>sounds nice: 2<br>heavy sounds: 2<br>*sound*, *monotonous (sound)* | difficult: 20<br>good foundation for learning: 3<br>expressive: 2<br>codes occurring once:<br>*flexible* | beautiful: 8<br>interesting: 6<br>different: 2<br>codes occurring once:<br>*pretentious*, *emotional*, *unique* | familiar/natural: 3<br>codes occurring once:<br>*traditional*, *taboo*, *less relevant* |

In contrast, their perception of their L2 could be easily viewed as enthusiastic (see Table 7). Swedish was perceived as not only simple and easy to learn, but also as interesting and beautiful. Out of the total of 101 code occurrences, 24% referred to sound—how melodic and nice to the ear the language was. The students expressed solely positive feelings about Swedish, as can be seen in the affective category of the codes. Interestingly, there were four instances in which Swedish was compared to other languages (including Polish, English and German). It was apparent that the students felt excited to be asked about Swedish—84% of the code occurrences were positive, with only 3% negative. Compared to the previous two groups of students, they seemed the most passionate about their chosen language.

**Table 7**

*Swedish Major Students' Perception of L2 (SWE)*

| LINGUISTIC | PRAGMATIC | AFFECTIVE | SYMBOLIC |
|---|---|---|---|
| simple (simpler than Polish or German ×3): 9<br>logical: 7<br>grammatical: 2<br>specific: 2<br>regular: 2<br>codes occurring once: *analytical*, *a mix of PL/ENG/GER*<br><br>codes relating to sound:<br>melodic: 17<br>sounds interesting: 4<br>nice for the ear: 3 | easy (to learn): 11<br>difficult: 3<br>codes occurring once: *methodic*, *practical*, *precise*, *straight-forward* | interesting: 12<br>nice: 5<br>beautiful: 3<br>funny: 3<br>unique: 3<br>codes occurring once: *exotic*, *joyful*, *rich*, *unusual*, *the best* | universal and useful (in Scandinavia): 2<br>codes occurring once: *self-development*, *nature* |

## Swedish Major Students' Imagined L1 and L2 Selves

The Swedish major students' profile of their imagined L1 self (see Figure 5) is very similar to the profiles of the German and English major students. In general, they felt natural, free, proud and confident in speaking Polish. Yet, looking at the profile of their imagined L2 self (see Figure 6), it is clear that their excitement observed in the previous section was transferred to this profile. They explicitly stated that they felt happy and excited speaking Swedish—pride also seemed to be an important component. What is interesting is the fact that the Swedish profile represents a mix of the elements present in the English and German profiles; for example, the Swedish major students observed that they feel more intelligent or wise speaking Swedish (see the L2 imagined self of the students of English). Comparing this profile to the German major students, the Swedish group admitted to sometimes feeling ambivalent in speaking Swedish (e.g., *unsure yet open*, *good yet unsure*, *unsure but happy*) as well as feeling anxious about committing mistakes in their target language.

**Figure 5**

*Swedish Major Students' Imagined L1 Self—A Profile*

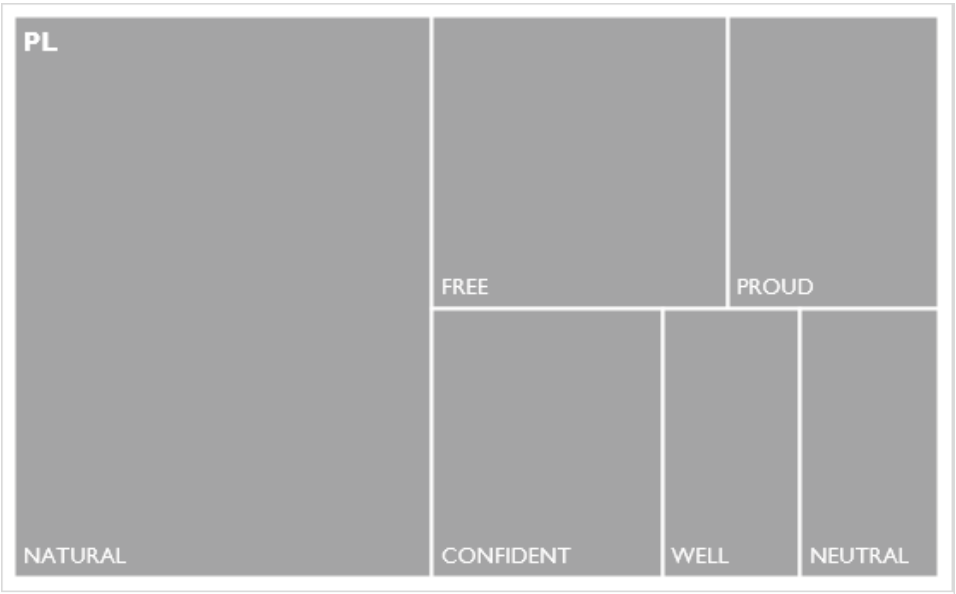
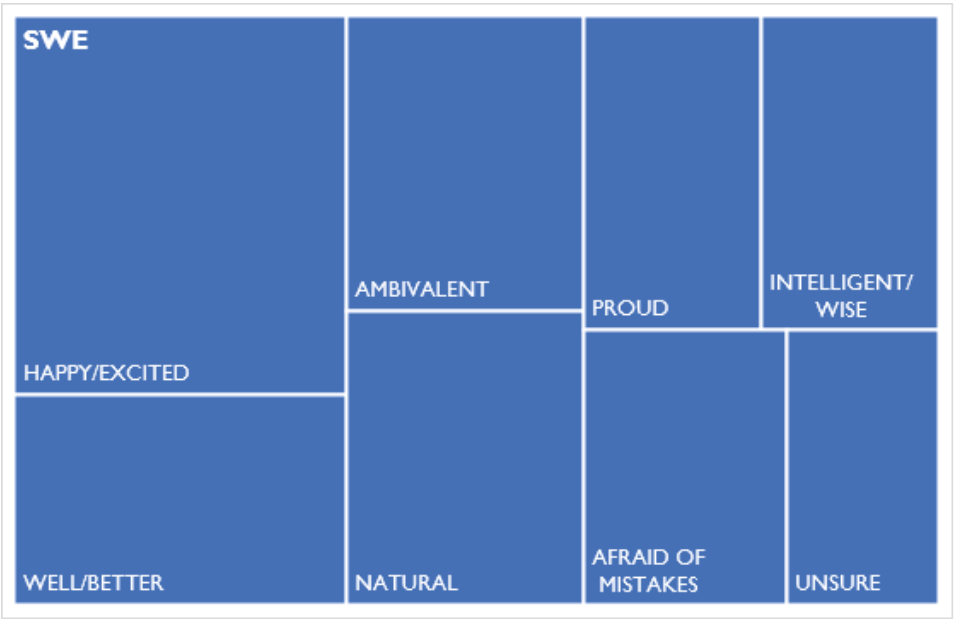



**Figure 6**

*Swedish Major Students' Imagined L2 Self—A Profile*



## Discussion

The research findings offer an interesting insight into the language identity and the imagined L1 and L2 selves of undergraduate students of three language majors: English, German and Swedish. Addressing the first part of the first research question (RQ1a), which enquired about the students' perception of their languages in question, it can be stated that they were aware language users, who often referred to the complex nature of the language systems and their structures. They seemed to be very sensitive to the way the languages sound (a total of 88 code occurrences in all three groups) and how the languages are used and learned (see the pragmatic category). The English major students generally considered their L2 to be a universal language which makes international communication easier. The German major students saw their L2 as demanding, yet logical and well-organised. The Swedish major students perceived their L2 as simple, easy and interesting. In the affective category, it can be observed that the students were much more positive about their target languages, as compared to their perception of their L1, about which they seemed undoubtedly more critical (see Pavlenko, 2013).

In terms of the second part of the first research question (RQ1b), which focused on the participants' feelings both experienced and expressed in using their L2s, it could be observed that the students felt, in general terms, better speaking their languages of choice, as compared to their L1 profiles. There were many instances of positive comparative adjectives, like *better*, *more intelligent*, *smarter*, *funnier*, *happier*, in the analysis of the profiles, with the Swedish group in particular feeling enthusiastic about Swedish. In addition, more students mentioned a sense of pride when speaking their L2s; in fact, only the Swedish group admitted to feeling proud in their L1.

It is interesting that the German major students seemed more emotional about their language of study, which stands in contrast with the findings in Gabryś-Barker (2019). Yet, their emotions describing their experience in speaking their L2 were not always simply positive or negative. Both the German and Swedish students exhibited instances of ambivalence (see Block, 2010; Darvin & Norton, 2014), where they expressed often contradictory emotions about their target languages.

Finally, concerning the differences in the way the students imagined themselves in their foreign languages (see RQ2), based on their L2 profiles, the English major students were much more focused on fluency—the possibility to express themselves freely and without any limitations; whereas the German and Swedish students concentrated more on accuracy and their fear of making mistakes. It would be interesting to further explore the potential influence of their perception of the imagined communities (Anderson, 1991) on the way

they express their imagined L2 selves. In other words, it could be argued that their language identity might be to some extent affected by the stereotypical images of the target language communities.

As the symbolic dimension of the language learners' L2 identity was highlighted (McKinney & Norton, 2008; Kramsch, 2021), it should also be emphasized that the students in all three categories recognized the higher status of English. The English major students explicitly addressed this issue by referring to English as a global language, which is not only popular but also necessary. Still, the symbolic status of English was also recognized by the other two groups. Although the German and Swedish students were only asked about their L1 and L2 languages, they often compared German and Swedish to English, for example, by observing that German is dominated by English. This might be accounted for by the fact that both German and Swedish major students were more multilingual and exhibited more sensitivity to the differences between languages and language use, with the Swedish students speaking, on average, more than three foreign languages. Despite the fact that English occupies the position of a *lingua franca*, the participants were also clear about the "social and economic gains" (Darvin & Norton, 2014, p. 57) that speaking foreign languages other than English could offer.

It can also be observed that the way the English and German major students perceived their languages of study overlapped with the common stereotypes about these languages and their L2 users, with English being connected to liberation, freedom and fluency, and German being associated with order, logic and accuracy. Only the participants studying Swedish did not attribute their language identity with commonly associated stereotypical traits, focusing only on the positive aspects of learning and speaking the language. They were, among these three groups, definitely the most excited about their L2s.

## Concluding Remarks

To conclude, the students of foreign language majors in Poland exhibited very positive imagined L2 selves and seemed passionate and excited about their languages of choice, even though their study might be challenging and demanding. What is surprising is that the students were rather critical about their L1, in contrast to their perception of their L2s. However, in general, the participants could be considered aware language users, recognizing the symbolic power behind the language use. The English and German major students' imagined L2 selves to some extent reflect the common stereotypes of the languages and

L2 users, but it would be interesting to explore some of the threads and patterns that emerged in the analysis in more detail.

When considering the research findings, there are three major teaching implications that could be drawn from this study. First, since the students exhibited highly positive attitudes towards their imagined L2 selves, there is, for example, a need for fostering the students' intrinsic, integrative motivation and supporting their investment in learning by offering them opportunities to participate more actively in the imagined communities of practice. Second, it is important to further explore to what extent the stereotypes language learners have about the languages they study affect their imagined L2 selves. It would be, thus, necessary to raise the students' (inter)cultural awareness and challenge the stereotypical views of the languages and the L2 users. Third, considering the complexity of the students' imagined L2 self profiles, it might be worthwhile to give them platform for expressing and exploring their multilingual identities by encouraging language journals or language biographies.

The present study is, however, not without certain limitations. First of all, it offers merely a snapshot of the issue under discussion—with only a single tool, there was limited amount of data that could be gathered, which resulted in rather simplified profiles and limited contexts. Due to the exploratory nature of the open-ended questions (Brown, 2009), the findings of the study should be treated more as the groundwork for future projects. What is more, the tool allowed the students' spontaneous associations with their languages to be captured at a specific moment in time—their language biography and prior experience was not taken into account. Finally, as the collected data captures a static "moment" in their complex language identities, it should be borne in mind that it is necessary to treat the results as guidelines for future research questions in an ethnographic project.

Since one of the primary aims of the study was to establish a foundation for future studies, some of the issues that emerged from the analysis of the gathered material could be further investigated by means of ethnographic projects or narratives, for instance, to explore the students' profiles in context and to observe how their language identities evolve over time. In particular, it would be interesting to learn about the role of the L2 teachers in the process of shaping students' imagined L2 selves, as well as the correlation between the students' language identity and their L2 motivation. Another intriguing aspect to explore would be the cultural dimension that emerged in the findings—in other words, to what extent the profiles are affected by the stereotypical thinking about the imagined communities of the languages under investigation, and to what extent this simplified image of a target language user is fostered within formal education in Poland. Finally, the results should be enriched by collecting data about languages from other language families.

# References


Anderson, B. (1991). *Imagined communities: Reflections on the origin and spread of nationalism*. Verso.

Block, D. (2010). Researching language and identity. In B. Paltridge & A. Phakti (Eds.), *Continuum companion to research methods in applied linguistics* (pp. 37–347). Continuum.

Brown, J. D. (2009). Open-response items in questionnaires. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics* (pp. 200–219). Palgrave Macmillan.

Canagarajah, S. (2004). Multilingual writers and the struggle for voice in academic discourse. In A. Pavlenko & A. Blackledge (Eds.), *Negotiation of identities in multilingual contexts* (pp. 266–289). Multilingual Matters.

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

Darvin, R., & Norton, B. (2014). Transnational identity and migrant language learners: The promise of digital storytelling. *Education Matters: The Journal of Teaching and Learning, 2*(1), 55–66.

Darvin, R., & Norton, B. (2015). Identity and a model of investment in applied linguistics. *Annual Review of Applied Linguistics, 35*, 36–56.

De Costa, P. I., & Norton, B. (2016). Identity in language learning and teaching: Research agendas for the future. In S. Preece (Ed.), *The Routledge handbook of language and identity* (pp. 586–601). Routledge.

Dołowy-Rybińska, N. (2016). Language attitudes and community engagement: Diwan–The Breton immersion high school through the eyes of its pupils. *Journal of Language, Identity & Education, 15*(5), 280–292. https://doi.org/10.1080/15348458.2016.1213134

Douglas Fir Group (2016). A transdisciplinary framework for SLA in a multilingual world. *The Modern Language Journal, 100*, 19–47.

Duff, P. A., & Uchida, Y. (1997). The negotiation of teachers' sociocultural identities and practices in postsecondary EFL classrooms. *TESOL Quarterly, 31*(3), 451–486. https://doi.org/10.2307/3587834

Evans, M., & Liu, Y. (2018). The unfamiliar and the indeterminate: Language, identity and social integration in the school experience of newly-arrived migrant children in England. *Journal of Language, Identity & Education, 17*(3), 152–167. https://doi.org/10.1080/15348458.2018.1433043

Furman, M., & Aleksandrzak, M. (2023). The reception and representation of English as an academic language among foreign students in Poland and its influence on (re)shaping individual identities. *Glottodidactica, 50*(1), 97–117. https://doi.org/10.14746/gl.2023.50.1.6.

Gabryś-Barker, D. (2012). *Reflectivity in pre-service teacher education. A survey of theory and practice*. Wydawnictwo Uniwersytetu Śląskiego.

Gabryś-Barker, D. (2019). Studying bilingual and multilingual language identities: Natural settings versus formal instruction. *Linguistica Silesiana, 40*, 341–359. https://doi.org/10.24425/linsi.2019.129417

Hryniuk, K. (2018). Expressing authorial self in research articles written by Polish and English native-speaker writers: A corpus-based study. *Studies in Second Language Learning and Teaching, 8*(3), 621–642. https://doi.org/10.14746/ssllt.2018.8.3.5.

Huang, J., & Benson, P. (2013). Autonomy, agency and identity in foreign and second language education. *Chinese Journal of Applied Linguistics, 36*(1), 7–28.

Johnson, K. E., & Golombek, P. R. (Eds.). (2011). *Research on second language teacher identity: A sociocultural perspective on professional development*. Routledge.

Khilkhanova, E., & Khilkhanov, D. (2004). Language and ethnic identity of minorities in Post-Soviet Russia: The Buryat case study. *Journal of Language, Identity & Education, 3*(2), 85–100. https://doi.org/10.1207/s1532770ljlie0302_1

King, K., & Ganuza, N. (2005). Language, identity, education, and transmigration: Chilean adolescents in Sweden. *Journal of Language, Identity & Education, 4*(3), 179–199. https://doi.org/10.1207/s1532770ljlie0403_1

Komorowska, H. (2019). Languages and the self. *Neofilolog, 52*(1), 15–28. https://doi.org/10.14746/n.2019.52.1.3

Kramsch, C. (2021). *Language as symbolic power*. Cambridge University Press.

Lehman, I. M. (2014). The role of identity in student writing in English and Polish. *Kwartalnik Neofilologiczny, 61*(3), 601–611.

Mazak, C. M. (2012). My cousin talks bad like you: Relationships between language and identity in a rural Puerto Rican community. *Journal of Language, Identity & Education, 11*(1), 35–51. https://doi.org/10.1080/15348458.2012.644119

McKinney, C., & B. Norton (2008). Identity in language and literacy education. In B. Spolsky & F. Hult (Eds.), *The handbook of educational linguistics* (pp. 192–205). Blackwell,.

Menard-Warwick, J. (2008). The cultural and intercultural identities of transnational English teachers: Two case studies from the Americas. *TESOL Quarterly, 42*(4), 617–640. https://doi.org/10.1002/j.1545-7249.2008.tb00151.x

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

Norton, B. (1997). Language, identity, and the ownership of English. *TESOL Quarterly, 31*(3), 409–429. https://doi.org/10.2307/3587831

Norton, B. (2006). Identity as a sociocultural construct in second language education. *TESOL in Context (Special issue)*, 22–33.

Norton, B. (2013). *Identity and language learning (2nd edition).* Multilingual Matters.

Norton, B., & Pavlenko, A. (2019). Imagined communities, identity, and English language learning in a multilingual world. In X. Gao (Ed.), *Second handbook of English language teaching* (pp. 703–718). Springer.

Pavlenko, A., & Blackledge, A. (2004). Introduction: New theoretical approaches to the study of negotiation of identities. In A. Pavlenko & A. Blackledge (Eds.), *Negotiation of identities in multilingual contexts* (pp. 1–33). Multilingual Matters.

Pavlenko, A. (2013). The affective turn in SLA: From 'affective factors' to 'language desire' and 'commodification of affect.' In D. Gabryś-Barker & J. Bielska (Eds.), *The affective dimension in second language acquisition* (pp. 3–28). Multilingual Matters.

Rallis, S. F., & Rossman, G. B. (2009). Ethics and trustworthiness. In J. Heigham & R. A. Croker (Eds.), *Qualitative research in applied linguistics* (pp. 263–287). Palgrave Macmillan.

Tajfel, H., & Turner, J. (1986). The social identity theory of intergroup behavior. In S. Worchel & W. Austin (Eds.), *Psychology of intergroup relations* (pp. 7–24). Nelson-Hall.

Ushioda, E. (2011). Language learning motivation, self and identity: Current theoretical perspectives. *Computer Assisted Language Learning, 24*(3), 199–210. https://doi.org/10.1080/09588221.2010.538701

Varghese, M., Morgan, B., Johnston, B., & Johnson, K. A. (2005). Theorizing language teacher identity: Three perspectives and beyond. *Journal of Language, Identity & Education, 4*(1), 21–44. https://doi.org/10.1207/s1532770ljlie0401_2

Yamamoto, K. (2017). Imagined community, imagined self: Identity construction in language socialization outside the classroom. *Studies in Linguistics and Language Teaching, 28,* 215–240.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/21
https://doi.org/10.31261/TAPSLA.16020

CC BY SA

**Katarzyna Morena**
https://orcid.org/0000-0002-2039-5009
University of the National Education Commission, Krakow, Poland
Institute of English Studies

# Academics' Motivation to Learn Foreign Languages: The Case of English

Abstract

Teaching and research are central aspects of university scholars' work. To provide a high quality of teaching and research, scholars are required to continually learn and develop professionally. Only recently have some empirical studies that focused on academics' motivations and emotions attempted to explain key outcomes related to their research, teaching, and professional development.

This study adopted a self-determination lens to explore academics' motivation to learn foreign languages, an area that has hardly been scrutinized due to an assumption that academics are highly motivated. The sample consisted of 593 academics (330 women) from nine public and one non-public higher education institutions in Krakow, Poland. The participants also self-rated their level of proficiency in English, French, German, Spanish, and Russian.

The results from the Polish version of the Language Learning Orientations Scale—intrinsic motivation, extrinsic motivation, amotivation (LLOS-IEA) (Noels et al., 2000) showed that these university academics were characterized by both intrinsic motivation and the most internalized form of extrinsic motivation (identified regulation). These have been considered to be optimal forms of motivation with positive outcomes. Statistically significant differences were found between scholars' motivation to learn languages and age, gender, and job seniority. The analysis performed for English indicated that scholars rated themselves as having a proficient level (C1) in reading and an independent level (B2) in writing skills.

*Keywords*: academics, motivation, self-determination theory, foreign languages

University scholars are the largest producers of innovative research and contributors to scientific progress (Stupnisky, BrckaLorenz, & Laird, 2019). This scientific advancement leads to university-industry knowledge transfer,

economic activity, and data-based government decision-making (Landry et al., 2003; Perkmann et al., 2013; Weinberg et al., 2014). Effective academic teaching and research are fundamental components of informed citizenship, quality of higher education, and institutional visibility (Javitz et al., 2010). Studies of academics' career goals show that most scholars enter the profession with expectations of contributing new knowledge and research to society, doing intellectually stimulating work, and delivering high-quality teaching (Watt & Richardson, 2020).

Teaching and research are central and time-intensive aspects of university scholars' work. While teaching primarily focuses on conveying existing knowledge to others, research focuses on creating and documenting new knowledge (Daumiller & Dresel, 2020b). Effective teaching is associated with the quality of student engagement, differences in student academic achievements, and persistence (BrckaLorenz et al., 2012), whereas effective research is associated with international collaboration and/or getting published in high-impact, peer-reviewed journals (Stupnisky, BrckaLorenz, & Laird, 2019; Willetts, 2019).

Tertiary education, and particularly the scientific domain, is where English dominates over other foreign languages (Björkman, 2013). English has become the lingua franca of science (Crystal, 2006). Thus, it is the primary language to access scientific information, participate in discussions, symposia, and congresses, or collaborate with other scholars in the field (Björkman, 2013). Demands on academics continue, yet little is known about their career goals, attitudes, or motivations. One examination suggests the reason for this may be that academics constitute a smaller labor force than teachers and, therefore, may not be studied very much in general (Daumiller, Stupnisky, & Janke, 2020). As scholars comprise a smaller group of professionals, a smaller sample generates less statistically significant and generalizable results. Another rationale for the lack of research is the perception that academics are highly motivated; given the efforts invested in obtaining the necessary degrees, or the low salary, people who persevere to become scholars are assumed to be highly motivated (Daumiller, Stupnisky, & Janke, 2020). The third reason is that researching academics' motivation can be socially intimidating and methodologically challenging for the researcher. For some, approaching colleagues or their teachers and asking how motivated they are to do their job, research, or develop might be uncomfortable. Besides, fellow academics may also be too busy or disengaged to participate in such empirical research. Another challenge relates to fellow scholars' scientific expertise and researcher concerns that the research method may receive criticism, or that colleagues will deduce the aims of the research and provide desirable responses (e.g., the Hawthorne effect).

Only recently, have some empirical studies that focused on the motivations and emotions of academics attempted to explain key outcomes related to research, teaching, and professional development (e.g., Stupnisky et al., 2017,

2018; Stupnisky, BrckaLorenz, & Laird, 2019; Daumiller & Dresel, 2020a, b). A conclusion drawn from recent research indicates that to achieve in research and teaching domains, scholars need optimal motivation (e.g., Daumiller & Dresel, 2020b; Stupnisky, BrckaLorenz & Laird, 2019; Stupnisky et al., 2018). Given that an exceptionally high level of fluency in English is the key to success in the academic world, scholars need to spend considerable time and effort to develop competencies in this language (Coleman, 2006; Crystal, 2006; Horn, 2017; Lopes-Navarro, 2015; Stockemer & Wigginton, 2019). Thus, it can be speculated that scholars who are optimally motivated to learn English (as a foreign language) may enjoy participation in career and knowledge advancement, increased mobility, and thus employability. In turn, successful and motivated scholars enhance institutional prestige, attract research and development funding, and demonstrate high-quality teaching abilities based on state-of-the-art scientific knowledge (Macaro, 2018).

## Self-determination Theory and Language Learning Motivation

Self-determination theory (SDT) assumes that the propensity to be curious about one's environment and interested in learning and developing one's knowledge is inherent in human nature. Individuals, according to SDT, are innately curious creatures who possess a natural love of learning and who desire to internalize knowledge, customs, and values that surround them (Niemiec & Ryan, 2009, p. 133).

Self-determination theory is a theory of human motivation that takes an interest in factors that either facilitate or undermine the assimilative and growth-oriented process in people. SDT emphasizes different types and sources of motivation that impact the quality and dynamics of behavior, and the SDT model involves general motivational regulations (orientations) that are autonomous, controlled, or amotivated (Noels et al., 2000; Ryan et al., 2019; Ryan & Deci, 2017, 2020). SDT has provided a framework for studying motivation in multiple contexts: education, health care, sport and physical activity, psychotherapy, virtual environments, or work and organizations (Ryan & Deci, 2017).

The motivation to learn languages conceptualized through the self-determination theory lens assumes that there are different types of motivation related to how much a learner engages in (learning) activity (Noels et al., 2001). In line with SDT, Noels, Lou et al. (2019, p. 97) defined motivation as "a set of reasons for learning the language, which provides a frame of reference for the learner within which she interprets her language learning experience and directs her attention and effort." These researchers further differentiated the

reasons for learning the language into three forms: intrinsic, extrinsic, and amotivation orientations.

**Table 1**

*Self-determination Theory Taxonomy of Foreign Language (FL) Learning Motivation Based on the Language Learning Orientations Scale—IEA (Noels et al., 2000)*

| AMOTIVATION | EXTRINSIC MOTIVATION | | | INTRINSIC MOTIVATION | | |
|---|---|---|---|---|---|---|
| | EXTERNAL REGULATION | INTROJECTED REGULATION | IDENTIFIED REGULATION | KNOW | ACCOMPLISH | STIMULATE |
| feeling of incompetence, no value to engage in learning | external rewards: salary increase, promotion prospects, meet other people's expectances | ego-relevant reasons, self-pressure to perform well feel proud, or avoid shame and guilt | value of learning and using the FL, FL is an important aspect of development and identity | pleasure and satisfaction from learning, exploring, and understanding something new | pleasure and satisfaction from attempts to accomplish something new | experience stimulating sensations of cognitive pleasure |

According to self-determination theory, intrinsic motivation to learn is based on the experience of enjoyment and vitality that stems from an individual's "inherent tendency to seek out novelty and challenges, to explore, to learn" (Ryan & Deci, 2000, p. 70). Thus, language learning could be pleasurable in and of itself, even if that activity is not tied to one's sense of self (Noels, Vargas Lascano, & Saumure, 2019). According to Vallerand et al. (1992) intrinsic motivation might be differentiated into more specific motives such as motivation to know, accomplish, and experience stimulation (see Table 1).

Learners are intrinsically motivated to know when they read a book or a paper for the sheer pleasure they experience while learning or discovering something new. Intrinsically motivated individuals approach language learning with an inherent interest (Noels, Lou et al., 2019). Intrinsic motivation to know thus refers to "the fact of performing an activity for pleasure and satisfaction that one experiences while learning, exploring, and trying to understand something new" (Vallerand et al., 1992, p. 1005). Another type of intrinsic motivation is motivation towards accomplishments. Learners who extend their work beyond requirements, achieve a goal, or surpass themselves might experience intrinsic motivation towards accomplishments (Noels et al., 2000). The third type of intrinsic motivation is intrinsic motivation to experience stimulation, that is, "the fact of engaging in an activity to experience stimulating sensations of pleasure, fun, and excitement" (Vallerand et al., 1992, p. 1005).

Learners who are characterized by this form of motivation engage in learning to experience excitement and feelings of cognitive pleasure (Vallerand et al., 1992; Noels et al., 2000).

Not all individuals feel intrinsically motivated to engage in language learning (Noels, Lou et al., 2019). Although the feelings about learning languages can be similar to feelings associated with intrinsic motivation that involve a sense of accomplishment and thriving, some people do not experience feelings of pleasure or joy. Learning a new language can be personally meaningful. It may facilitate the attainment of important goals, such as improving one's occupational performance, or ability to communicate effectively within one's professional context (Noels, Lou et al., 2019). Such individuals learn languages because of identified reasons. Identified regulation occurs when an individual integrates the value of learning and using the language. At this point, individuals invest energy in learning because of personally relevant reasons; learners feel that fluency in a foreign language is an important aspect of their development and identity (Noels et al., 2000). The person will persist in learning as long as the goal is important (Noels et al., 2001). These relatively self-determined regulations can be contrasted with more controlled forms, including introjected regulation and external regulation.

Individuals might learn a foreign language less because it is important or valuable and more because they feel they should know a language. Such motives arise because the person feels internal pressure, or social circumstances require them to engage in language learning (Noels, Lou et al., 2019). It may be the case that an individual self-imposed pressure and compelled themselves to learn (Noels et al., 2000). Such individuals learn languages because of introjected reasons. According to Noels et al. (2000), introjected regulation is controlled by ego-relevant reasons such as self-pressure to perform well, or to avoid shame or guilt for not doing so. Even if the value of learning the language is not evident to a person in their personal lives, they might feel it is expected for an educated person to know the language. Learning also occurs if individuals feel the need to reduce negative feelings (e.g., guilt or shame) (Noels et al., 2000). Yet, some people engage in learning the language, not because of self-imposed pressures or because it is meaningful or congruent with their goals and identity, but because of external factors, for example, rewards such as a salary increase or promotion prospects, or because other people expect them to do so (Noels et al., 2000). Such individuals learn languages because of external reasons. Accordingly, external regulation is the most controlled regulation by sources external to the person. However, once that pressure or reward is removed, the learner might be expected to stop putting effort into language learning (Noels et al., 2001).

These four forms (integrated, identified, introjected, external) of regulations fall under the umbrella term of extrinsic motivation. The reasons (or contingen-

cies) may be more or less internalized—motives change along the continuum from external regulation to identified regulation (Noels, Lou et al., 2019).

Some individuals might see no purpose of any kind for learning the language. This type of situation is referred to as *amotivation*. In amotivation, an individual may feel incompetent and/or see no value in engaging in the activity (Noels, Lou et al., 2019). Such learners would be expected to quit at their earliest convenience (Noels et al., 2001).

Research conducted in second and foreign language learning using the SDT framework has consistently shown that language learners who endorse a self-determined orientation are more likely to engage deeply and positively with the learning process, and they are more likely to experience a positive outcome. Compared to less self-determined learners, they exhibit greater competence, confidence, and persistence, better use of learning strategies, and greater intention to continue studying the language even after the course is completed (Busse & Walter 2013; Lou & Noels, 2018; Noels et al., 2001; Pae, 2008; Vandergrift, 2005).

## Method

This study adopted a quantitative research design. The questionnaire consisted of demographic information, the LLOS-IEA scale, and language-specific questions based on the Common European Framework of Reference for Languages (2003). A self-assessment grid and detailed descriptors of the levels (A1–C2, 2017) were provided for the participants. The questionnaire was distributed in paper format and also sent out via email to the offices of the ten biggest universities in Krakow. The participation was voluntary and anonymous. A total of 616 questionnaires were returned, however, due to the incompleteness, 21 were removed from the analysis, and the final sample consisted of 593 respondents (330 females).

The participants were academics from nine public and one non-public higher education institutions in Krakow. Participants' ages ranged from 25 to 65+. The most numerous group included participants aged 35–44 ($N$ = 289), and 25–34 ($N$ = 133), whereas the least numerous group included participants 65+ ($N$ = 19). There was a broad range of titles and degrees among participants with the most numerous group being PhDs ($N$ = 265). The academics were employed as teaching staff, research staff, or both. Table 2 shows this data.

**Table 2**

*Participants' Titles, Degrees, and Positions*

| | *N* | [%] |
|---|---|---|
| Titles and degrees | | |
| Master of Arts (MA) | 37 | 6.2 |
| Master of Science in Engineering (MSc Eng.) | 17 | 2.8 |
| Doctor of Philosophy (PhD) | 265 | 44.6 |
| Doctor of Engineering (Dr. Eng.) | 145 | 24.4 |
| Associate Professor (Assoc. Prof.) | 109 | 18.3 |
| Professor (Prof.) | 15 | 2.5 |
| Other, e.g., Medical Doctor (MD) | 5 | 0.8 |
| Positions | | |
| Teaching staff | 41 | 6.9 |
| Research staff | 35 | 5.9 |
| Teaching and research staff | 500 | 84.3 |
| Other, e.g., research-technical | 17 | 2.8 |

*Note.* $N = 593$.

Apart from demographic data (age, gender, job seniority, title & degree), the respondents filled out the Polish version of the Language Learning Orientations Scale—intrinsic motivation, extrinsic motivation, amotivation (LLOS-IEA) (Noels et al., 2000). The Polish adaptation of the LLOS-IEA was conducted by a group of two translators independently, who performed the translation into Polish. The translators were experienced teachers of the English language; one worked mostly with adolescents and the second with adult learners. The two translations were compared with the original instrument and evaluated (in terms of wording, discrepancies, and variations) by a licensed psychologist and an academic teacher of foreign languages. Next, the instrument's translated version was blindly translated back into English by two other experienced English teachers and compared with the original tool. Two consecutive studies were conducted to analyze the internal consistency of the scale. Study 1 ($N = 117$) and study 2 ($N = 81$) showed good internal consistency varying between 0.57 and 0.84. Minor wording changes were introduced into the scale after study 1 and retested. After study 2, the final version of LLOS-IEA was established.

**Table 3**

*The Cronbach Alpha Index of Internal Consistency for LLOS-IEA Studies 1 and 2*

| Subscales | Items | The Cronbach alpha study 2 | Remarks | The Cronbach alpha study 1 |
|---|---|---|---|---|
| Amotivation | p1, p8, p15 | 0.833 | | 0.792 |
| External regulation | p2, p6, p16 | 0.233 | After removing p2, $\alpha$ = 0.837 | 0.570 |
| Introjected regulation | p3, p10, p17 | 0.629 | | 0.713 |
| Identified regulation | p4, p11, p18 | 0.843 | | 0.801 |
| Intrinsic motivation–knowledge | p5, p12, p19 | 0.778 | After removing p19, $\alpha$ = 0.840 | 0.827 |
| Intrinsic motivation–accomplish | p9, p13, p20 | 0.79 | | 0.842 |
| Intrinsic motivation–stimulation | p7, p14, p21 | 0.667 | | 0.796 |

It is of note that these results were similar to other studies that adapted LLOS-IEA (Ardasheva, Tong, & Tretter, 2012; Lou & Noels, 2018; Vandergrift, 2005).

The LLOS-IEA consists of seven subscales: amotivation, external regulation, introjected regulation, identified regulation, intrinsic motivation–knowledge, intrinsic motivation–accomplishment, and intrinsic motivation–stimulation, which are arranged on a continuum, that is, less self-determined regulation are inversely related to those more-self-determined regulations. Each subscale consists of three items rated on a 7-point scale. The maximum points on each subscale were 21.

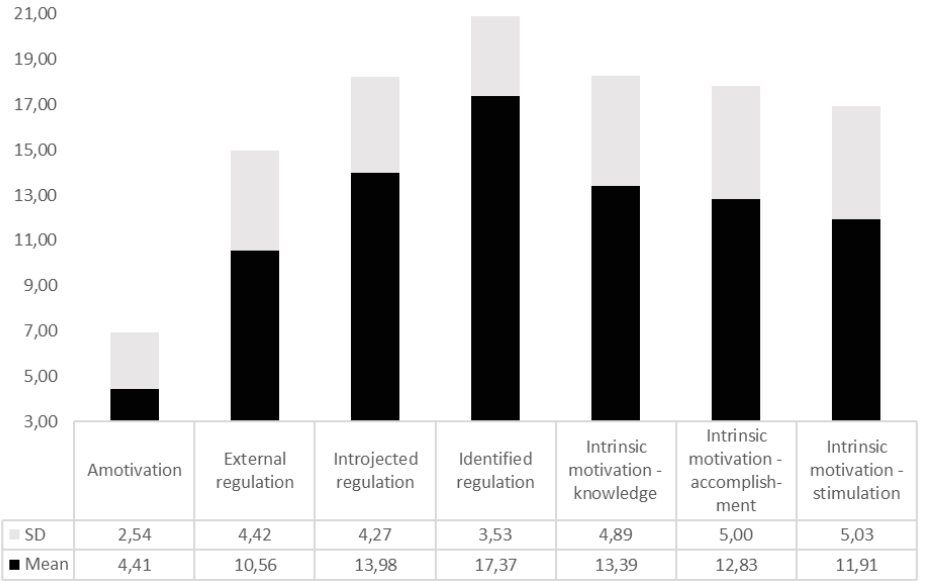
## Results

The main aim of the study was to determine what type of motivation university teachers had to learn foreign languages. The means scores obtained through statistical analysis from the LLOS-IEA (Noels et al., 2000) showed the highest means for identified regulation ($M$ = 17.34; $SD$ = 3.52), followed by similar levels of Introjected Regulation ($M$ = 13.98; $SD$ = 4.27) and intrinsic motivation–knowledge ($M$ = 13.38; $SD$ = 4.88). Next, intrinsic motivation–accomplishment ($M$ =12.82, $SD$ = 5.00) and intrinsic motivation–stimulation ($M$ = 11.91;

*SD* = 5.03) followed by external regulation (*M* = 10.56; *SD* = 4.41). The lowest mean was observed for amotivation (*M* = 4.40; *SD* = 2.54). The post hoc test for multiple comparisons using Bonferroni correction showed statistically significant differences ($p < 0.05$) between all subscales of the LLOS-IEA, except for Introjected regulation and intrinsic motivation–knowledge. It suggests that the levels of introjected regulation and intrinsic motivation–knowledge are similar.

**Figure 1**

*Mean Scores for Each of the Subscales of the LLOS-IEA*




| | Amotivation | External regulation | Introjected regulation | Identified regulation | Intrinsic motivation - knowledge | Intrinsic motivation - accomplish-ment | Intrinsic motivation - stimulation |
|---|---|---|---|---|---|---|---|
| SD | 2,54 | 4,42 | 4,27 | 3,53 | 4,89 | 5,00 | 5,03 |
| Mean | 4,41 | 10,56 | 13,98 | 17,37 | 13,39 | 12,83 | 11,91 |

*Note.* *N* = 534, *SD* = standard deviation.

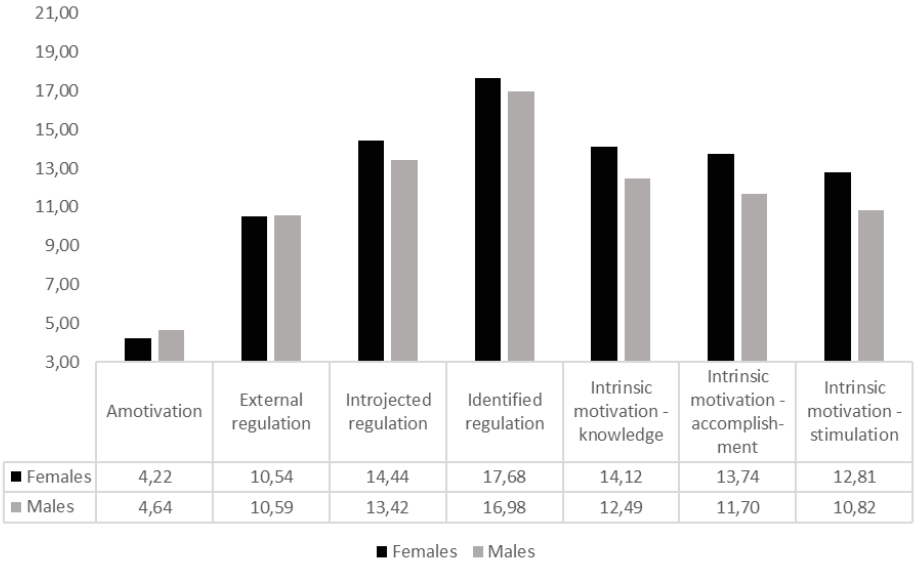
Another analysis found statistically significant correlations and differences between participants' motivation to learn foreign languages and their age, job seniority, and gender.

The analysis, using Pearson-product moment correlation found a statistically significant negative correlation between participants' age and external regulation ($r = –0.28$, $p < 0.01$) and identified regulation ($r = –0.098$, $p < 0.05$). These were weak correlations but statistically significant. There were no statistically significant correlations found between age and other types of regulations. Similarly to participants' age, a weak but statistically significant negative correlation was found between participants' job seniority and external regulation ($r = –0.27$, $p < 0.01$) and identified regulation ($r = –0.13$, $p < 0.05$). No other statistically significant correlations were observed between job seniority and other regulations ($p > 0.05$).

Another analysis using the t-test for differences between female and male academics and the motivation to learn foreign languages showed statistically significant higher motivation in the group of female scholars and introjected regulation ($t = 2.75$; $p < 0.01$), identified regulation ($t = 2.29$; $p < 0.05$), intrinsic motivation–knowledge ($t = 3.86$; $p < 0.001$), intrinsic motivation–accomplishment ($t = 4.77$; $p < 0.001$) and intrinsic motivation-stimulation ($t = 4.61$; $p < 0.001$) than in male academic group. The results can be seen in Figure 2.

**Figure 2**

*Mean Values for Female and Male Academics on the LLOS-IEA Subscales*



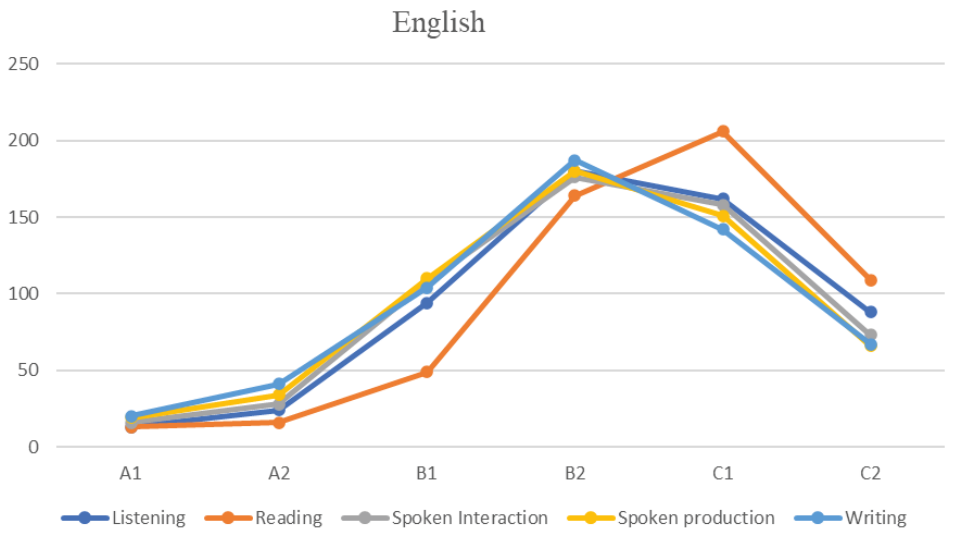

| | Amotivation | External regulation | Introjected regulation | Identified regulation | Intrinsic motivation - knowledge | Intrinsic motivation - accomplishment | Intrinsic motivation - stimulation |
|---|---|---|---|---|---|---|---|
| Females | 4,22 | 10,54 | 14,44 | 17,68 | 14,12 | 13,74 | 12,81 |
| Males | 4,64 | 10,59 | 13,42 | 16,98 | 12,49 | 11,70 | 10,82 |

*Note.* $N = 534$.

Due to space limitations, only data analysis for English will be presented. Almost all participants declared they knew English (94.7%). The overall level of proficiency in the English language ranged between B2 and C1, that is, independent, or proficient users, with reading skills at the highest level of C1 as indicated by participants (Figure 3).

**Figure 3**

*Self-rated Levels of Proficiency in Language Skills in English*




*Note.* $N = 562$.

Self-assessment of five language skills: writing, reading, listening, spoken interaction, and production, showed that academics' proficiency in the English language ranged between B2 and C1. Almost half of the sample was actively learning a foreign language at the time of the study. This study showed that many academics rated their skill of reading in English at the C1 level, known as effective operation proficiency. According to CEFR (2003), a learner who has attained the C1 level can understand a wide range of literary writings, journals, magazine articles, and specialized academic or professional publications. Reading in English at this level involves a detailed reading for information/argument, of lengthy, complex texts encountered in social, professional, or academic life, and the ability to identify finer points of detail, including attitudes and implied as well as stated opinions.

As far as the skill of writing in the English language is concerned, most scholars in this study rated it at the B2 level and fewer at the C1 level. In line with CEFR (2003), the B2 level in writing indicates that a learner can write specific pieces of writing (e.g., article, chapter, essay, letter) while passing on information or providing reasons in support of or against a particular point of view. Whereas the C1 level in writing suggests that a learner can form a well-structured, lengthy text about complex subjects, underlining salient issues with good expression and accuracy.

This study also showed that many scholars rated their speaking skills (production and interaction) at a B2 level, which indicates that academics can speak

about a wide range of subjects related to their field of interest, and expand and support ideas with subsidiary points and relevant examples. When delivering presentations to the audience, scholars can give a clear, systematically developed speech, highlighting significant points and relevant details and reacting to questions raised by the audience, showing fluency and ease of expression. During an interaction, at this level, a scholar can interact with the audience with a degree of fluency and spontaneity, which poses no strain for either the scholar or the audience (CEFR, 2018).

## Discussion

### Academics' Motivation to Learn Foreign Languages

The primary finding of this study was that scholars in this sample were both extrinsically motivated by identification (i.e., identified regulation) and introjection (i.e., introjected regulation), as well as intrinsically motivated (intrinsic motivation–knowledge). It is of note that people usually hold more than one reason (regulation) for engaging in specific behaviors at any one time (Howard, Gagne, & Morin, 2020), which means that types of regulation co-occur, and learning language can be energized by more than one type of motivation (Ryan, 2019).

Following the SDT continuum, some scholars displayed introjected orientation to learn languages, which implies that these individuals engage in behavior to earn self- and other-approval (or avoid disapproval). These scholars feel a personal obligation to learn the language because of self and/or normatively imposed expectations. These people do not feel fully volitional, and their behaviors are motivationally unstable and only weakly related to long-term commitment and performance.

Most academics in this study were, however, motivated to learn languages for identified reasons. These individuals personally endorse and truly identify with the value of learning languages. They recognize it as something personally important and worthwhile for themselves and their careers (Ryan & Deci, 2017). In other words, a person learns the language because they decided to do so and the activity has value for their chosen goals. As long as the goal is important, the learner persists in language learning (Noels et al., 2000). Research has shown that identified orientation is a relatively stable regulation and unlikely to change over time because it might be related to enduring values for a language learner (e.g., pursuing a personally important goal) (Noels et al., 2019).

The third type of motivation most frequently indicated by scholars in this study as the reason for learning foreign languages was intrinsic motivation–

knowledge. Scholars with this type of motivation engage in language learning because it has inherent appeal, interest, and satisfaction derived from learning, exploring new ideas, and developing knowledge (Noels et al., 2000; Ryan et al., 2019). Intrinsic motivation has been associated with higher quality learning and performance (Deci et al., 2017), greater intensity of effort (Busse & Walter, 2013), greater self- and linguistic confidence, and low anxiety (Lou & Noels, 2018).

The combination of extrinsic orientations (introjected and identified) and intrinsic motivation–knowledge suggests that academics learn languages for instrumental reasons on the one hand, which might be related to conducting research and publishing in foreign languages, teaching international students, or effectively communicating at conferences abroad. On the other hand, scholars' intrinsic motivation to learn languages might be related to developmental reasons such as exploring and acquiring knowledge through and about another language and deriving satisfaction from learning foreign languages. Extrinsic motivation is more common in the academic context than intrinsic motivation (Dresel & Hall, 2013). Extrinsic motivation might positively affect learning and achievement behavior, particularly when the learning activity or outcome is expected to be short-lived (Dresel & Hall, 2013). However, it is worth highlighting that intrinsic motivation is a strong determinant of learners' self-confidence and a reason for life-long learning of a second language (Pae, 2008).

## Motivation, Age, Gender, Job Seniority, and English

This study found that younger academics were more externally motivated than older scholars. The finding is perhaps not surprising as a foreign language for younger scholars is a vehicle to achieve other desired outcomes, for example, obtain external funding to conduct research, increase research productivity, and publish research findings (Lechuga & Lechuga, 2012). Therefore, for younger academics learning the language might be associated with external contingencies such as recognition, that is, research success (Stupnisky, BrckaLorenz, & Laird 2019). Conversely, older academics were less externally motivated to learn foreign languages. Research has shown that older learners tend to have a greater intrinsic goal orientation, and the value and worth of what is learned are more important for them than for their younger counterparts (Wlodkowski & Ginsberg, 2017).

As we age, our motivations to be recognized for our achievements, and to rack up more and more achievements, tend to decline, that is, older adults tend toward being more intrinsically than extrinsically motivated (Levitin, 2020, p. 174). Therefore, younger scholars tend to be more externally motivated, that is, achievement and recognition-oriented, than older scholars. In their profession, academics also progress through a series of career stages (e.g., junior, mid-

career, senior scholar), which collectively refer to the term seniority. Job seniority differentiated academics in terms of motivation to learn foreign languages.

Female scholars were found to differ from male scholars on three subscales of motivation, that is, introjected regulation, identification regulation, and all the intrinsic motivation subscales. The biggest differences between females and males were noted in intrinsic motivation–accomplishment and intrinsic motivation–stimulation. Female scholars were more intrinsically motivated than their male colleagues. These results suggest that for female academics, learning languages is associated with pleasant sensations related to the learning activity itself, satisfaction from improving performance, trying to reach new personal objectives, and accomplishing tasks (Noels et al., 2000).

Female scholars also scored higher than male academics on identified regulation—the most frequently indicated motivation for learning languages in this sample. This result implies that for females, more than for males, knowing the language may help attain important goals, such as improving their occupational performance (e.g., in research, teaching activities, or communication with other scholars) (Noels et al., 2019). In the study of the effects of motivation on the research success of professors, Stupnisky, BrckaLorenz, and Laird (2019) found differences between female and male professors in all three types of motivation (autonomous, introjected, external). Male scholars were found to be much more externally motivated than female scholars. However, the most considerable effect was noted among males who reported more perceived success than females.

The results indicated that scholars with lower seniority were more externally motivated to learn languages than their colleagues with longer job seniority. A parallel can be drawn to an age dimension, as these findings share a degree of similarity. Scholars who are at an early stage of their career are focused on obtaining tenure, and therefore, motivation is derived from extrinsic motivation to obtain tenure (Austin et al., 2007). At this stage, a foreign language might serve as a means to achieve this goal, and thus learning the language is extrinsically motivated. Some studies suggest that conducting research and publishing have become an important criterion for promotion, tenure, and career success (Lechuga & Lechuga, 2012). In turn, this situation poses more pressure on younger scholars in the early stages of their scientific careers. These academics face other pressures that might affect their motivation for learning languages. A study conducted by Tien (2008) showed that financial rewards play an important motivational role for younger scholars. However, a lack of financial support can function as a demotivating factor. For instance, one participant in this study commented that she "found no time for learning foreign languages because she focused her attention and time on finding additional income outside her university" (open comment). Recent research conducted among PhDs employed at Polish universities showed that most of these scholars (72% of women, and 79% of men) engaged in additional work outside their

institutions (Siemienska et al., 2019). Therefore, some younger academics may view a foreign language as a valuable instrument to achieve desired goals but adopt an extrinsic orientation towards learning it.

The seniority perspective can provide valuable information about motivation to learn languages because it can be seen as a fluent construct that changes over time. In the mid- and late career, academics are less concerned about tenure and more concerned about international cooperation with other scholars or disseminating research internationally (Kwiek, 2015). Bugaj (2016, p. 31) found that academics at a peak and late in their career focused on supervising research projects, providing expertise to the organization or other scholars, supporting their own and other research and development projects, and supporting younger colleagues. Thus, other factors can contribute to motivation for learning at this stage.

Not surprisingly, English is the foreign language most known among scholars (94.7%). Reading and writing are inherent (skills) to academic activity and perhaps the most important skills to affirm scientific and career progress. Few studies confirm this claim, for example, Macaro (2018) highlights that "academics experience fewer concerns about literacy skills as they have read large quantities of materials written in English as well as written papers in English themselves" (p. 83). In a study of Spanish university teachers, the majority of respondents said they felt equipped to read literature in English in their field (88.9%) (Fortanet-Gomez, 2012). Similar findings were reported from the study of Vietnamese university teachers, who felt confident about their literacy skills, and reading and writing were considered relatively unproblematic (Vu & Burns, 2014).

While reading is a receptive skill, speaking and writing are productive skills. Both skills have an important function in many academic and professional fields (oral presentations, written studies, reports), and particular social value is attached to them. Scholars are evaluated based on what they have submitted in writing, and how fluent they are in speaking, in particular when addressing an audience. What is more, fluency in formal production (writing and speaking) is not acquired naturally; instead, it is a literacy process learned through education and experience, mainly focused on conventions of the academic genre (CEFR, 2018). According to Macaro (2018), complex, academic discourse requires advanced levels of language, and spoken interactions with audiences also require a mastery of the language.

So far, some research has focused on scholars whose first language is not English. In the study of academic staff in UK business schools, Śliwa and Johansson (2014) found that academics who considered their communication skills inferior to the first language users also considered themselves as professionally less competent lecturers. Spanish senior academics also reported a particular sense of disadvantage in relation to spoken communication at

conferences (Perez-Llantada, Plo, & Ferguson, 2011). These authors found that academics' conference presentations tended to be unfairly assessed, not based on their scientific content, but on linguistic style (e.g., pronunciation, lack of use of humor in their presentations, and not being able to field as adroitly as first language users). Horn (2017) also showed that scholars with less than excellent English expression feel disadvantaged compared to first-language users. However, they acknowledge the importance of English as a dominant language in academia, but difficulties in disseminating knowledge through English evoke stigmatization and hardship. Scholars in other research worried that their oral proficiency was so low that it could negatively affect their students' English, or their comprehension of the lecture content (Vu & Burns, 2014). Besides, students who expected their teachers to speak like the first language users were not satisfied and negatively evaluated their teachers (Macaro, 2018; Śliwa & Johansson, 2014). In the Italian context, university teachers expressed particular concerns related to teaching through English, such as the inability to improvise in L2 easily and effectively as in their L1 and problems with the use of English in social and informal situations (Guarda & Helm, 2016).

Some research, however, shows that scholars have sufficient levels of English proficiency. A study of university teachers in Austria, Italy, and Poland (Dearden & Macaro, 2016) found that English was the language of articles, textbooks, and teaching materials for some scholars in Poland. One science teacher used scientific papers mostly written in English, and as a result, she found it "easier to talk in English than in Polish" (p. 467). In a comparative study of scholars in non-English-speaking European countries in the context of English-taught Programs, Lam and Maiworm (2014) reported that 97% of Polish scholars assessed their language proficiency as good or very good. Also, Danish, younger scholars declared having the necessary skills to teach in English and felt confident in delivering content in English (Jensen & Thogersen, 2011). However, some admitted that universities should not assume that all faculty could use English effectively in lecturing.

Most scholars in this study knew English and self-rated their proficiency as B2 or C1 according to the CEFR (2003). Similarly, in Airey's (2011) study, Swedish university lecturers rated their English skills as either B2 or C1. These levels seem satisfactory for scholars' regular academic activities; however, it seems that level C1 is thought to be the limit below which language training should be necessary.

## Conclusions

The motivation of academics to learn foreign languages is complex and not unitary. On the one hand, younger scholars (in terms of age and seniority) learn languages for instrumental reasons, perhaps to progress with their careers, increase research productivity, or cope with the internal or external pressures of being considered a professional and recognized locally and globally through international publications. On the other hand, older and female academics are more intrinsically motivated to learn languages, and the main reasons involve acquiring knowledge and developing through the means of that language. Scholars in their late careers might derive more pleasure from learning in general and be motivated by internal rewards such as satisfaction of curiosity, volition, and/or simply the joy of life-long learning, whereas their younger counterparts are extrinsically motivated by recognition, achievements, and successful progression through a series of career stages. Thus, English serves as a useful tool for attaining these. However, as Horn (2017) posits, scholars who wish to be internationally recognized must develop Anglophone fluency at an exceptionally high level.

In this study scholars' English proficiency ranged between B2 and C1 (only for reading skills) indicating independent or proficient users. At the time of the study, half of the sample was still actively involved in learning a foreign language. The reason for this is the use and spread of the English language as the "lingua franca" of science (Crystal, 2006) and a stronger focus on the internationalization of research, global academic exchange, and mobility. These factors have exerted on academics the need to engage in learning foreign languages as never before (Coleman, 2006).

Learning a foreign language is not an easy task, therefore, optimal motivation is a prerequisite to initiate, direct, and persevere in learning. In turn, scholars who are optimally motivated to learn foreign languages may enjoy participation in career and knowledge advancement, increased mobility, and thus higher employability. In addition, successful and motivated academics contribute to the prestige of the institution and demonstrate high-quality research and teaching skills based on the latest scientific knowledge (Macaro, 2018).

## Limitations and Recommendations for Future Research

This study was conducted in only one academic city. Extending research nationally perhaps would deliver a more complete and/or varied picture of scholars'

motivation. In future empirical inquiry, the inclusion of more qualitative and mixed methods in the study of university academics is recommended. Also, future studies could focus on academic mobility and foreign language development. In particular, research should attempt to determine whether there is a relationship between academic mobility and motivation to learn languages.

## References


Airey, J. (2011). Talking about teaching in English. Swedish university lecturers' experiences of changing their teaching language. *Iberica, 22*, 35–54.

Ardasheva, Y., Tong, S. S., & Tretter, T. R. (2012). Validating the English Language Learner Motivation Scale (ELLMS): Pre-college to measure language learning motivational orientations among young ELLs. *Learning and Individual Differences, 22*(4), 473–483. https://doi.org/10.1016/j.lindif.2012.03.001

Autstin, A. E, Sorcinelli, M. D., & McDaniels, M. (2007). Understanding new faculty: Background, aspirations, challenges, and growth. In R. P. Perry & J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 39–89). Springer.

Björkman, B. (2013). *English as an academic lingua franca*. De Gruyter Mouton.

Brckalorenz, A., Ribera, T., Kinzie, J., & Cole, E. R. (2012). 10: Examining effective faculty practice. *To Improve the Academy, 31*(1), 148–159. https://doi.org/10.1002/j.2334-4822.2012.tb00679.x

Bugaj, J. M. (2016). *Uniwersyteckie strategie rozwoju zawodowego nauczycieli akademickich w Polsce.* Wydawnictwo Uniwersytetu Jagiellońskiego.

Busse, V., & Walter, C. (2013). Foreign language learning motivation in higher education: A longitudinal study of motivational changes and their causes. *Modern Language Journal, 97*, 435–456. https://doi.org/10.1111/j.1540-4781.2013.12004.x

CEFR (Common European Framework of Reference for Languages: Learning, Teaching, Assessment). (2003). *Europejski system opisu kształcenia językowego: uczenie się, nauczanie, ocenianie*. Wydawnictwo Centralnego Ośrodka Doskonalenia Nauczycieli.

CEFR (Common European Framework of Reference for Languages: Learning, Teaching, Assessment). (2018). *Companion volume with new descriptors.* Council of Europe. https://www.coe.int/lang-cefr

Coleman, J. A. (2006). English-medium teaching in European higher education. *Language Teaching, 39*(1), 1–14. https://doi.org/10.1017/s026144480600320x

Crystal, D. (2006). English and the communication of science. *Moderna Språk 1*, 20–33.

Daumiller, M., & Dresel, M. (2020a). Researchers' achievement goals: Prevalence, structure, and associations with job burnout/engagement and professional learning. *Contemporary Educational Psychology, 61* (101843). https://doi.org/10.1016/j.cedpsych.2020.101843

Daumiller, M., & Dresel, M. (2020b). Teaching and research: Specificity and congruence of university faculty achievement goals. *International Journal of Educational Research, 99* (101460). https://doi.org/10.1016/j.ijer.2019.08.002

Daumiller, M., Rinas, R., Olden, D., & Dresel, M. (2020). Academics' motivations in professional training courses: Effects of learning engagement and learning gains. *International Journal for Academic Development, 25,* 1–17. https://doi.org/10.1080/1360144X.2020.1768396

Daumiller, M., Stupnisky, R., & Janke, S. (2020). Motivation of higher education faculty: Theoretical approaches, empirical evidence, and future directions. *International Journal of Educational Research, 99*, (101502). https://doi.org/10.1016/j.ijer.2019.101502

Deci, E. L., Olafsen, A. H., & Ryan, R. M. (2017). Self-determination theory in work organizations: The state of a science. *Annual Review of Organizational Psychology and Organizational Behavior, 4*(1), 19–43. https://doi.org/10.1146/annurev-orgpsych-032516-113108

Dearden, J., & Macaro, E. (2016). Higher education teacher's attitudes towards English as the Medium of Instruction. A three country comparison. *Studies in Second Language Learning and Teaching, 6(3)*, 455–486. https://doi.org/10.14746/sllt.2016.6.3.5

Dresel, M., & Hall, N. C. (2013). Motivation. In T. Goetz & N. C. Hall, (2013). *Emotion, motivation, and self-regulation a handbook for teachers* (pp. 57–113). Emerald Group Publishing Limited.

Fortanet-Gomez, I. (2012). Academics; beliefs about language use and proficiency in Spanish multilingual higher education. *AILA Review, 25*, 48–63. https://doi.org/10.1075/aila.25.04for

Guarda, M., & Helm, F. (2016). "I have discovered new teaching pathways": The link between language shift and teaching practice. *International Journal of Bilingual Education and Bilingualism, 20*(7), 897–913. https://doi.org/10.1080/13670050.2015.1125848

Hardré, P. L. (2012). Community college faculty motivation for basic research, teaching research, and professional development. *Community College Journal of Research and Practice, 36*(8), 539–561.

Horn, S. A. (2017). Non-English nativeness as stigma in academic settings. *Academy of Management Learning and Education, 16*(4), 579–602. https://doi.org/10.5465/amle.2015.0194

Howard, J. L., Gagné, M., & Morin, A. J. (2020). Putting the pieces together: Reviewing the structural conceptualization of motivation within SDT. *Motivation and Emotion, 44*(6), 846–861. https://doi.org/10.1007/s11031-020-09838-2

Javitz, H., Grimes, T., Hill, D., Rapoport, A., Bell, R., Fecso, R., & Lehming, R. (2010). *U.S. academic scientific publishing. Working paper SRS 11–201.* Arlington, VA: National Science Foundation, Division of Science Resources Statistics.

Jensen, C., & Thogersen, J. (2011). Danish university lecturers' attitudes towards English as the Medium of Instruction. *Ibérica, 22*, 13–34.

Kwiek, M. (2015). *Uniwersytet w dobie przemian*. PWN.

Lam, Q. K. H., & Mainworm, F. (2014). English in the classroom and beyond. In B. Wächter & F. Maiworm (Eds.), *English-taught programmes in European higher education the state of play in 2014* (pp. 98–118). Lemmens.

Landry, R., Lamari, M., & Amara, N. (2003). The extent and determinants of the utilization of university research in government agencies. *Public Administration Review, 63*(2), 192–205. https://doi.org/10.1111/1540-6210.00279

Lechuga, V. M., & Lechuga, D. C. (2012). Faculty motivation and scholarly work: Self-determination and self-regulation perspectives. *The Journal of Professoriate, 6*(2), 60–97.

Levitin, D. (2020). *The changing mind: A neuroscientist's guide to aging well*. Penguin Life.

López-Navarro, I., Moreno, A. I., Quintanilla, M. Á., & Rey-Rocha, J. (2015). Why do I publish research articles in English instead of my own language? Differences in Spanish researchers' motivations across scientific domains. *Scientometrics, 103*(3), 939–976. https://doi.org/10.1007/s11192-015-1570-1

Lou, N. M., & Noels, K. A. (2018). Western and heritage cultural internalizations predict EFL students' language motivation and confidence. *International Journal of Bilingual Education and Bilingualism,* 1–15. https://doi.org/10.1080/13670050.2018.1508277

Macaro, E. (2018). *English Medium Instruction. Content and language in policy and practice,* Oxford University Press.

Niemiec, C. P., Ryan, R. M. (2009). Autonomy, competence, and relatedness in the classroom: Applying self-determination theory to educational practice. *Theory and Research in Education,* 7, 133–144. https://doi.org/10.1177/1477878509104318

Noels, K. A., Clément, R., & Pelletier, L. (2001). Intrinsic, extrinsic, and integrative orientations of French Canadian learners of English. *Canadian Modern Language Review, 57*, 424–442. https://doi.org/10.3138/cmlr.57.3.424

Noels, K. A., Lou, N. M., Vargas Lascano, D. I., Chaffee, K. E., Dincer, A., Zhang, Y. S. D., & Zhang, X. (2019). Self-determination and motivated engagement in language learning. In M. Lamb, K. Csizer, A. Henry & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 95–115). Palgrave Macmillan. https://doi.org/10.1007/978-3-030-28380-3

Noels, K. A., Pelletier, L., Clément, R., & Vallerand, R. J. (2000). Why are you learning a second language? Orientations and Self-Determination Theory. *Language Learning, 50*, 57–85. http://dx.doi.org/10.1111/0023-8333.00111

Noels, K. A., Vargas Lascano, D. I., & Saumure, K. (2019). The development of self-determination across the language course: Trajectories of motivational change and the dynamic interplay of psychological needs, orientations, and engagement. *Studies in Second Language Acquisition*. https://doi.org/10.1017/S0272263118000189

Pae, T. I. (2008). Second Language Orientation and Self-Determination Theory. *Journal of Language and Social Psychology 27*(1), 5–27. https://doi.org/10.1177/0261927X07309509

Pérez-Llantada, C., Plo, R., & Ferguson, G. R. (2011). "You don't say what you know, only what you can": The perceptions and practices of senior Spanish academics regarding research dissemination in English. *English for Specific Purposes, 30*(1), 18–30. https://doi.org/10.1016/j.esp.2010.05.001

Perkmann, M., Tartari, V., McKelvey, M., Autio, E., Brostrom, A., D'Este, P., & Krabel, S. (2013). Academic engagement and commercialisation: A review of the literature on university–industry relations. *Research Policy, 42*, 423–442. https://doi.org/10.2139/ssrn.2088253

Ryan, R. M. (Ed.). (2019). *The Oxford handbook of human motivation* (2nd ed.). Oxford University Press.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology, 25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Publishing.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology, 61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Ryan, R. M., Ryan, W. S., Di Domenico, S. I., & Deci, E. L. (2019). The nature and the conditions of human autonomy and flourishing: Self-determination theory and basic psychological needs. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (2nd ed., pp. 89–110). Oxford University Press.

Stockemer, D., & Wigginton, M. J. (2019). Publishing in English or another language: An inclusive study of scholar's language publication preferences in the natural, social, and interdisciplinary sciences. *Scientometrics, 118*(2), 645–652. https://doi.org/10.1007/s11192-018-2987-0

Stupnisky, R. H., BrckaLorenz, A., & Laird, T. F. (2019). How does faculty research motivation type relate to success? A test of self-determination theory. *International Journal of Educational Research, 98*, 25–35. https://doi.org/10.1016/j.ijer.2019.08.007

Stupnisky, R. H., BrckaLorenz, A., Yuhas, B., & Guay, F. (2018). Faculty members' motivation for teaching and best practices: Testing a model based on self-determination theory across institution types. *Contemporary Educational Psychology, 53*, 15–26. https://doi.org/10.1016/j.cedpsych.2018.01.004

Stupnisky, R. H., Hall, N. C., Daniels, L. M., & Mensah, E. (2017). Testing a model of pretenure faculty members' teaching and research success: Motivation as a mediator of balance, expectations, and collegiality. *The Journal of Higher Education, 88*(3), 376–400. https://doi.org/10.1080/00221546.2016.1272317

Śliwa, M., & Johansson, M. (2014). How non-native English-speaking staff are evaluated in linguistically diverse organizations: A sociolinguistic perspective. *Journal of International Business Studies, 45*(9), 1133–1151. https://doi.org/10.1057/jibs.2014.21

Tien, F. F. (2008). What kind of faculty are motivated to perform research by the desire for promotion? *Higher Education, 55*(1), 17–32. https://doi.org/10.1007/s10734-006-9033-5

Vallerand, R. J., & Blssonnette, R. (1992). Intrinsic, extrinsic, and amotivational styles as predictors of behavior: A prospective study. *Journal of Personality, 60*(3), 599–620. https://doi.org/10.1111/j.1467-6494.1992.tb00922.x

Vandergrift, L. (2005). Relationships among motivation orientations, metacognitive awareness and proficiency in L2 listening. *Applied Linguistics, 26*(1), 70–89. https://doi.org/10.1093/applin/amh039

Vu, N. T. T., & Burns, A. (2014). English as a medium of instruction: Challenges for Vietnamese tertiary lecturers. *The Journal of Asia TEFL, 11*(3), 1–31.

Watt, H. M., & Richardson, P. W. (2020). Motivation of higher education faculty: (How) it matters! *International Journal of Educational Research, 100*(101533). https://doi.org/10.1016/j.ijer.2020.101533

Weinberg, B. A., Ownen-Smith, J., Rosen, R. J., Schwarz, L., Allen, B. M., Weiss, R. E., et al. (2014). Science funding and short-term economic activity. *Science, 344*(6179), 41–43. https://doi.org/10.1126/science.1250055

Willetts, D. (2019). *A university education*. Oxford University Press.

Wlodkowski, R., J., & Ginsber, M. B. (2017). *Enhancing adult motivation to learn* (4th ed.). Jossey-Bass.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/26
https://doi.org/10.31261/TAPSLA.16404

CC BY SA

**Adelina Castelo**
https://orcid.org/0000-0002-9079-1066
Universidade Aberta
Centro de Linguística da Universidade de Lisboa
Portugal

# Model for Using Music in Pronunciation Teaching (MOMUP): New Validation and Revision

Abstract

Songs have been used in the context of additional language teaching with different purposes, such as cultural knowledge, motivation, vocabulary development, and pronunciation teaching. A *Model for Using Music in Pronunciation Teaching* (MOMUP), proposed in 2018, provides guidelines to help teachers to create or adjust song-based learning materials to teach pronunciation in a more efficient way, but it needs further validation. Consequently, this paper aims at (i) presenting and discussing new validation data on the MOMUP, and (ii) revising the model in accordance with that discussion. The paper includes: a literature review on creation and validation of guiding models for teachers; the MOMUP's presentation; the description of new validation data on the model, collected through an online questionnaire to 30 teachers of additional languages; the model's revision in accordance with the validation data, as well as further literature review. The output of this process, the revised version of MOMUP, is simpler than the previous one and is composed of thirteen principles/criteria associated to each one of the topics *what for?*, *which song?*, and *how?* (3, 5 and 5 principles, respectively).

*Keywords*: songs, pronunciation, additional language teaching, guidelines for teachers, validation

The use of songs in the context of teaching an additional language (either a foreign language, a second language or a heritage language; hereafter referred to as L2) has since long been advocated for several reasons, namely its value for promoting learners' motivation, cultural knowledge, vocabulary, grammar, and pronunciation (e.g., Betti, 2012; Engh, 2013; Keskin, 2011; Ludke, 2009; Medina, 2002; Moncada Cevallos & Chancay Cedeño, 2023; Omolara, 2023; Santos Asensi, 1996). For instance, according to the learners' responses in questionnaires, their interest/motivation to learn the language increased thanks to the use of songs in the class (e.g., Anton, 1990; Yuhariyah & Syafryadin, 2023),

which might be considered a more gamified strategy. The songs are integrated into a cultural environment and their lyrics constitute an authentic text including cultural references, so the use of songs allows the learners to know more about the target culture (e.g., Engh, 2013; Keskin, 2011). Also, several studies argue that songs are a good resource to promote vocabulary growth and consolidation of grammar structures (e.g., Ludke, 2009; Medina, 2002). In fact, the songs' lyrics may include many words which are new to the learners and some studies show that music reinforces memorization of new vocabulary (e.g., Schön et al., 2008). In terms of grammar structures, both the fact that a certain structure is often repeated in the lyrics and the fact that the learners tend to listen to the same song several times contribute to promote the consolidation of grammar structures (e.g., Anton, 1990).

In terms of pronunciation, different authors consider that songs are particularly adequate for teaching it (e.g., Ashtian & Zafarghandi, 2015; Nobre-Oliveira, 2007; Saldıraner & Cintara, 2021; Zemlianska, 2021). In fact, songs include spoken (or better, sung) speech, thus allowing to practice both listening comprehension (improving this skill also seems to constitute a basis for progressing in pronunciation—e.g., Odisho, 2016) and pronunciation (e.g., Ludke, 2006; Omolara, 2023). As music, the songs foster a "repeated behaviour" (Upa et al., 2021, p. 115), of being listened to and sung several times, and thus an increased practice in hearing the correct pronunciation (which might function as a model) and producing the sounds and prosodic properties featured in them: "Through the songs, students learn rhythm, intonation, and pronunciation in a natural way as they listen to the music over and over and then attempt to reproduce the sounds they hear" (Anton, 1990, p. 1169). Even the rhythmic nature of songs can lead to an improvement in fluency and a natural flow of language: "[a song] can be highly beneficial because the verse's rhythm encourages learners to place emphasis where it belongs, fosters a natural flow of language, and increases fluency to help with pronunciation" (Yuhariyah & Syafryadin, 2023, p. 320). Also, some studies have empirically proved the positive impact of using songs in the L2 pronunciation (e.g., Ashtian & Zafarghandi, 2015; Upa et al., 2021).

Besides, even learners have been found to be aware of this benefit of songs. For example, Yuhariyah and Syafryadin (2023, p. 322) report that, in the questionnaire they administered to 26 students, all respondents either "strongly agree" or "agree" that "[their] pronounciation skills have improved [due] to YouTube songs videos" and "[u]sing YouTube song videos to learn English pronunciation is effective." Ananda (2023) shows that 97% of the 21 respondents in his questionnaire either "strongly agree" or "agree" that English songs help them to improve their listening comprehension and their pronunciation skills.

Also importantly, several studies mention the need to provide teachers with adequate guiding models or principles that allow them to prepare, change or improve their educational practices or teaching materials (e.g., Kanuka, 2002; Rahman et al., 2016).

However, to the best of our knowledge, so far there is only one guiding model for helping teachers to use songs in pronunciation teaching: the guiding model for the use of songs to teach a second language and emphasize the phonetic-phonological development proposed by Castelo (2018) and hereafter referred to as Model for Usings Music in Pronunciation Teaching or MOMUP. As the model was based only on a literature review, it should be validated. Besides, subsequent work with the model has shown that it can be improved, for instance, through simplification. Consequently, this paper has two goals: to present and discuss new validation data on the MOMUP; to propose its revision based on that discussion.

## Creation and Validation of Guiding Models for Teachers

Several authors recognize the need to develop and validate guiding models or principles or even complete frameworks that help teachers to prepare, change or improve their teaching practices or materials (e.g., Beckford, 1980; Kanuka, 2002; Rahman et al., 2016; Makina, 2020). Definitions and advantages of those guiding models or principles have been proposed in different studies, as some examples can show.

> [...] the use of a principled approach framed within a model for problem solving and change could be effective. Models have shown to be effective at facilitating change in that they can be used to clarify our thinking about a relatively complex phenomenon. Accordingly, using a model for change could provide guidance in the development of distance [or other modality] teaching and learning activities [...] (Kanuka, 2002, p. 164)

> Teaching and learning principles are statements on the scholarship of teaching and learning and a reference guide to good practice. These principles represent the shared view within an institution of the processes and conditions that contribute to a high-quality teaching and learning process [...]. [...] teaching and learning principles can be a guide to the maintenance and enhancement of teaching and learning standards. (Rahman et al., 2016, p. 127)

A review of some studies proposing guiding models or principles for teachers (with different goals) reveals that several sources can be used as a basis for those models or principles: only literature review (Beckford, 1980); semi-structured interviews with people involved in the problem, literature review, researcher's observations and experiences of the problem (recorded in a reflective journal), and scholarly critical analysis (Kanuka, 2002); literature review,

examples of guiding principles from selected institutions, and input from ten key informants (Rahman et al., 2016); literature review, followed by personal reflection and categorization of the data gathered in the literature (Makina, 2020).

The validation methods are also several but often include a survey by questionnaire and/or an expert evaluation or discussion. In those questionnaires, Likert-type opinion scales are frequently used to identify the participants' level of agreement with a statement, or their opinion on the importance and clarity of the proposed guiding principles. For instance, Beckford (1980) prepared a list of guiding principles that were evaluated by two juries in terms of their validity, importance, and clarity (firstly, one jury of university experts, secondly, another jury composed of teacher educators). Kanuka (2002) used both a focus group discussion by well-informed participants and a consensus survey questioning experts in the field (being each principle validated only when a minimal level of acceptance by the respondents in the survey was reached). Rahman et al. (2016) also made use of a questionnaire survey. Makina (2020) prepared a first validation of her principles through a critical analysis (with discussion at group conversations or conferences) and recommends a piloting implementation process.

## The MOMUP Guiding Model

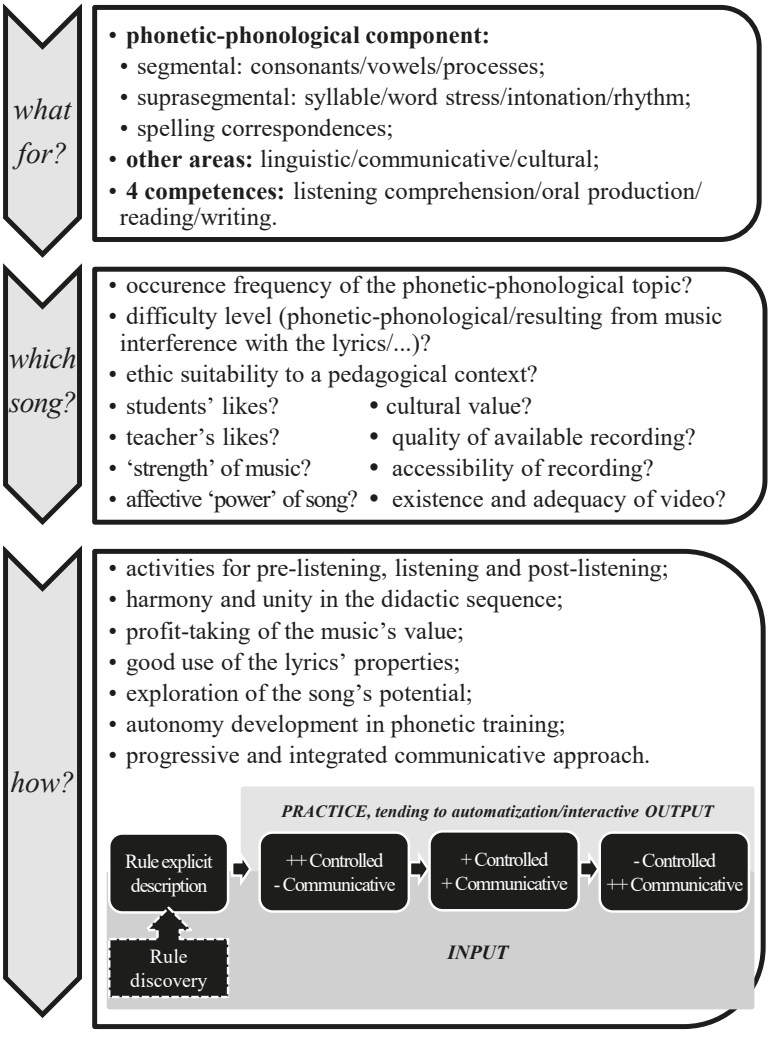
As mentioned before, Castelo (2018) proposed a guiding model for the use of songs to teach an L2 and emphasize the phonetic-phonological development, as the promotion of this language component is crucial to teach pronunciation.[1] This model, whose name is abbreviated in this paper as MOMUP (Model for Usings Music in Pronunciation Teaching) and which is presented in Figure 1, aims at helping teachers to prepare activities and/or materials to teach pronunciation through songs in a more efficient way. Its development was based in a systematization of principles and ideas gathered in a literature review of (i) research on the teaching of pronunciation and L2 in general (especially Celce-Murcia et al., 2010; Derwin & Munro, 2014; Ellis, 2005; Grant, 2014;

[1] As mentioned in Castelo (2018), the L2 pronunciation includes, at least, the language features which are segmental (i.e., sounds' properties and rules) and suprasegmental or prosodic (i.e., properties related to units larger than the speech sound, such as syllables, intonation, or stress) (although authors like Grant, 2014, mention four types of features in pronunciation: segmental; suprasegmental; peripheral; global). Segmental and suprasegmental features are part of the phonology of a language, "the abstract system of signs and rules which is the basis for the use (both in production and perception) of phonetic realities (physical, concrete sound elements) with a linguistic meaning. Consequently, in order to teach pronunciation, we have to promote the development of the whole phonetic-phonological domain or language component" (note 3, p. 8).

Hashemian & Fadaei, 2011; Hişmanoğlu, 2006; Kruk & Pawlak, 2014; Wrembel, 2011) and of (ii) many works advocating the employment of songs in an L2 teaching (especially Ashtian & Zafarghandi, 2015; Betti, 2012; Coelho de Souza, 2014; Engh, 2013; Keskin, 2011; Ludke, 2009; Medina, 2002; Montaner, 2006; Santos Asensi, 1996; Schön et al., 2008; Simpson, 2015).

**Figure 1**

*A Guiding Model for the Use of Songs to Teach a Second Language and Emphasize the Phonetic-Phonological Development (Castelo, 2018, p. 18)*




*Note*. The model is referred to as MOMUP (*Model for Usings Music in Pronunciation Teaching*) in this paper.

Summarizing MOMUP, it consists of a set of criteria/ideas which can guide teachers when they prepare activities or didactic materials to teach pronunciation through songs (for a fully detailed presentation, see Castelo, 2018). In a way, these criteria/ideas, sometimes presented in the form of questions, correspond to principles showing the properties these materials should have to respond to specific needs. The several criteria/ideas are organized into three main questions or topics:

> (i) 'what for?' (to identify the objects of study to address and the specific objectives to be achieved); (ii) 'which song?' (to select the basis of their teaching sequence according to the most relevant criteria); (iii) 'how?' (to consider the principles associated with a choice of the most appropriate strategies). (Castelo, 2018, p. 19)

When deciding the goals of the activities or materials (topic *what for?*), the teacher should choose not only the phonetic-phonological topic to address, but also which other areas to explore and which basic skills to develop. As shown in the Figure 1, the phonetic-phonological topic to address might be segmental (i.e., related either to specific sounds as consonants and vowels or to phonological processes such as vowel reduction in unstressed syllable), suprasegmental (when it consists of some phonetic-phonological aspect associated with units larger than the segment, such as the syllable, the word or the intonational sentence) or related to the spelling correspondences (i.e., what are the connections between specific sound(s) and their orthographic representation(s)). All these aspects are linked to pronunciation (a part of the linguistic competence which also greatly influences the communication success) and can be pedagogically explored through a song. However, the song may also be the basis for promoting other linguistic areas (such as morphosyntax or lexicon), as well as communicative competence and cultural knowledge. Finally, the goals of the didactic sequence might deal with different language skills: some sequences might specially aim at promoting the listening comprehension or the speaking, while others focus more on developing the reading and writing abilities. So, the goals defined for a didactic sequence based on a specific song might be, for example, to develop the learners' intonation (phonetic-phonological topic), their linguistic proficiency in using the passive sentence and their cultural knowledge on a certain tradition (linguistic and cultural areas), and their speaking and writing skills (competences).

Concerning the song's choice (topic *which song?*), the teacher should take into consideration several criteria/principles. They should consider the occurrence frequency of the phonetic-phonological topic to address, to ensure that the chosen song presents many instances of the target pronunciation element (to be listened to and reproduced by the learners). Also, the song's difficulty

level must be adequate to the learners: for example, it should not present (i) too difficult cases of coarticulation or connected speech that might hinder the listening comprehension, (ii) a speaking speed that is too fast for the learners' language level, (iii) cases where the song's melody or rhythm, in a way, overlaps with its lyrics and makes pronunciation less clear and more demanding in terms of understanding (i.e., a kind of *music interference with the lyrics*), or (iv) lexicon, grammar or pragmatic elements deemed too demanding for the learners' language proficiency level. In addition, it is important that the song is ethically suitable to a pedagogical context, avoiding, for example, the proposal of intolerant attitudes or topics that might hurt students for some reason.

Since motivation and personal interests are very important in language learning, it is also advisable (i) that the song matches both the students' and teacher's likes, as much as possible, (ii) that it easily captures attention, featuring the potential to cause the "songstuck-in-my-head" phenomenon (Ashtian & Zafarghandi, 2015) (something that is designated in MOMUP as *'strength' of music*), and (iii) that it is able to trigger an emotional or affective response from the learners (named in MOMUP as *affective 'power' of song*). Another criterion to take into account is the song's cultural value: even when the didactic goals do not include the cultural area, using a song recognised as relevant in the target culture (for political, social, historical or artistic reasons) adds an extra value to the teaching sequence. Finally, whenever possible (after considering the principles previously mentioned), the teacher should also choose a song with a good quality recording that is accessible and a suitable videoclip. Although one of the main goals of this kind of song-based didactic sequences is to promote phonetic-phonological training, it is recommendable as well that the song has a videoclip: a visual component not only reinforces the song's impact and can improve the learners' motivation, but also can be useful for pronunciation activities like mirroring, where the learners imitate speech and actions of a certain character or, in this case, a specific singer.

The third main topic of the model deals with the principles about how to use the song (topic *how?*). As recommended by several authors (e.g., Keskin, 2011; Montaner, 2006), the song's didactic exploration should include activities for pre-listening, listening and post-listening. The didactic sequence should also exhibit harmony and unity among the different tasks: this unity can be seen in the fact that a specific theme is present in every activity (e.g., romantic love, daily life) or that the progression from one activity in the sequence to another is natural, expected, and meaningful.

In addition, some principles to use songs in language teaching in a relevant way can be derived from the fact that a song is a combination of music and lyrics (e.g., Coelho de Souza, 2014). As a result, it is important to take the best of these three elements: music (the melody and rhythm), lyrics (the text), and song (the combination of music and lyrics). For instance, the music promotes

the learners' aesthetic sensitivity and musical intelligence. Consequently, although these abilities are not the major goals in L2 learning, a song-based didactic sequence gains an extra value if it explores these potential benefits, for example, by helping the students to be aware of the musical instruments being played and the properties that characterize a certain musical style. The lyrics consist of a short authentic text, frequently displaying literary features and cultural references, as well as a certain vocabulary and grammar structures. So, it is pertinent that the didactic sequence tries to explore as much as possible all this potential associated with the lyrics. Specific way of interaction between music and lyrics also unlocks a great potential: for instance, the manner music and lyrics are interconnected conveys meanings and the teacher can bring the students to discover those meanings.

Two final principles are proposed in MOMUP to describe how to use a song as a basis for pronunciation training: (i) *autonomy development in phonetic training* and (ii) *progressive and integrated communicative approach* [in phonetic training]. As advocated by several authors (e.g., Kruk & Pawlak, 2014), it is crucial to develop learners' autonomy in phonetic training, since the time spent in class is not enough for all the necessary pronunciation practice at an individual level. So, if the song explored in class is available for listening and rehearsal outside of class, the students are being given the opportunity to autonomously continue to practice their pronunciation. As far as the progressive and integrated communicative approach is concerned, it is noteworthy that this approach starts with the description of the pronunciation rule (and possibly its discovery by the learners), continues with tasks of more controlled and less communicative output (allowing the learners to have intensive pronunciation practice with smaller units), and leads to more communicative tasks presenting more spontaneous speech. During this process, the learners receive large amounts of input and are given many opportunities to produce output that is progressively more spontaneous and communicative.

MOMUP was not validated before being published, but Castelo et al. (2022) applied it to the analysis of three didactic sequences using songs to teach pronunciation. These sequences had been previously validated with the target learners or teachers. So, if their analysis under MOMUP's light reveals there is a match between the model's principles and the didactic sequences' properties, then it shows that the model's guidelines are relevant and functional. The main results of this analysis indicate that the model is generally adequate and useful, since almost all criteria/principles proposed in MOMUP are met in the didactic sequences as well. However, the analysis also shows that some small changes might simplify and improve the model. More specifically, for the topic *which song?*, the criteria *'strength' of music* and *affective 'power' of song* show a high degree of subjectivity, while the criteria *students' likes* and *teacher's likes* depend on the people involved. So, as each of these four criteria is not so

informative in the process of creating activities or materials to teach pronunciation through songs, the authors propose a simplification: its replacement by a unique criterion like *song's appeal to the listeners*. For the topic *how?*, the authors also advocate a simplification: keeping the criterion *good use of the lyrics' properties* (as the lyrics' textual and linguistic nature makes it extremely important to develop competence in a language) and including aspects related to the music's exploration (melody, instruments, etc.) into the criterion *exploration of the song's potential* (interaction of music and lyrics). Like this, three criteria would be reduced to two.

## New Validation Data for MOMUP

In this section, the new data for the validation of MOMUP are presented, namely the method used for its collection and its main results obtained.

### Method

The data were obtained via a questionnaire addressed to teachers of L2. This is a simple method of having access to the opinion and evaluation of a larger number of in-service professionals and thus gaining new insights into the validity of the model under analysis thanks to different professional experiences and views. As seen before, questionnaire surveys constitute a very frequently used method for validating guiding principles, models, or frameworks. The questionnaire under analysis included mainly closed questions (since the responses to them are easier to give, analyse and compare) as well as some open questions (allowing the teachers to explain better their opinions or share suggestions and comments).

This validation instrument (which constitutes Appendix 1) was composed of two parts. In part 1, eight questions allowed us to characterize the participants in terms of work experience as well as general beliefs and practices important to understand how experienced these participants were in creating materials, using songs, and addressing pronunciation. Part 2 results from a conversion of the three main topics of MOMUP into three complex questions. Each question asks to classify each criterion/idea of a model's topic in terms of its clarity and importance/usefulness (being the topics *what for?*/goals, *which song?*/song's choice, and *how?*/didactic sequence). For each criterion/idea the respondents had to choose an option for clarity (*totally clear*, *partially clear*, *not so clear*, *or unclear*) and another for importance/usefulness (*totally impor-*

*tant/useful*, *partially important/useful*, *not so important/useful*, *or unimportant/useless*). It means that a Likert-scale was used (as it occurs often in this type of validation method) and it included 4 points to avoid having an intermediate level, which might tend to be chosen by default or in cases of greater hesitation. After each question, the participants had the option to add comments or suggestions related to that topic.

The questionnaire was answered anonymously and online via a form in the Qualtrics platform during 2022, by participants who volunteered to do so, at their best convenient time and place, after receiving an email invitation to participate in the study either from the author or another colleague. The email invitations were sent to many teachers of additional languages who were acquaintances of the author and who were asked to further disseminate the questionnaire among their own colleagues. In the end, only thirty teachers of L2 answered it. As the only inclusion criterion for participants was to be a teacher of an additional language, we got answers from teachers who were diverse in terms of taught languages, teaching levels, and work experience. Also, most probably many of the respondents did not know MOMUP before taking the questionnaire, which allowed us to get answers from 'naïve' respondents. The diversity of this convenience sample was intended to mirror the diversity of the MOMUP's possible users and to try to receive different ideas and insights about it.

## Results

The characterisation of the participants is visible in Tables 1 and 2.

**Table 1**

*Characterisation of the Participants (I): Languages and Contexts (Counts)*

| Teacher's L1 | Taught L2 | Type of L2 | Type of students |
|---|---|---|---|
| Portuguese (19) | Portuguese (23) | Foreign language (21) | College/university (13) |
| Portuguese/English (1) | English (4) | Second language (5) | Adults (5) |
| Portuguese/French (1) | French (1) | Several types (4) | College + adults (10) |
| Portuguese/Spanish (1) | German (1) | | High school (1) |
| Chinese (3) | Spanish (1) | | Middle school (1) |
| Hungarian (2) | | | |
| German (1) | | | |
| Russian (1) | | | |
| Spanish/Catalan (1) | | | |

The L1 of most teachers was Portuguese (19 monolinguals and 3 bilinguals with Portuguese), but there were also respondents whose native language was Chinese (3), Hungarian (2) or another (1 German, 1 Russian, 1 Spanish/Catalan). The L2 they taught was mainly Portuguese (23), although there were also teach-

ers of English (4) and other languages (French, German and Spanish). Their teaching context corresponded primarily to L2 as a foreign language (21 cases) and at the college/university level or with adults in other contexts (13 teachers at college, 5 teachers of adults, 10 teachers of both college students and adults).

**Table 2**

*Characterisation of the Participants (2): Teaching Experience and Probability of Using MOMUP (Means, Medians, Modes)*

| | Teaching experience (number of years) | Number of materials created per month (scale 0 to > 6) | Number of songs used per month (scale 0 to > 6) | Importance assigned to pronunciation (scale 0–7) | Probability of using MOMUP (scale 0–21) |
|---|---|---|---|---|---|
| Mean | 14.3 | 4.3 | 1.8 | 4.8 | 10.9 |
| Median | 15 | 5 | 2 | 5 | 11 |
| Mode | 20 | > 6 | 1 | 5 | 8 |

In terms of experience, most teachers were experienced, as they had been teaching for more than a decade (mean 14.3 years; median 15 years; mode 20 years; only four reported having less than 5 years of experience), and created around 4–5 materials per month (mean 4.3; median 5; mode > 6). They also reported giving some importance to pronunciation in their teaching practice (around 5 on a scale with 7 as a maximum level of importance). However, they were not frequent users of songs, as they only included 1–2 songs per month in their teaching practice (mean 1.8; median 2; mode 1). Table 2 also shows the probability of using MOMUP (as a guiding model to create didactic materials based on songs to train pronunciation). This last measure, with a scale ranging from 0 to 21, corresponds to the sum of participants' answers for the number of materials created per month (0 to > 6, this latter option converted into 7), the number of songs used per month (0 to > 6, this latter option converted into 7), and the importance assigned to pronunciation in the L2 teaching (scale 0–7). The probability of using MOMUP allows us to distinguish the respondents whose answers should be taken into more consideration while analysing the results for the model's validation. As shown in the table, the probability of using MOMUP is located around the middle of the scale (mean 10.9; median 11; mode 8), and this result is mainly related to the fact that the teachers do not use songs in their teaching very often.

To analyse the results in part 2 (the questions on the clarity and importance/usefulness of each criterion/idea of the model), the following tables include three measures: mean (considering that the answer options were converted into 0 for "not clear/important," 1 for "not so clear/important," 2 for "partially clear/important" and 3 for "totally clear/important"); mode (with the same con-

versions); percentage of total agreement (percentage of answers "totally clear/important"). However, more importance will be given to the percentage of total agreement, as this measure shows better the level of satisfaction and agreement with the criterion's clarity and importance. Tables 3, 4 and 5 show the results obtained for the model's topics *what for?*, *which song?* and *how?*, respectively.

**Table 3**

*Clarity and Importance/Usefulness of Criteria/Ideas Related to "what for?" (Question 1 in Part 2)*

| | Clarity | | | Importance/Usefulness | | |
|---|---|---|---|---|---|---|
| | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] |
| Phonetic-phonological topic | 2.57 | 3 | 64 | 2.28 | 2 | 41 |
| Other areas | 2.71 | 3 | 71 | 2.44 | 2 | 44 |
| Four communicative competences | 2.68 | 3 | 75 | 2.48 | 3 | 56 |

As far as the topic *what for?* is concerned, the results reveal no problems in terms of clarity, since in all cases the mode is 3 (the maximum agreement) and the level of total agreement is around 60% and 70%. However, in terms of importance/usefulness, the participants consider the criteria *phonetic-phonological topic* and *other areas* less important/useful in the model: 41% and 44%, respectively, of total agreement. In the open question, a teacher also added that the phonological-phonetic topics are not important when the goal is to teach a domain other than oral production, while another one mentioned to have selected the songs according to the phonological-phonetic topics, the grammar (especially the verb tenses), and the vocabulary.

For the topic *which song?*, the results allow us to identify two problems of lack of clarity: in the criteria *teacher's likes* (only 24% of total agreement) and *'strength' of music* (40% of total agreement). In terms of importance, besides the problems in previous criteria (*teacher's likes* with 7% of total agreement and *'strength' of music* with 17%), several other were considered not so important: *students' likes* and *affective 'power' of the song* (both with 46% of total agreement), *cultural value* (43%) and *existence and adequacy of videoclip* (32%). Some teachers added comments on the criteria to choose a song: two mentioned the importance of having a videoclip (especially for students of lower proficiency levels); another two underlined the need for intelligibility (either of the singer's voice or of the input, which should be adjusted to the learners' proficiency level); one referred to the context (course) and the topic

to be addressed as important criteria to choose a song; one advocated that sounds are more important than content for the beginners.

**Table 4**

*Clarity and Importance/Usefulness of Criteria/Ideas Related to "which song?" (Question 2 in Part 2)*

| | Clarity | | | Importance/Usefulness | | |
|---|---|---|---|---|---|---|
| | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] |
| Occurrence frequency of the phonetic-phonological topic | 2.54 | 3 | 61 | 2.50 | 3 | 61 |
| Difficulty level | 2.79 | 3 | 82 | 2.71 | 3 | 71 |
| Ethic suitability to a pedagogical context | 2.69 | 3 | 79 | 2.64 | 3 | 75 |
| Students' likes | 2.43 | 3 | 57 | 2.43 | 2 | 46 |
| Teacher's likes | 1.60 | 2 | 24 | 1.27 | 1 | 7 |
| 'Strength' of music | 2.04 | 3 | 40 | 1.86 | 2 | 17 |
| Affective 'power' of the song | 2.59 | 3 | 63 | 2.43 | 2 | 46 |
| Cultural value | 2.56 | 3 | 67 | 2.36 | 2 | 43 |
| Quality of available recording | 2.33 | 3 | 67 | 2.27 | 3 | 57 |
| Existence and adequacy of videoclip | 2.15 | 3 | 50 | 1.82 | 3 | 32 |

Finally, for the topic *how?*, there is some disagreement with the criterion *exploration of the song's potential* (only 46% of total agreement on its clarity and 39% on its importance). The other criteria are considered both clear and important/useful by the respondents. One teacher commented that the pre-listening and post-listening activities depend on the learners' proficiency level, but are crucial for the beginners.

Table 6 presents a comparison between scores by all participants (mean in the Likert-scale of agreement with a statement) and by the more probable MOMUP's users (mean of agreement and percentage of total agreement).

**Table 5**

*Clarity and Importance/Usefulness of Criteria/Ideas Related to "how?" (Question 3 in Part 2)*

| | Clarity | | | Importance/Usefulness | | |
|---|---|---|---|---|---|---|
| | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] | Mean (scale 0–3) | Mode (scale 0–3) | Total Agreement [%] |
| Activities for pre-listening, listening, and post-listening | 2.86 | 3 | 86 | 2.82 | 3 | 82 |
| Harmony and unity in the didactic sequence | 2.25 | 3 | 50 | 2.36 | 3 | 50 |
| Good use of the lyrics' properties | 2.71 | 3 | 79 | 2.71 | 3 | 75 |
| Exploration of the song's potential | 2.18 | 3 | 46 | 2.14 | 3 | 39 |
| Autonomy development in pronunciation training | 2.61 | 3 | 71 | 2.75 | 3 | 75 |
| Progressive and integrated communicative approach | 2.50 | 3 | 61 | 2.48 | 3 | 62 |

**Table 6**

*Clarity and Importance/Usefulness of All Criteria/Ideas: Comparison between Scores by All Participants and More Probable MOMUP's Users*

| | | Clarity | | | Importance/Usefulness | | |
|---|---|---|---|---|---|---|---|
| | | All (mean) | More probable MOMUP's users (mean) | More probable users (Total Agreement) [%] | All (mean) | More probable MOMUP's users (mean) | More probable users (Totally Agreement) [%] |
| Q1 | Phonetic-phonological topic | 2.57 | 2.67 | 75 | 2.28 | 2.38 | 46 |
| | Other areas | 2.71 | 2.92 | 92 | 2.44 | 2.42 | 42 |
| | Four communicative competences | 2.68 | 2.75 | 83 | 2.48 | 2.50 | 67 |
| Q2 | Occurrence frequency of the phonetic-phonological topic | 2.54 | 2.67 | 75 | 2.50 | 2.62 | 69 |
| | Difficulty level | 2.79 | 2.92 | 92 | 2.71 | 2.62 | 62 |
| | Ethic suitability to a pedagogical context | 2.69 | 2.85 | 85 | 2.64 | 2.77 | 77 |

*Table 6 continued*

| | | Clarity | | | Importance/Usefulness | | |
|---|---|---|---|---|---|---|---|
| | | All (mean) | More probable MOMUP's users (mean) | More probable users (Total Agreement) [%] | All (mean) | More probable MOMUP's users (mean) | More probable users (Totally Agreement) [%] |
| | Students' likes | 2.43 | 2.58 | 75 | 2.43 | 2.46 | 54 |
| | Teacher's likes | 1.60 | 2.09 | 45 | 1.27 | 1.64 | 7 |
| | 'Strength' of music | 2.04 | 2.70 | 80 | 1.86 | 2.07 | 21 |
| | Affective 'power' of the song | 2.59 | 2.83 | 83 | 2.43 | 2.69 | 69 |
| | Cultural value | 2.56 | 2.92 | 92 | 2.36 | 2.46 | 46 |
| | Quality of available recording | 2.33 | 3.00 | 100 | 2.27 | 2.64 | 71 |
| | Existence and adequacy of videoclip | 2.15 | 2.64 | 73 | 1.82 | 1.92 | 38 |
| Q3 | Activities for pre-listening, listening, and post-listening | 2.86 | 2.83 | 83 | 2.82 | 2.77 | 77 |
| | Harmony and unity between in the didactic sequence | 2.25 | 2.25 | 67 | 2.36 | 2.38 | 54 |
| | Good use of the lyrics' properties | 2.71 | 2.67 | 83 | 2.71 | 2.62 | 69 |
| | Exploration of the song's potential | 2.18 | 2.00 | 42 | 2.14 | 2.23 | 38 |
| | Autonomy development in pronunciation training | 2.61 | 2.50 | 67 | 2.75 | 2.69 | 69 |
| | Progressive and integrated communicative approach | 2.50 | 2.67 | 75 | 2.48 | 2.69 | 69 |

Results show the same tendencies mentioned previously as far as the topics *what for?* and *how?* are concerned. However, for the topic *which song?*, the more probable users only show some disagreement with the criteria *teacher's likes* (only 45% of total agreement on its clarity and 7% on its importance), *'strength' of music*, *cultural value*, and *existence and adequacy of a videoclip* (only 21%, 46%, and 38% of total agreement on its importance, respectively). This means that the set of more probable MOMUP's users accept as clear and important/useful more criteria than the set of all respondents.

## Discussion of Validation Data and Model's Revision

The results of the validation in Castelo et al. (2022) showed that the MOMUP is generally adequate and useful, but it could be improved with small changes consisting mainly of simplification. The authors made two suggestions: (i) replacing the criteria *'strength' of music*, *affective 'power' of song*, *students' likes* and *teacher's likes* by a unique criterion like *song's appeal to the listeners*; (ii) including aspects related to the music's exploration (criterion *profit-taking of the music's value*) into the criterion *exploration of the song's potential.*

The new validation data presented in this paper reveal the same tendencies. Overall, the criteria included in the model are considered clear and useful, with some exceptions related to the following criteria: *phonetic-phonological topic* and *other areas* (topic *what for?*); *'strength' of music*, *teacher's likes*, *cultural value*, and *existence and adequacy of videoclip* (topic *which song?*); *exploration of the song's potential* (topic *how?*).

These results combined with further reflection and literature review support the revision of MOMUP, that is showed in Figure 2.

In the revised version of MOMUP, the principles/criteria are stated under the form of full declarative sentences (instead of only questions or nominal phrases) with an enumeration of criteria—e.g., *The didactic goals should include: 1. ...*; *2. ...*; *3. ...* This choice intends to make the meaning of the principles even more transparent, although they were normally considered very clear.

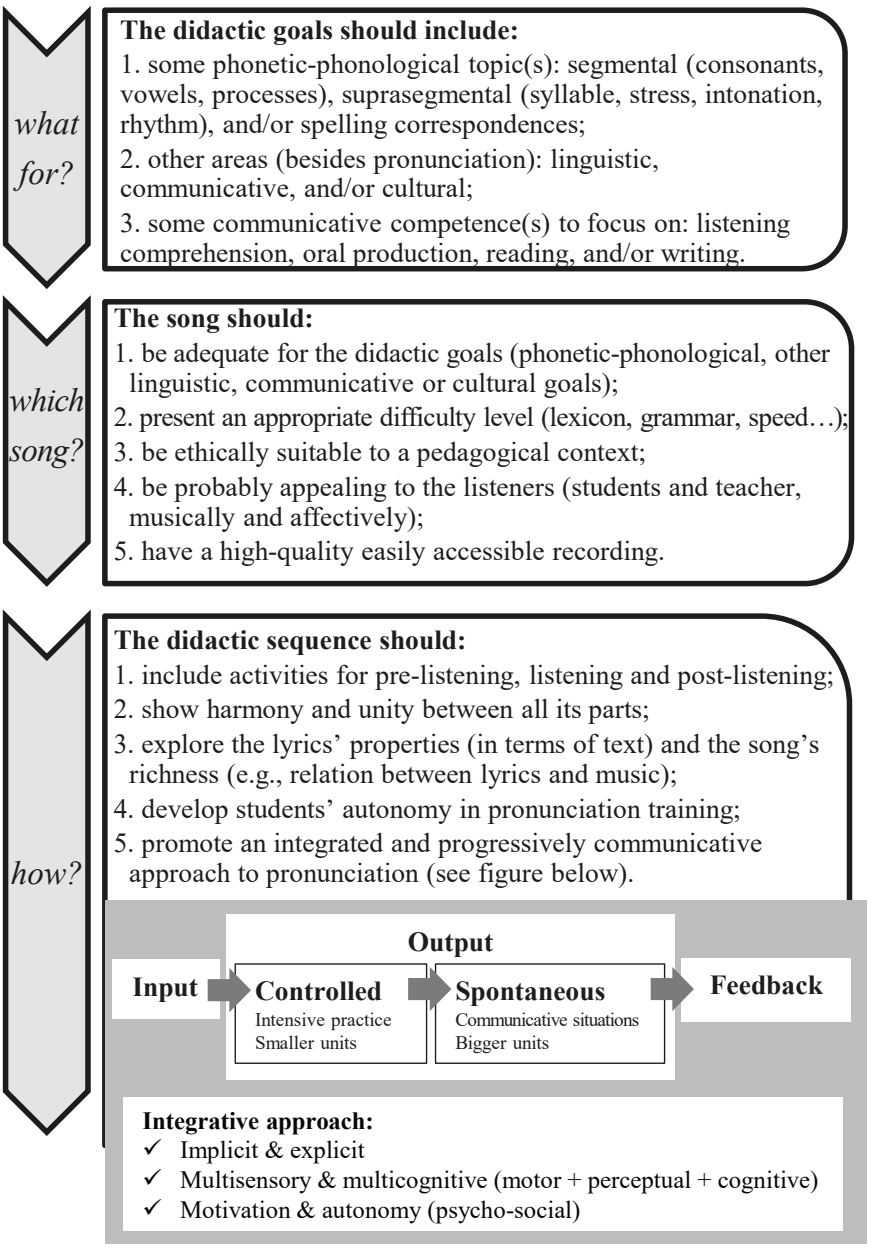
In terms of the topic *what for?*, the criteria of *phonetic-phonological topic* and *other areas* are kept in the model, despite the new validation data. Firstly, the inclusion of goals related to the phonetic-phonological component is essential in a model associated with training pronunciation through songs. Secondly, as mentioned by several authors, the songs are a good means to promote other areas of knowledge besides pronunciation and communication skills, such as grammar and culture knowledge (e.g., Degrave, 2019; Kumar et al., 2022; Ludke, 2009; Santos Asensi, 1995).

The topic *which song?* undergoes a considerable simplification. The several criteria identified in the before-mentioned validation processes as not so clear or so relevant are integrated into other principles, giving rise to a reduction from eleven to five criteria. So, the criterion of *cultural value* is included in the principle stating that *The song should be adequate for the didactic goals*. Several other criteria become a part of the principle *The song should be probably appealing to the listeners*, which means that as far as possible it should show a great potential for pleasing the students and the teacher, for having a music that captures attention and is easily remembered after a few listening opportunities and for promoting the affective involvement of the students. Finally, the criteria of *quality of available recording*, *accessibility of recording*

and *existence and adequacy of videoclip* are combined into a single principle: *The song should have a high-quality easily accessible recording.*

**Figure 2**

*The Revised MOMUP*



As for the topic *how?*, there was also some reduction. All components of lyrics, music and song (association of music and song) are often deemed as very important while exploring a song didactically (e.g., Delgrave, 2019; Kumar et al., 2022). Consequently, all of them are kept and merged into a single principle: *The didactic sequence should explore the lyrics' properties (in terms of text) and the song's richness (e.g., relation between lyrics and music).*

The scheme for the communicative approach to pronunciation, which is a part of the principle *The didactic sequence should promote an integrated and progressively communicative approach to pronunciation*, is also updated following more literature review. It adopts the synthesis presented in Castelo (2017, 2022) and enriched with Odisho's (2007, 2016) proposals. According to Castelo's synthesis (which was mainly based on the proposals by Celce-Murcia et al., 2010; Derwin & Munro, 2014; Ellis, 2005; Grant, 2014; Hashemian & Fadaei, 2011; Kruk & Pawlak, 2014), a communicative approach to pronunciation should start with much input, continue with controlled output (production of smaller units, corresponding to an intensive practice of the target sound structures) and move progressively to spontaneous output (free expression of sentences or texts in communicative situations, with no time for controlling the pronunciation). The last step is feedback, which is essential to guarantee that the learners are aware of their strong and weak points and focuses their future attention in specific aspects of pronunciation. All this process should follow an integrative approach that combines implicit and explicit focus on pronunciation, develops multisensory and multicognitive knowledge and fosters both motivation and autonomy. In fact, both implicit and explicit ways of teaching pronunciation seem helpful to reach a better pronunciation (e.g., Hashemian & Fadaei, 2011). Besides, Odisho (2007, 2016), based on his long experience as a pronunciation teacher of L2 English, advocates the use of a multisensory and multicognitive approach of the L2 pronunciation. According to this proposal, the process of becoming able to perceive, recognize and produce a new L2 sound structure demands that the learners overcome the phonological filter of their native language: different senses (auditory, visual, and tactile/kinesthetic/proprioceptive) should sustain several cognitive tasks (such as thinking, remembering, analysing, and comparing) that will "reprogram" the mental phonological filter. The senses are used, for example, in listening (to learn how to identify the sound), as well as seeing and feeling the gestures (to recognise the needed articulatory movements). It is noteworthy that the proposal by Grant (2014) includes different terms but is in line with Odisho's claims. Grant distinguishes the four levels a pronunciation teacher should focus on: the motor one (related to articulation), the perceptual (associated with distinguishing sound contrasts which are new in the L2), the cognitive (promoting the formation of new mental categories for the sounds existing only in L2), and the psycho-social (consisting of fostering positive conscious and unconscious learners' attitudes related to their pronuncia-

tion training, such as the motivation to improve their pronunciation and their autonomy in this endeavour).

## Concluding Remarks

In this paper, further validation data, literature review and reflection serve as a basis for revising the MOMUP, which as a result of this process is simplified and improved.

It must be recognized that this work presents some limitations, which remain as tasks for further research. More specifically, the new validation data introduced in this paper are based on a questionnaire applied to only 30 respondents and this sample should be further expanded. Besides, only two cases of validation are considered (a validation by application of the model presented in Castelo et al., 2022, and the present validation by questionnaire). As there are different ways of improving a model (application, questionnaires, focus groups, among others), ideally the MOMUP should be the object of more validation experiences including also other methods such as using focus groups.

In spite of these limitations, this revision process results into a model that is already based on different sources of information: literature review, reflection, and validation both by application and questionnaire. As such, MOMUP should already be considered a research-based instrument that can be very useful to teachers who want to prepare activities and/or materials to teach pronunciation through songs in a more efficient way.

## References

Ananda, F. A. (2023). English song-based learning strategies to improve English students' listening skills of English Education Department at the University of Mataram. *SHS Web of Conferences, 173*, 03004. https://doi.org/10.1051/shsconf/202317303004

Anton, R. J. (1990). Combining singing and psychology. *Hispania, 73*, 1166–1170.

Ashtian, F. T., & Zafarghandi, A. M. (2015). The effect of English verbal songs on connected speech. Aspects of adult English learners' speech production. *Advances in Language and Literary Studies, 6*(1), 212–226. https://doi.org/10.7575/aiac.alls.v.6n.1p.212

Beckford, I. C. (1980). *The development and validation of guiding principles and guideline criteria for field experiences in Jamaican teacher colleges*. [Doctoral dissertation, Florida State University, Tallahassee, FL].

Betti, S. (2012). La canción moderna en una clase de E/LE [The modern song in a class of Spanish as a Foreign Language]. *Cuadernos Cervantes de la Lengua Española, 50*, 26–31.

Castelo, A. (2017). Ensino da componente fonético-fonológica: Uma síntese e um exemplo de português para estrangeiros [Teaching the phonetic-phonological component: A synthesis and an example of Portuguese for foreigners]. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto, 12*, 41–72.

Castelo, A. (2018). Songs for pronunciation teaching: A guiding model. *Journal of Innovation in Psychology, Education and Didactics, 22*(1), 7–22.

Castelo, A. (2022). FunEasyLearn: An App for Learning Pronunciation? In A. Mesquita, A. Abreu, & J. V. Carvalho (Eds.), *Perspectives and Trends in Education and Technology. Selected Papers from ICITED 2021 (Smart Innovation, Systems and Technologies, vol. 256)* (pp. 395–405). Springer. https://doi.org/10.1007/978-981-16-5063-5_32

Castelo, A., Coelho de Souza, J., & Morelo, B. (2022). Ensino de pronúncia através de canções: Análise de material didático à luz de um modelo orientador [Teaching pronunciation through songs: A teaching material analysis in light of a guiding model]. *Études Romanes de Brno, 43*(2), 165–192. https://doi.org/10.5817/ERB2022-2-10

Celce-Murcia, M., Brinton, D. M., Goodwin, J. M., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide*. Cambridge University Press.

Coelho de Souza, J. (2014). *Letra e música: Uma proposta para o ensino da canção na aula de Português como língua adicional* [Lyrics and music: A proposal for the teaching of song in the class of Portuguese as an additional language]. [PhD dissertation, Federal University of Rio Grande do Sul, Porto Alegre, Brazil].

Degrave, P. (2019). Music in the foreign language classroom: How and why? *Journal of Language Teaching and Research, 10*(3), 412–420. http://doi.org/10.17507/jltr.1003.02

Derwing, T. M., & Munro, M. J. (2014). Myth 1: Once you have been speaking a second language for years, it's too late to change your pronunciation. In L. Grant, D. M. Brinton, T. Derwing, M. J. Munro, J. Field, J. Gilbert, J. Murphy, R. Tomson, B. Zielinski, & L. Yates (Eds.), *Pronunciation myths. Applying second language research to classroom teaching* (pp. 34–55). University of Michigan Press.

Ellis, R. (2005). Principles of instructed language learning. *System, 33*(2), 209–224. https://doi.org/10.1016/j.system.2004.12.006

Engh, D. (2013). Why use music in English language learning? A survey of the literature. *English Language Teaching, 6*(2), 113–127. https://doi.org/10.5539/elt.v6n2p113

Grant, L. (2014). Prologue to the myths: What teachers need to know. In L. Grant, D. M. Brinton, T. Derwing, M. J. Munro, J. Field, J. Gilbert, J. Murphy, R. Tomson, B. Zielinski, & L. Yates, *Pronunciation myths. Applying second language research to classroom teaching* (pp. 1–33). University of Michigan Press.

Hashemian, M., & Fadaei, B. (2011). A comparative study of intuitive-imitative and analytic-linguistic approaches towards teaching English vowels to L2 learners. *Journal of Language Teaching and Research, 2*(5), 969–976. https://doi.org/10.4304/jltr.2.5.969-976

Hişmanoğlu, M. (2006). Current perspectives on pronunciation, learning and teaching. *Journal of Language and Linguistic Studies, 2*(1), 101–110. https://www.jlls.org/index.php/jlls/article/view/26/28

Kanuka, H. (2002). Guiding principles for facilitating higher levels of web-based distance teaching and learning in post-secondary settings. *Distance Education, 23*(2), 163–182. https://doi.org/10.1080/0158791022000009187.

Keskin, F. (2011). Using songs as audio materials in teaching Turkish as a foreign language. *TOJET: The Turkish Online Journal of Educational Technology, 10*(4), 378–383. https://files.eric.ed.gov/fulltext/EJ946647.pdf

Kruk, M., & Pawlak, M. (2014). The use of internet resources in developing autonomy in learning English pronunciation: A qualitative study. In M. Pawlak, J. Bielak, & A. Mystkowska-

Wiertelak (Eds.), *Classroom-oriented research. Second language learning and teaching* (pp. 169–184). Springer International Publishing.

Kumar, T., Akhter, S., Yunus, M., & Shamsy, A. (2022). Use of music and songs as pedagogical tools in teaching English as foreign language contexts. *Education Research International, 2022*, article ID 3384067, 1–9. https://doi.org/10.1155/2022/3384067

Ludke, K. M. (2009). *Teaching foreign languages through songs.* University of Edinburgh.

Makina, A. (2020). Developing a framework for managing the quality use of podcasts in open distance and e-learning environments. *Open Praxis, 12*(1), 67–81. https://doi.org/10.5944/openpraxis.12.1.990

Medina, S. L. (2002). Using music to enhance second language acquisition: From theory to practice. In J. Lalas & S. Lee (Eds.), *Language, literacy and academic development for English language learners*. Pearson Educational Publishing.

Moncada Cevallos, C., & Chancay Cedeño, C. (2023). Songs for enhancing oral communication through the acquired of vocabulary. *Revista UNESUM-Ciencias, 7*(1), 23–36. https://doi.org/10.47230/unesum-ciencias.v7.n1.2023.707

Montaner, S. (2006). Songs in the EFL classroom. *APAC*, *56*. http://www.apac.es/publications/documents/SONGS%20IN%20THE%20EFL%20CLASSROOM.pdf

Nobre-Oliveira, D. (2007). Produção de materiais para o ensino de pronúncia por meio de músicas [Production of materials for teaching pronunciation through songs]. In V. J. Leffa (Ed.), *Produção de materiais de ensino: Teoria e prática* (pp. 109–123). Educat.

Odisho, E. (2007). A multisensory, multicognitive approach to teaching pronunciation. *Linguística: Revista de Estudos Linguísticos da Universidade do Porto, 2*, 3–28. https://ojs.letras.up.pt/index.php/EL/article/view/4015

Odisho, E. (2016). The weight of phonological vs. phonetic accent in teaching pronunciation: Implications and applications. *Linguarum Arena, 7*, 31–48. https://ojs.letras.up.pt/index.php/LinguarumArena/article/view/1766

Omolara, I. (2023). The method of teaching English using music. *International Journal Papier Public Review, 4*(1), 1–7. https://doi.org/10.47667/ijppr.v4i1.187

Rahman, N. A., Masuwai, A., Tajudin, N. M., Tek, O. E., & Adnan, M. (2016). Validation of teaching and learning guiding principles instrument for Malaysian higher learning institutions. *Malaysian Journal of Learning and Instruction, 13*(2), 125–146. https://doi.org/10.32890/mjli2016.13.2.5

Saldıraner, G., & Cintara, E. (2021). Using songs in teaching pronunciation to young EFL learners. *PASAA, 62*, 119–141. https://files.eric.ed.gov/fulltext/EJ1335005.pdf

Santos Asensi, J. (1996). Música española contemporánea en el aula de español [Contemporary Spanish music in the Spanish classroom]. In F. J. Grande Alija, J. M. Loyer, M. Rueda Rueda, & E. Prado Ibán (Eds.), *Actuales tendencias en la enseñanza del español como lengua extranjera II : Actas del VI Congreso Internacional de ASELE (León 5–7 de octubre de 1995)* (pp. 367–378). Universidad de León.

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition, 106*(2), 975–983. https://doi.org/10.1016/j.cognition.2007.03.005

Simpson, A. J. (2015). How to use songs in the English language classroom. *British Council.* https://www.britishcouncil.org/voices-magazine/how-use-songs-english-language-classroom

Upa, R., Ola, T., & Wahyono, E. (2021). Correlation between students' learning habit in listening English song and their pronunciation capacity. *Ethical Lingua, 8*(1), 110–116.

Wrembel, M. (2011). Metaphonetic awareness in the production of speech. In M. Pawlak, E. Waniek-Klimczak, & J. Majer (Eds.), *Speaking and instructed foreign language acquisition* (pp. 169–182). Multilingual Matters.

Yuhariyah, F., & Syafryadin, S. (2023). Students' perception on the use of Youtube song lyrics in constructing English language teaching activities: A teacher-student interaction. *JOLLT Journal of Languages and Language Teaching, 11*(2), 318–327. https://doi.org/10.33394/jollt.v11i2.6785

Zemlianska, N. (2021). Using pop songs in teaching English language to students majoring in music arts. Phonological aspect. *ART-platFORM, 4*(2), 45–59. https://doi.org/10.51209/platform.2.4.2021.45-59

Appendix

# Content of the Online Questionnaire Completed by the Participants

This questionnaire is part of a research study by Adelina Castelo (adelina.castelo@uab.pt). There is a **model to guide teachers when they create pedagogical materials based on songs with the goal of improving pronunciation** (among other skills) **in L2 (foreign, second, or heritage language)**. This questionnaire aims at collecting teachers' reflections on this model in order to improve it. There are no correct or incorrect answers. Please just try to reply as honestly as you can.

The questionnaire has 2 parts, will take you approximately 20 minutes to complete, and **can be completed in your mother tongue (L1)**. By replying to it anonymously, you agree that the researcher uses this data out of charge and only for research goals.

I thank you in advance for your generous contribution!

**Part 1**

1. Your L1 [you can write 2 languages if you are a "true" bilingual]
2. L2 that you teach
3. For your students, the L2 that you teach is: [you can select more than one]

– a foreign language
– a second language
– a heritage language

4. Your students are: [you can select more than one]

– primary schoolers
– middle schoolers
– high schoolers
– college/university students
– adults (with several educational backgrounds)

5. Your experience in L2 teaching (approximate number of years)
6. How often do you create pedagogical materials for L2 teaching?

– never
– seldom
– around 2 materials / month
– around 3 materials / month
– around 4 materials / month
– around 5 materials / month

– around 6 materials / month
– more than 6 materials / month

7. How often do you use songs in your L2 teaching?
– never
– seldom
– around 2 songs / month
– around 3 songs / month
– around 4 songs / month
– around 5 songs / month
– around 6 songs / month
– more than 6 songs / month

8. How important is pronunciation instruction in your teaching practice?
– 0 – not important at all
– 1
– 2
– 3
– 4
– 5
– 6
– 7 – one of the most important areas

**Part 2**

Question 1:

**Before choosing a song to train pronunciation, the teacher should decide:**

[Please classify each criterion/idea below in terms of (1) its **clarity** and (2) **importance/usefulness** for the teacher using songs to teach pronunciation.]

| | (1) This criterion/idea is... | (2) This criterion/idea is... |
|---|---|---|
| – which phonetic-phonological topic to train (i.e., consonants, vowels, processes, syllable, word stress, intonation, rhythm, spelling correspondences); | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – which other areas (besides pronunciation) to promote (i.e., linguistic, communicative, cultural); | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – which communicative competences to promote especially (i.e., listening comprehension, reading, speaking, and/or writing). | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |

Comments or suggestions related to Question 1 [optional]

Question 2:

**To train pronunciation, the teacher should choose a song that:**

[Please classify each criterion/idea below in terms of (1) its **clarity** and (2) **importance/usefulness** for the teacher using songs to teach pronunciation.]

| | (1) This criterion/idea is... | (2) This criterion/idea is... |
|---|---|---|
| – has many examples of the relevant phonetic-phonological topic to train; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – presents an adequate difficult level (in terms of vocabulary, grammar, pronunciation speed…); | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – is ethically suitable to a pedagogical context (e.g., does not promote racism); | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – will probably correspond to students' likes; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – is in accordance with teacher's likes; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – has a music that is easily remembered; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – is likely to trigger the students' affective response/reaction; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – is important in terms of L2 culture; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – has a high-quality easily accessible recording; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – has a videoclip that is suitable to a pedagogical context. | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |

Comments or suggestions related to Question 2 [optional]

Question 3:
**A pedagogical sequence including a song to train pronunciation should:**

[Please classify each criterion/idea below in terms of (1) its **clarity** and (2) **importance/usefulness** for the teacher using songs to teach pronunciation.]

| | (1) This criterion/idea is... | (2) This criterion/idea is... |
|---|---|---|
| – include activities for pre-listening, listening of the song, and post-listening; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – show harmony and unity between all its parts; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – make a good use of the lyrics' properties, in terms of text; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – explore the song's richness (e.g., relation between lyrics and music); | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – develop students' autonomy in pronunciation training; | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |
| – promote an integrated and progressively communicative approach to pronunciation (i.e., starting with explanation or discovery on the phonetic topic and moving forward from input and controlled output activities to more spontaneous and communicative speaking activities). | – totally clear.<br>– partially clear.<br>– not so clear.<br>– unclear. | – totally important/useful.<br>– partially important/useful.<br>– not so important/useful.<br>– unimportant/useless. |

Comments or suggestions related to Question 3 [optional]

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/23
https://doi.org/10.31261/TAPSLA.15784

CC BY SA

**Andrew Schenck**
https://orcid.org/0000-0002-3864-6267
Fort Hays State University, Kansas, USA

# Gleaning Insights about Input-Based and Output-Based Form-Focused Instruction through Meta-Analysis of Korean EFL Learners

Abstract

Form-Focused Instruction (FFI) has been extensively studied, yet past research often yields contradictory results. Overly simplistic examination of grammatical features, along with inadequate exploration of EFL learners at different English proficiency levels, has limited understanding of how and when to use different forms of FFI. Due to a need for further research, the present study was designed to holistically examine input and output-based forms of FFI with various grammatical features, along with Korean EFL learners at different proficiency levels. Eighteen experimental studies with Korean EFL learners were selected for metanalysis and research synthesis. The effects of type of instruction (input vs. output) were analyzed along with complexity of a target feature and learner proficiency in English. Results suggested that input-based instruction is more effective with grammatical features that involve a single phrase (e.g., basic morphology and vocabulary), while output-based instruction is more effective with more complex phrases and clauses with a syntactic component (e.g., relative clauses and conditionals). Input may activate form-meaning mapping, which is more useful for the acquisition of basic morphology or vocabulary, whereas output may help learners order constituents, thereby aiding in the acquisition of syntactic phrases and clauses. Proficiency also appears to influence the efficacy of FFI. When proficiency of participants was low and complexity of a grammatical feature was high, the influence of input tended to be stronger. When proficiency was intermediate and complexity of a grammatical feature was low, output-based instruction tended to result in larger effect sizes.

*Keywords*: form-focused instruction, input, output, Korean, L1, L2, EFL, grammar

Form-Focused Instruction (FFI), which refers to "any pedagogical practice aimed at drawing learners' attention to language form" (Collins & Ruivivar, 2020, p. 472), has been extensively studied (Piggott, 2019; Sun & Zhang, 2021, 2022). While FFI techniques may have a common purpose, they differ in their emphasis of either input or output. Whereas some techniques use input (written or oral) to draw attention to a grammatical feature, other techniques compel learners to produce output (speech or writing). Research suggests that both techniques may be effective (Kaivanpanah, Alavi, & Ravandpour, 2020). In a recent experimental study of 54 intermediate EFL learners in Iran, both input-based and output-based FFI treatments had a significant impact on production of vocabulary (Namaziandost, Dehkordi, & Shafiee, 2019).

Input-based instruction comes in many different forms. Input flood emphasizes a grammatical feature by placing many target features within a text, while input enhancement highlights a target feature within text through underlining or other forms of highlighting. Both techniques are minimally intrusive methods to focus attention on grammar. Regarding techniques of input enhancement like bolding, underlining, and italicizing, there is some research suggesting that it aids in language acquisition (Alanen, 1995; Lee, 2007; Rassaei, 2012, 2015; Sarkhosh et al., 2013). Input flood may also have an effect, albeit less so than input enhancement, which is more obtrusive (Rassaei, 2012). In addition to input enhancement and input flood, attention may be focused on a grammatical feature through Processing Instruction (PI). This pedagogical strategy helps learners to identify a relationship between form and meaning through structured input activities (Soruç, 2018). The technique is executed in three steps: providing information about the target linguistic form or structure, informing learners about input processing strategies that may negatively affect the target structure, and carrying out input-based activities that help the learner understand and process form during comprehension (Nassaji & Fotos, 2011, p. 24).

Like the modification of input, output may be emphasized to improve the acquisition of grammatical features. Having students produce more output, via speaking or writing, can enhance noticing of a grammatical feature. As pointed out by Swain (1998), output-based instruction can allow a language learner to test hypotheses about a grammatical feature or use "metatalk" to reflect on language use. Several studies suggest that tasks requiring output help learners to acquire grammatical features and vocabulary (Izumi, 2002; Rassaei, 2012; Shintani, 2011). In a study of 129 Persian learners, output-based instruction had the most significant impact on writing production of *so* and *such* clauses. Learners using this form of instruction outperformed peers who received both text enhancement and input enrichment (Rassaei, 2012). In a study of English relative clauses, Izumi (2002) also found that learners who engaged in output-based activities outperformed learners who received visual input.

## Inconsistency of FFI Results

Although both input-based and output-based FFI have had a positive influence on acquisition of grammatical features, this efficacy is, by no means, consistent. Some research suggests that techniques such as input enhancement and input flood have little to no impact on speaking or writing (Cho, 2010; Lee, 2007; Lee & Huang, 2008; Leow et al., 2003; Rassaei, 2012). Other research suggests that the effects of PI on grammatical accuracy are limited with respect to production in speech or writing. While learners with PI treatment tend to perform better on interpretation tasks, their improvement on production tasks is often lackluster (Benati, 2005; Comer & deBenedette, 2011). Such research, along with other studies that report more significant outcomes for output-based instruction (Izumi, 2002; Rassaei, 2012, 2015), cast some doubt on the efficacy of input-based instruction.

As with input-based FFI, studies of output-based instruction often yield inconsistent results. Regarding vocabulary, Shintani (2011) found that both comprehension (input) and production (output) activities had a positive effect on production, yet the input-based tasks allowed for richer interaction that led to higher gains on a comprehension test. Benati (2001) also found that input-based interventions were more effective than output-based instruction which uses grammar rules followed by written or oral practice. Some studies of output-based FFI even suggest that the technique may be ineffectual. A study by Izumi and Bigelow (2000), for example, concluded there were no unique effects from providing an output-based task. Another study by Izumi and Izumi (2004), which examined the use of oral output, also found a nonsignificant result for output-based instruction, concluding that "re-examination of the treatment tasks revealed that the output task failed to engage learners in the syntactic processing that is necessary to trigger L2 learning, while the task for the nonoutput group appeared to promote better form-meaning mapping" (p. 587).

Past research concerning both input-based and output-based instructional emphasis of grammar has led to variable results, making interpretation of findings problematic. Without a clear ability to predict educational outcomes, teachers cannot be assured that FFI will be effective. More research is needed to provide a holistic perspective of FFI and associated influences, which may then allow educators to choose FFI techniques that are both consistent and reliable.

## Timing in FFI and the Need for Further Research

Conflicting beliefs concerning the efficacy of different forms of FFI reveal a fundamental gap in existing research. Although FFI techniques may be effective, precisely how and when these techniques should be introduced to ensure consistent effectiveness is not yet known. Recently, the importance of timing in the implementation of FFI has been considered. In a study by Xu and Li (2021), FFI given before a task led to heightened efficacy for a difficult grammatical feature, whereas FFI given after a task led to heightened efficacy for an easier target feature. Such study illustrates the importance of providing FFI at an optimal time based upon grammatical feature, thereby enhancing effectiveness. While a positive step forward, it does not adequately examine issues of timing associated with different grammatical features or varied proficiency levels.

When considering the timing of FFI, characteristics of diverse grammatical features and L2 English proficiency must also be carefully considered. Research suggests that learners must be at a specific level of proficiency before a target feature can be successfully taught and acquired (Dyson, 2018; Dyson & Håkansson, 2017). According to the Processability Theory, the effectiveness of FFI is influenced by two variables, grammatical complexity and L2 English proficiency. As learners become more proficient in English as an L2, they also gain the ability to acquire (and be explicitly taught) more complex grammatical features (Pienemann & Lenzing, 2015). First, learners gain the ability to acquire intra-phrasal features that modify a single phrase. An example would be changing a present tense verb into the past tense (e.g., *eat* to *ate*; *walk* to *walked*) or making a noun into a plural (e.g., *book* to *books*). In each case, internal modification of the phrase is required. Next, learners can acquire inter-phrasal target features. This grammar requires the modification of multiple phrases on a syntactic or morphological level. An example would be inverting the auxiliary verb *do* with the subject, a noun phrase. In addition to question formation, morphological modification of a verb for the third person singular (e.g., He read*s*) reveals a cognitive connection between two phrases (the subject noun phrase and the verb). Following the acquisition of inter-phrasal features, learners who are the most proficient in English can acquire clausal features, which require the creation of larger subordinate clauses. An example of this type of grammatical feature is the conditional (e.g., *If I hadn't eaten breakfast*, *I would have been hungry right now*). Due to the morphosyntactic complexity of clausal features, they tend to be acquired last (Pienemann, 2005). As revealed by stages of the Processability Theory, heightened grammatical complexity may have a significant impact on acquisition, suggesting that learners need to be

at a certain English proficiency level to benefit from FFI for a specific target feature (Gholami & Zeinolabedini, 2018).

The Processability Theory appears to have implications about timing for the use of specific types of FFI. For example, there may be an optimum time to introduce input- and output-based FFI according to grammatical complexity and L2 English proficiency. Input-based instruction promotes noticing of a grammatical feature (Winke, 2013), which may prime the lexicon or serve as a scaffold when learners have had little exposure to a target feature. Output-based instruction moves learners away from "semantic, open-ended" forms of comprehension toward "complete grammatical processing needed for accurate production" (Swain, 1995, p. 128). Essentially, output pushes the learner to produce a target feature, which may be more effective when learners have already had adequate exposure to the feature and sufficient proficiency in English (Namaziandost, Dehkordi, & Shafiee, 2019). Input and output-based forms of FFI may help learners acquire a grammatical feature at an opportune time. In addition, different forms of FFI may target different forms of grammar (e.g., syntactic vs. morphological). Whereas input could provide valuable information for form-meaning mapping, output-based instruction may prompt learners to examine syntactic relationships and word order, since learners are pushed to order constituents of a phrase or sentence via this approach (Gass & Selinker, 2009). More research is needed to examine the effects of both timing and the use of different forms of FFI. This research must include a variety of target features and learners at different English proficiency levels. Currently, the correct time to introduce different forms of FFI remains unclear. Further examination of the effects of timing may explain inconsistent results of past experimentation, clarifying how and when different FFI techniques should be utilized.

## Research Questions

The present meta-analysis and research synthesis was designed to investigate the effectiveness of input and output-based FFI when used with different grammatical features and learners of varied English proficiency. To aid in the investigation, the following questions were posed:

1. What types of instruction (input-based or output-based) are most effective with each type of grammatical feature (intra-phrasal, inter-phrasal, and clausal)?
2. Does the effectiveness of an instructional type (input-based or output-based) differ according to the English L2 proficiency level of a learner?

Through investigating the questions above, it was hoped that a holistic understanding could be obtained to reform curricula or enhance automated language learning systems, thereby tailoring instruction to learner needs.

## Method

The present meta-analysis synthesized several experimental studies to evaluate the influence of grammatical complexity (intra-phasal, inter-phrasal, and clausal), type of instruction (input-based vs. output-based), and learner background (English proficiency) on accuracy of production in speech or writing. Google was systematically searched by using the keywords *Korean* with various search terms for grammatical features (*plural*, *past tense*, *past regular*, *past irregular*, *passive*, *third person*, *questions*, *article*, *definite article*, *indefinite article*, *phrasal verb*, *verb particle*, *conditional*) and types of FFI treatments (*form-focused instruction*, *focus-on-form*, *focus-on forms*, *PI*, *text enhancement*, *dictogloss*, *output*, *input*, *control group*). Following the search, full texts for each study were obtained for further examination.

There are key differences between explicit knowledge and actual performance in production (implicit knowledge). Therefore, only studies that elicited responses in speech and writing were selected. To help ensure that production reflected implicit knowledge of a target feature, testing was selected that communicated ideas, not rules; put pressure on learners to prevent conscious correction of language errors; focused on meaning not form; and avoided use of metalanguage (Ellis, 2009). In order to be included within the present meta-analysis, each experimental study needed to have:

1. An input-based or output-based treatment (including time for treatment and methods of delivery);
2. Pretest and Posttest measures of production (either oral or written);
3. Information about the type of grammatical feature targeted;
4. Participants with Korean as their L1.

Information about proficiency was often variable and inconsistent. Studies that did include this variable were placed in basic categories such as beginner, intermediate, and advanced for further examination in this study.

Originally 59 experimental studies with South Korean EFL learners were located. From this group, only 18 studies met the criteria for inclusion. Many studies lacked adequate assessment of productive, implicit knowledge, leading to exclusion (See Appendix B for more information on assessments used in the selected studies). In addition to problems with assessment of production, some studies lacked sufficient knowledge to understand the methodology or length

of treatment. Thus, they were excluded. Totally, the present meta-analysis contained 18 studies for analysis. In total, there were 47 treatment groups.

## Grammatical Feature Type

After studies were compiled, results from each study were organized and evaluated according to grammatical feature. Types of grammatical feature were grouped for comparison. As revealed by the Processability Theory, grammatical complexity could vary based on whether a target feature is intra-phrasal (e.g., a verb or noun and associated morphological features like the regular past *-ed* or plural *-s*), inter-phrasal (e.g., question inversion or phrasal verbs), or clausal (e.g., relative clauses or conditionals) (Pienemann & Lenzing, 2015). Grammatical features from the selected studies were taken from Appendix A and separated as in Table 1.

**Table 1**

*Separation of Grammatical Features Based on Processability Theory*

| | Target Features (Korean Studies) |
|---|---|
| Intra-phrasal | Single adjectives, verbs, and nouns (2 Studies)<br>Present Perfect (4 Studies)<br>Collocations (despair, earnestly, etc.) (1 Study)<br>Verb tenses (1 Study)<br>Participial adjectives (e.g., *boring/bored*) (1 Study) |
| Inter-phrasal | Comparative adjectives and *than* (smarter than X) (1 Study)<br>Verb + Pronoun + to + V (I want her to visit) (1 Study) |
| Clausal | Conditional (5 Studies)<br>That-Trace Filter (1 Study)<br>Relative Clauses (1 Study) |

In total, 47 treatment groups were obtained from the selected studies. Because nine of the control groups contained collocations with more than one grammatical feature from a different complexity level, these treatment groups were excluded from analysis of grammar type, leaving a remainder of 38 treatment groups.

## Proficiency Level

Studies chosen for meta-analysis could include different proficiency designations such as low beginner, beginner, high beginner, intermediate low, intermediate, or high intermediate (See Appendix A). However, this variation was not systematic or complete. As an example, the beginner level included only students who were evaluated as "low English proficiency" by using TOEIC scores that ranged from 250 to 320. Due to variability in how students were assessed as either high or low, level designations were collated into three general groups: beginner, intermediate, and advanced. Of the 18 studies with Korean participants, only 14 had information for designation of proficiency level (8 beginner treatment groups and 31 intermediate treatment groups). Information provided (or not provided) concerning proficiency levels confirms assertions by Liu and Brown (2015), who contend that methodological inconsistencies have limited our understanding of ESL or EFL instruction.

While studies selected for meta-analysis did generally include some assessment of levels, they should be interpreted with caution. Because different instruments were used to assess proficiency, discrepancies between assessment of proficiency may be expected. Ultimately, assignment of proficiency levels in selected studies reflects the researcher's judgment, meaning that this variable cannot be interpreted as an empirical variable based on a standardized assessment. It must be considered a construct of the researcher, like variables obtained from survey data. Classifications of proficiency level may reveal some trends that can be confirmed later through follow-up experimental research with more standardized instruments.

## Input-Based Output-Based Definitions

Studies designed to evaluate the efficacy of either input or output were selected and separated based on instructional type (See Appendix B for more information about treatments). Any studies that had a major emphasis on both input and output in the FFI treatment were excluded from analysis. Whereas treatments primarily designed to emphasize the impact of input (e.g., input flood, IE, and PI) were assigned to the input category, tasks that emphasized output (e.g., text reconstruction or dictogloss) were assigned to the output category. The input vs. output distinction was used to analyze differences in effect sizes, along with variables such as grammatical complexity, learner proficiency, and

learner L1. For input, there were a total of 30 treatment groups; for output, there were a total of 17 groups.

Overall, output-based treatments included a variety of both written and spoken tasks. Among these tasks, the dictogloss encouraged both written and spoken output. As a text is being reconstructed, learners may use metatalk to discuss the target feature. Other studies of output-based assessments like that of Kim (2014) utilized images to elicit verbal responses about a target feature. With the exception of studies that used only the dictogloss, there was little standardization of the techniques used to elicit output. While forms of production did vary, Swain (1998) points out that all production tasks give learners the ability to use and test hypotheses about a target feature. Some output treatments did include a degree of explicit information or guidance to conduct the activity, which was a type of input. In each treatment, however, emphasis was placed on producing output, rather than providing input. Studies that used a dictogloss, for example, required input before the story was reconstructed. However, the main goal of the activity was output, as reflected by procedures that included notetaking, meta-talk, and story construction. Although some input may have been provided with output treatment groups, the main goal of these groups was to produce either an oral or written product. Therefore, a clear emphasis was placed on production, rather than input, allowing learners time to test hypotheses as envisioned by Swain (1998). Some studies provided an emphasis on both input and output in the same treatment groups. These treatment groups were excluded from analysis. Control groups with no treatment were also excluded from analysis.

## L1 Influence

Research suggests that L1 transfer does indeed have an impact on acquisition, thereby impacting the effectiveness of FFI (Luk & Shirai, 2009; Maleki, 2006; Shin, 2015; Yang et al., 2017). Therefore, EFL learners with the same native language (Korean) were studied. Examination of learners with the same L1 helped to ensure that differences in experimental results were related to grammatical feature type and L2 English proficiency, rather than differences due to L1 transfer. Grammatical features targeted in experimental studies emphasized English grammatical features that were very different from Korean. Studies emphasized grammatical features like relative clauses or conditionals, which use a head final parameter (opposite from the predominant head initial parameter in English). Other studies examined verb tenses. Korean verb tenses for the simple past resemble English in some ways, yet the present perfect and

present perfect progressive tenses lack equivalent morphosyntactic structures in Korean. Instead, Korean uses a variety of morphological verb endings and adverbials to express aspect. Other features such as participials (e.g., *I'm bored* vs. *The book is boring*) and verb/pronoun combinations (e.g., *ask me to*) also lack equivalent structures in Korean. Overall, English grammar emphasized within FFI studies was significantly different from grammatical features found in the Korean L1 (See Appendix A for more information on target features used for each experimental study).

## Procedure

To compare results from individual studies, effect sizes needed to be calculated for each study. Whereas *p* values reveal whether results are significant (not a result of chance), effect sizes determine the magnitude of a difference between groups (Sullivan & Feinn, 2012). Since significant *p* values may not actually reveal a large effect (e.g., large numbers of participants or amounts of data may reveal only a small significant difference), effect size is needed to understand how effectual a treatment is. Calculation of effect size also provides a consistent way to compare different studies, since the calculations are standardized measures.

Cohen's *d* was selected since it is widely used to compare two means within meta-analysis studies and provides "an important complement to traditional null hypothesis statistical significance testing" (Diener, 2010, p. 1). It was also used in the formative meta-analysis of FFI conducted by Spada and Tomita (2010), which analyzed effects of explicit and implicit instruction on the acquisition of simple and complex grammatical features in English. In the current study, effect size was calculated by inserting pretest scores (M2), posttest scores (M1), and associated standard deviations (SD2 and SD1) into the Cohen's *d* formula for effect size (Spada & Tomita, 2010, p. 307):

$$d = [M1 - M2] / [SQRT[(SD1SD1 + SD2SD2]/2]$$

After effect sizes were calculated for each treatment group, results were collated based upon the variables, allowing for further analysis. For grammatical complexity, effect sizes were collated based upon whether an intra-phrasal, inter-phrasal, or clausal feature was emphasized. For proficiency level, effect sizes were collated according to researcher designations of beginner or intermediate (there were no advanced). Average effect sizes for both grammatical complexity and proficiency level were then subdivided based upon instructional type (input or output) for further analysis.

# Results and Discussion

## Grammatical Complexity and the Efficacy of FFI

Research of grammatical features revealed interesting differences between the use of input-based and output-based instruction. Table 2 revealed that input-based instruction was more effective when less complex intra-phrasal features were emphasized, while output-based instruction was more effective for inter-phrasal and clausal features. As for intra-phrasal features, input-based instruction had a larger average impact of .42 over its output-based counterpart. Because prior research of meta-analysis suggests that a small effect size is $d > 0.2$, a medium effect size is $d > 0.5$, and a large effect size is $d < 0.8$ (Rice & Harris, 2005), this value would represent a small but substantial difference. Inter-phrasal features revealed a much higher value for output-based instruction, which represented a difference in effect size of .67, which would be a medium effect. Although significantly different, this value should be interpreted with caution. The output-based score is determined through only one study by Hwang (2018), which examined comparative adjectives with *than*. Concerning clausal features, output-based instruction was more effective by .09.

**Table 2**

*Average Effect Size by Grammatical Complexity and Type of Instruction*

| | Input-Based vs. Output-Based | Mean | *N* | Std. Deviation |
|---|---|---|---|---|
| Intra-phrasal | Input | 1.9548 | 10 | 1.08530 |
| | Output | 1.5345 | 8 | .70399 |
| Inter-phrasal | Input | 1.4015 | 6 | .63283 |
| | Output | 2.0699 | 1 | . |
| Clausal | Input | 1.2421 | 7 | 1.11637 |
| | Output | 1.3345 | 6 | 1.10216 |

Intra-phrasal features tended to benefit more from input-based instruction. Experimental studies of vocabulary and collocations by Lee (2002) and Lee (2003) both show higher effect sizes for input over output. Experimental studies that examine verb tenses also show a greater impact for input (Yang, 2004, 2008; Hwang, 2018). Findings may be explained by the characteristics of target features, which have a large number of lexical elements that require form-meaning mapping. Input may provide the semantic and morphological information necessary to prime the learner, thereby aiding in the acquisition of form-

meaning mappings. This perspective may also explain the results of a study by Kim and Nam (2017), which revealed that emphasis of inter-phrasal collocations such as "shake a leg" or "hit the sack" via input-based instruction yielded very high effect sizes that averaged 3.25.

Input may provide scaffolding for morphological and lexical grammatical features that require form-meaning mapping early in the acquisition process. As the complexity of syntax increases, output-based instruction may be more effective, pushing learners to arrange syntactic elements into the correct order. Concerning clausal features, output-based instruction was more effective for both inter-phrasal and clausal features. This finding may reflect an effect of output-based instruction on more syntactically complex features that require ordering of constituents. Experimental studies which had both input and output-based treatments for conditionals (Song, 2007; Shin, 2011) had higher effect sizes for output. A study of relative clauses by Hwang (2018) likewise shows a larger effect size for output over input. When producing output, the learner is forced to attend to word order, thereby producing a correct oral or written utterance. Output-based activities may activate cognitive mechanisms that deal with syntax.

Input or output-based instruction may emphasize different characteristics of grammar. Since the difficulty posed by lexical variation or syntactic complexity is also partially determined by L2 proficiency, characteristics of the learner must be further studied to clarify the impact of input- and output-based FFI.

## English Proficiency and the Efficacy of FFI

Separation of effect size based upon proficiency suggests that output-based instruction is more effective at the beginner level, while input-based instruction is more effective at the intermediate level (Table 3).

**Table 3**

*Mean Effect Sizes for Input and Output-Based Instruction at Two Proficiency Levels*

| | Input-Based vs. Output-Based | Mean | *N* | Std. Deviation |
|---|---|---|---|---|
| Beginner | Input | 2.1042 | 4 | .82168 |
| | Output | 2.3800 | 4 | .97856 |
| Intermediate | Input | 1.8029 | 21 | 1.25458 |
| | Output | 1.0781 | 10 | .38060 |

At the beginner level, the difference is .28, a small magnitude. Although this finding appears to suggest that input-based instruction is more beneficial for beginner learners, the findings should be interpreted with caution. Results for the beginner levels were based upon just one study by Hwang (2018), which analyzed verb tenses (intra-phrasal), comparative adjectives with *than* (inter-phrasal), and relative clauses (clausal). The study revealed that input-based instruction was more effective for both verb tenses (difference in effect size of .47) and comparatives (difference in effect size of .18). Output-based instruction was more effective for relative clauses (difference in effect size of 2.28). Interestingly, the effect of output-based instruction systematically increases, and the effect of input-based instruction decreases as the grammatical feature becomes more complex. Thus, the study lends further for the idea that input-based instruction is effective for intra-phrasal morphology, whereas output-based instruction is more effective for inter-phrasal and clausal grammatical features.

At the intermediate proficiency level, the difference is substantial (a difference of .72), which suggests a medium effect. The results were heavily influenced by studies like those of Yang (2004, 2008), who emphasized the present perfect progressive, and Kim and Nam (2017), who emphasized inter-phrasal lexical collocations. All of these studies had higher effect sizes for input, lending further support for the idea that input is more effective for form-meaning mapping of morphological and lexical features.

In studies by Kim (2002) and Yeo (2002), which emphasized participial adjectives (*bor**ing**/bor**ed***), learners performed better when output-based instruction was used (differences in effect sizes of 1.00 and .52, respectively). Unlike the more complex morphosyntax of present perfect progressive tenses emphasized by Yang (2004, 2008) or inter-phrasal collocations emphasized by Kim and Nam (2017), participial phrases are simple, binary morphemes. Perhaps low complexity of the target feature, along with intermediate English L2 proficiency explains the findings. Output may be more effective with simplistic morphological features when English L2 proficiency is high enough. In the study by Kim (2002), an inter-phrasal feature (verb + *me* + *to*) was also included in the same treatment and assessment of participial phrases. Inclusion of a syntactic feature may explain why the output-based instruction was nearly twice as effective than that of Yeo (2002), who examined only participial adjectives. While a potentially valid interpretation of research results, it is important to note that the studies by Yeo (2002) and Yang (2004) did not include information about learner proficiency. Further research is needed to ensure that proficiency level influences FFI outcomes for target features such as participial adjectives or present perfect tenses.

## Conclusion

The present study was designed to facilitate understanding of how input- and output-based FFI can be utilized in a more timely and effective way. Results appear to suggest that efficacy of either input or output-based instruction is dependent on characteristics of the target feature. Simplistic intra-phrasal features (morphological and lexical features) may benefit more from input. Input-based instruction allows learners to consider form-meaning mappings, without a need to attend to issues of word order. More complex inter-phrasal and clausal features may benefit more from output-based instruction. In contrast to input-based FFI, output-based instruction forces the learner to address issues related to word order, which may explain why inter-phrasal and clausal syntactic features (e.g., subject/verb inversion, relative clauses, or conditionals) had larger effect sizes for this instructional style.

There may also be a relationship between L2 English proficiency and the efficacy of different types of FFI. While collation of effect sizes appears to suggest that beginner levels benefit more from output and intermediate levels benefit more from input, systematic exploration of individual studies appears to suggest the opposite. Using average effect sizes for proficiency does not include consideration of the challenges posed by complexity of a target feature. Results from individual studies provide some evidence that simplicity of a target feature and intermediate proficiency make output-based instruction more effective. At lower proficiency levels, input appears to be more effective when complexity of a grammatical feature is high. Such findings reveal a close link between grammatical feature and proficiency level. Both variables may need to be considered before the effectiveness of different FFI types can be accurately assessed. Although conclusions about proficiency are interesting, they must be interpreted with caution. Assessment of English proficiency was not consistent within past studies, which may have influenced the findings.

Results of the present study have implications for both instruction and the development of educational technology. Input may need to be modified (or remain unmodified) based upon the target feature chosen for emphasis. Whereas modifying input can promote acquisition of intra-phrasal features related to verb tense or noun forms, it may be less effective for more complex inter-phrasal and clausal features. Rather than input-based FFI, enhanced output-based tasks may be used to promote the acquisition of more complex syntactic features such as question inversion or relative clauses. While a target feature partially dictates whether input- or output-based FFI should be used, a learner's proficiency level should also be considered. Learners must be at a proficiency level which allows for the acquisition of a target feature. As proficiency rises, cognitive resources or abilities may also increase, meaning that different tasks (input vs.

output) could be effective at different times. In the future, language instruction and educational technology will need to diversify FFI, carefully controlling administration of techniques based upon target feature and proficiency level. In this way, English instruction may be tailored to the unique needs of each individual learner.

While some very insightful information was gleaned from existing studies of FFI, there were key limitations that need to be addressed in future research. The number of grammatical features emphasized in these studies was limited, as was the variety of L2 proficiency levels examined in the studies. Clearly, more research is needed. Experimental studies conducted to date often examine the same grammatical features or use learners at similar proficiency levels (e.g., intermediate). As a result, the impact of different target features and timing (introduction at the right level of L2 proficiency) is unknown. We must fully examine the range of grammatical features presented at a variety of proficiency levels, thereby leading to a concrete understanding of how and when to use FFI.

## References

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 259–302). University of Hawaii Press.

Benati, A. (2001). A comparative study of the effects of processing instruction and output-based instruction on the acquisition of the Italian future tense. *Language Teaching Research, 5*(2), 95–127. https://doi.org/10.1177/136216880100500202

Benati, A. (2005). The effects of processing instruction, traditional instruction and meaning—Output instruction on the acquisition of the English past simple tense. *Language Teaching Research, 9*(1), 67–93. https://doi.org/10.1191/1362168805lr154oa

Cho, M. Y. (2010). The effects of input enhancement and written recall on noticing and acquisition. *Innovation in Language Learning and Teaching, 4*(1), 71–87. https://doi.org/10.1080/17501220903388900

*Cho, Y. (2011). Effects of different types of focus-on-form instruction and learner readiness on the learning of perfective modals. *Foreign Languages Education, 18*, 1–27.

Collins, L., & Ruivivar, J. (2020). Form-Focused Instruction. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics* (pp. 472–502). John Wiley and Sons. https://doi.org/10.1002/9781405198431.wbeal0432.pub2

Comer, W. J., & deBenedette, L. (2011). Processing instruction and Russian: Further evidence is in. *Foreign Language Annals, 44*(4), 646–673. https://doi.org/10.1111/j.1944-9720.2011.01155.x

Diener, M. J. (2010). Cohen's *d*. In *The Corsini encyclopedia of psychology*. https://doi.org/10.1002/9780470479216.corpsy0200

Dyson, B. (2018). Developmental sequences. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–8). John Wiley & Sons.

Dyson, B. P., & Håkansson, G. (2017). *Understanding second language processing: A focus on processability theory* (Vol. 4). John Benjamins.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*(3), 221–246. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Gass, S. M., & Selinker, L. (2009). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Gholami, J., & Zeinolabedini, M. (2018). Learnability and teachability hypothesis. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–7). John Wiley & Sons.

*Hwang, H. J. (2018). A study on the Korean EFL learners' grammatical knowledge development under input-enhanced FFI and output-enhanced FFI conditions. *Journal of Digital Convergence*, *16*(5), 435–443. https://doi.org/10.14400/JDC.2018.16.5.435

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*(4), 541–577. https://doi.org/10.1017/S0272263102004023

Izumi, S., & Bigelow, M. (2000). Does output promote noticing and second language acquisition?. *TESOL Quarterly, 34*(2), 239–278. https://doi.org/10.2307/3587952

Izumi, Y., & Izumi, S. (2004). Investigating the effects of oral output on the learning of relative clauses in English: Issues in the psycholinguistic requirements for effective output tasks. *Canadian Modern Language Review, 60*(5), 587–609. https://doi.org/10.3138/cmlr.60.5.587

*Jeong, H., & Lee, H. (2018). The effects of two types of written output tasks on noticing and learning of vocabulary. *Korean Journal of Applied Linguistics, 34*(4), 75–102. https://doi.org/10.17154/kjal.2018.12.34.4.75

Kaivanpanah, S., Alavi, S. M., & Ravandpour, A. (2020). The effect of input-based and output-based tasks with different and identical involvement loads on Iranian EFL learners' incidental vocabulary learning. *Cogent Psychology, 7*(1), 1–17. https://doi.org/10.1080/23311908.2020.1731223

*Kang, D. (2011). Noticing and L2 learning in input-based focus-on-form instruction. *Foreign Languages Education, 18*(2), 29–49. UCI: G704-000287.2011.18.2.017

*Kang, D. H. (2003). Focus on form instructions and L2 learners' instructional preferences. *Foreign Languages Education*, *10*(1), 57–82. UCI: G704-000287.2003.10.1.007

*Kang, N. (2009). Effects of form-focused instructions on the learning of English verb complementation by Korean EFL learners. *English Teaching, 64*(1), 3–25. https://doi.org/10.15858/engtea.64.1.200903.3

*Kim, B. (2002). The role of focus on form in the EFL classroom setting: Looking into recast and CR tasks. *English Teaching, 57*(4), 267–296. UCI: G704-000189.2002.57.4.022

*Kim, J. E. (2014). Timing of form-focused instruction and development of implicit vs. explicit knowledge. *English Teaching, 69*(2), 123–147.

*Kim, J. E., & Nam, H. (2017). The pedagogical relevance of processing instruction in second language idiom acquisition. *International Review of Applied Linguistics in Language Teaching, 55*(2), 93–132. https://doi.org/10.1515/iral-2015-0027

*Lee, H. (2002). Communicative output as a mode of focus on form. *English Teaching, 57*(3), 171–192. UCI: G704-000189.2002.57.3.006

*Lee, H. (2003). The effects of production and comprehension for focus on form and second language acquisition. *Korean Journal of Applied Linguistics, 19*(2), 41–68. UCI: G704-000491.2003.19.2.010

Lee, S. K. (2007). Effects of textual enhancement and topic familiarity on Korean EFL students' reading comprehension and learning of passive form. *Language Learning, 57*(1), 87–118. https://doi.org/10.1111/j.1467-9922.2007.00400.x

Lee, S. K., & Huang, H. T. (2008). Visual input enhancement and grammar learning: A meta-analytic review. *Studies in Second Language Acquisition, 30*(3), 307–331. https://doi.org/10.1017/S0272263108080479

Leow, R. P., Egi, T., Nuevo, A. M., & Tsai, Y. C. (2003). The roles of textual enhancement and type of linguistic item in adult L2 learners' comprehension and intake. *Applied Language Learning, 13*(2), 1–16.

*Lim, J. (2007). Input enhancement in the EFL learning of present perfect and the lexical aspect. *Modern English Education, 8*(3), 38–60. UCI: G704-001860.2007.8.3.005

Liu, Q., & Brown, D. (2015). Methodological synthesis of research on the effectiveness of corrective feedback in L2 writing. *Journal of Second Language Writing, 30*, 66–81. https://doi.org/10.1016/j.jslw.2015.08.011

Luk, Z. P. S., & Shirai, Y. (2009). Is the acquisition order of grammatical morphemes impervious to L1 knowledge? Evidence from the acquisition of plural -s, articles, and possessive 's. *Language Learning, 59*(4), 721–754. https://doi.org/10.1111/j.1467-9922.2009.00524.x

Maleki, A. (2006). The impact of the head-initial/head-final parameter on reading English as a foreign language: A hindering or a facilitating factor?, *The Reading Matrix, 6*, 154–169.

Namaziandost, E., Dehkordi, E. S., & Shafiee, S. (2019). Comparing the effectiveness of input-based and output-based activities on productive knowledge of vocabulary among pre-intermediate EFL learners. *Asian-Pacific Journal of Second and Foreign Language Education, 4*, 1–14. https://doi.org/10.1186/s40862-019-0065-7

Nassaji, H., & Fotos, S. S. (2011). *Teaching grammar in second language classrooms: Integrating form-focused instruction in communicative context*. Routledge.

Pienemann, M. (2005). *Cross-linguistic aspects of Processability Theory* (Vol. 30). John Benjamins Publishing.

Pienemann, M., & Lenzing, A. (2015). Processability Theory (2nd ed.). In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 159–179). Routledge. https://doi.org/10.4324/9780203628942

Piggott, L. (2019). *First meaning then form: A longitudinal study on the effects of delaying and reducing explicit form-focused instruction for young adolescent EFL learners*. Doctoral dissertation, Utrecht University, The Netherlands. LOT Dissertation Series, number 545. LOT.

Rassaei, E. (2012). The effects of input-based and output-based instruction on L2 development. *TESL-EJ*, *16*(3), 1–26. https://tesl-ej.org/wordpress/issues/volume16/ej63/ej63a2/

Rassaei, E. (2015). Effects of textual enhancement and input enrichment on L2 development. *TESOL Journal, 6*(2), 281–301. https://doi.org/10.1002/tesj.149

Rice, M. E., & Harris, G. T. (2005). Comparing effect sizes in follow-up studies: ROC Area, Cohen's d, and r. *Law and Human Behavior*, *29*(5), 615–620. https://doi.org/10.1007/s10979-005-6832-7

Sarkhosh, M., Taghipour, B., & Sarkhosh, H. (2013). Differential effect of different textual enhancement formats on intake. *Procedia-Social and Behavioral Sciences, 70*, 544–559. https://doi.org/10.1016/j.sbspro.2013.01.092

*Shin, M. K. (2011). Effects of output tasks on Korean EFL learners' noticing and learning of English grammar. *Foreign Language Education, 18*(2), 127–163. UCI: G704-000287.2011.18.2.014

Shin, M. K. (2015). The resetting of the head direction parameter. *Studies in Foreign Language Education, 18*, 17–35. http://hdl.handle.net/10371/95840

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*(2), 137–158. https://doi.org/10.1177/1362168810388692

*Song, M. J. (2007). Getting learners' attention: Typographical input enhancement, output, and their combined effects. *English Teaching, 62*(2), 193–215. https://doi.org/10.15858/engtea.62.2.200706.193

*Song, M. J., & Suh, B. R. (2008). The effects of output task types on noticing and learning of the English past counterfactual conditional. *System, 36*(2), 295–312. https://doi.org/10.1016/j.system.2007.09.006

Soruç, A. (2018). The mediating role of explicit information in processing instruction and production-based instruction on second language morphological development. *Educational Sciences: Theory & Practice, 18*(3), 279–302. https://dergipark.org.tr/en/pub/buje/issue/55539/760450

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning, 60*(2), 263–308. https://doi.org/10.1111/j.1467-9922.2010.00562.x

Sullivan, G. M., & Feinn, R. (2012). Using effect size—or why the *P* value is not enough. *Journal of Graduate Medical Education, 4*(3), 279–282. https://doi.org/10.4300/JGME-D-12-00156.1

Sun, Q., & Zhang, L. J. (2021). A sociocultural perspective on English-as-a-foreign-language (EFL) teachers' cognitions about form-focused instruction. *Frontiers in Psychology, 12*, 1–14. https://doi.org/10.3389/fpsyg.2021.710736

Sun, Q., & Zhang, L. J. (2022). Understanding novice and experienced teachers' cognitions and practices for sustainable teacher development: the case of form-focused instruction in English language teaching. *Sustainability, 14*(8), 1–17. https://doi.org/10.3390/su14084711

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidlhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125–144). Oxford University Press.

Swain, M. (1998). Focus on form through conscious reflection. In C. Doughty & J. Williams (Eds.), *Focus on form in classroom second language acquisition* (pp. 64–81). Cambridge University Press.

Winke, P. M. (2013). The effects of input enhancement on grammar learning and comprehension: A modified replication of Lee (2007) with eye-movement data. *Studies in Second Language Acquisition, 35*(2), 323–352. https://doi.org/10.1017/S0272263112000903

Xu, J., & Li, C. (2021). The effects of the timing of form-focused instruction on EFL learners' learning of difficult and easy grammatical features: A comparative study. *System, 101*. https://doi.org/10.1016/j.system.2021.102612

*Yang, M. (2004). A comparison of the efficacy between input-based instruction and output-based instruction in focus on form. *English Teaching, 59*(2), 145–164. UCI: G704-000189.2004.59.2.012

*Yang, M. (2008). The effects of task difficulty on learners' attention to meaning and form during focus-on-form instruction. *Foreign Languages Education, 15*(3), 27–52. UCI: G704-000287.2008.15.3.006

Yang, M., Cooc, N., & Sheng, L. (2017). An investigation of cross-linguistic transfer between Chinese and English: A meta-analysis. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(15), 1–21. https://doi.org/10.1186/s40862-017-0036-9

*Yeo, K. (2002). The effects of dictogloss: A technique of focus on form. *English Teaching, 57*(1), 149–167. UCI: G704-000189.2002.57.1.011

## Declarations:

**Availability of Data and Materials**: All data obtained for the meta-analysis is accessible through examination of published research studies.
**Competing Interests**: The author declares that he has no competing interests.
**Funding**: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
**Author's Contribution**: The author has written and reviewed this paper.
**Acknowledgement**: N/A
**Ethics Statement**: No human subjects were used in this study.

Appendix A

# Studies of Input/Output Using Learners with Korean L1

| Authors | Duration | Learner Proficiency (Number) | Grammar Feature |
|---|---|---|---|
| Kang (2011) | 3 hours in one week | TOEIC 420 to 930 (Average 675) (*n* = 15) | conditional |
| Kang (2003) | 12 hours / 6 hours for two weeks | no info (*n* = 134) | conditional |
| Song and Suh (2008) | 3 sessions over a month period | Intermediate (*n* = 52) | conditional |
| Yang (2008) | 2-hour treatment | approximately 140 to 185 on reading TOEIC (*n* = 70) | present perfect progressive |
| Kim (2014) | 3 sessions of treatment over 3 days | Intermediate level self-rating (*n* = 42) | **Who do you think that married Sarah last year?* / that-trace filter |
| Kang (2009) | eight 25-minute form focused treatment over 4 weeks | Intermediate (*n* = 150) | Verb + NP + to infinitive / ask me to |
| Song (2007) | 3 sessions of about 30 minutes over a 3-week period | intermediate TEPS 501 to 700 (average 600.5) (*n* = 140) | conditional |
| Yang (2004) | 2-hour class | no info (*n* = 70) | present perfect progressive |
| Kim (2002) | eight 25-minute focus on form treatments over four weeks | low intermediate (*n* = 72) | participial adjectives (*boring*) and Verb + Pronoun + to + V (I want her to visit my place) |
| Yeo (2002) | 20 to 30 minutes twice (once each week for two weeks) | no info (*n* = 90) | participial adjectives |

Appendix A continued

| | | | |
|---|---|---|---|
| Cho (2011) | 25 minutes in one session | no info (*n* = 117) | present perfect (for regret) |
| Lee (2003) | One session (reading session 5 mins long and interaction session 8 to 10 mins) | low to mid intermediate (*n* = 68) | words like *eyesight* and *adequate* |
| Jeong and Lee (2018) | Treatment session 1 (30 min) + Treatment session 2 (after 5 min. break / 30 min) | high intermediate (*n* = 66) | two adjectives and five nouns |
| Lim (2007) | two 50-minute classes a week for 3 weeks (total of 300 minutes) - 6 sessions over a three-week period | Intermediate (*n* = 62) | Present Perfect |
| Shin (2011) | Over 2 weeks | both low and high mixed (Based on high school midterm exam) (*n* = 90) | Conditional |
| Hwang (2018) | One class period | 250 to 320 TOEIC (low English Proficiency) (*n* = 122) | Verb tenses and relative clauses |
| Lee (2002) | One class period | low intermediate (*n* = 124) | Collocations such as: I had better do/at an American college / smart enough to / I had better do / turned to / in despair |
| Kim and Nam (2017) | All on same day / Pretest – Treatment of around 20 to 30 minutes | high intermediate or low advanced by self-report (*n* = 100) | 10 Idiomatic expressions like “shake a leg” and “hit the sack” |

Appendix B

# Number of Treatments, Treatment Types, and Assessment

Korean Studies

| Authors<br>Number of Groups –<br>Treatment Type – Effect Size | Assessment |
|---|---|
| Kang (2011)<br>1 – Input (Low proficiency input enhancement and consciousness raising) (*d* = 3.3120)<br>1 – Input (High proficiency input enhancement and consciousness raising) (*d* = .7846) | Dictogloss (written) |
| Kang (2003)<br>1 – Output (All output and production activities) (*d* = 1.6461)<br>1 – (Processing Instruction) (*d* = 1.5110) | Sentence Production Task (written) |
| Song and Suh (2008)<br>1 – Output (Reconstruction task) (*d* = .7611)<br>1 – Output (Picture-cued writing) (*d* = 1.0041) | Contextual sentence completion task (production) / TLU Analysis Used (written) |
| Yang (2008)<br>1 – Input (Textual Enhancement) (*d* = 3.1854)<br>1 – Output (Dictogloss) (*d* = 2.6561) | Production test where Korean asked for English response / ten minutes / (written) |
| Kim (2014)<br>1 – Input (60 target sentences with images and rule explanation in Korean) (*d* = 1.1846)<br>1 – Input (60 target sentences with images and rule explanation in Korean) (*d* = .7914)<br>1 – Input (Input Enhancement - 60 target sentences with images) (*d* = .7770) | Oral imitation test |
| Kang (2009)<br>1 – Input (negative feedback plus one session of rule presentation) (*d* = 2.0896)<br>1 – Input (Input Enhancement) (*d* = 1.3206) | Picture Description Task (written) |
| Song (2007)<br>1 – Input (input enhancement) (*d* = –.0654)<br>1 – Output (picture cued writing) (*d* = 1.0042) | Production test about 30 minutes / explained in Korean with prompt (written) |
| Yang (2004)<br>1 – Input (in-depth reading, comprehension questions, and feedback - student metatalk but not implicit treatment given) (*d* = 3.1854)<br>1 – Output (a modified dictogloss) (*d* = 2.6140) | Sentence production test / translate Korean sentences into English (written) |

Appendix B continued

| | |
|---|---|
| Kim (2002)<br>1 – Output (communicative tasks designed to elicit target form) (*d* = 1.2692)<br>1 – Input (Consciousness Raising - Rules of target forms / comprehension and checkup questions) (*d* = .2630) | Elicited oral production test and picture description |
| Yeo (2002)<br>1 – Input (Input Enhancement) (*d* = .6998)<br>1 – Output (Dictogloss) (*d* = 1.2229) | Open ended oral questions |
| Cho (2011)<br>1 – Input (input flood and input enhancement combined) (*d* = 1.0470)<br>1 – Output (Image and sentence completion) (*d* = 1.5110) | Fill-in-the-blank and open-ended questions (written) |
| Lee (2003)<br>2 – Input (input with target feature) (*d* = 1.1159)<br>2 – Output (production activities) (*d* = 1.0066) | Uses some comprehension measures but has a major production component (written) |
| Jeong and Lee (2018)<br>1 – Input (Input Enhancement with pictures – body part circled) (*d* = 3.7538) | Fill in the blank production test about parts of the body (written) |
| Lim (2007)<br>1 – Input (Input with visual) (*d* = 1.4194) | Blanks of one to up to three words to assess productive knowledge / 14 verbs from three aspectual categories and 20 distractors (written) |
| Shin (2011)<br>1 – Output (Guided Essay) (*d* = 1.0495)<br>1 – Output (Reconstruction task) (*d* = .6293)<br>1 – Input (input with target feature) (*d* = .2237) | Picture-Cued Production test and fill-in-the-blank (contextualized pictures)—8 targets and 4 distractors (written) |
| Hwang (2018)<br>3 – Input (Input Enhancement—narratives with bolded text) (*d* = 1.7438)<br>3 – Output (Audio-recorded narratives with dictogloss) (*d* = 2.2880) | Production with fill-in-the-blank and open-ended questions (written) |
| Lee (2002)<br>2 – Output (read 146-word story out loud and worked in groups to retell story) (*d* = 1.5252)<br>2 – Input (with picture prompts) (*d* = 1.7565) | Production through translation and sentence completion (5 minutes given) (written) |
| Kim and Nam (2017)<br>3 – Input (explicit information provided) (*d* = 3.5602)<br>2 – Input (with target feature) (*d* = 2.7907) | Production task with ten short contexts and blank for idiomatic expression (written) |

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/22
https://doi.org/10.31261/TAPSLA.16021

CC BY SA

**Petra Langerová**
Masaryk University, Czechia
https://orcid.org/0000-0002-9886-8865

# The Learning Styles of Technical Students and Their Role in Learning English as a Foreign Language

Abstract

The paper presents the results of mixed-methods research into the English language learning styles of a sample group of students from the Faculty of Electrical Engineering and Communication (FEEC) of the Brno University of Technology (BUT). In the quantitative phase of the research, students were questioned using an adapted Ehrman and Leaver Questionnaire (2003) with a bipolar scale of ten learning styles to identify commonly shared dimensions of learning styles of English among the sample group. The questionnaire also determined how flexible or rigid students were in their preferences and identified students with strong preferences for particular dimensions. The research then moved to a qualitative phase in which selected students who had stated a preference for commonly shared learning styles in the questionnaire were asked to participate in a semi-structured interview in which they discussed how their learning styles are reflected in their experiences of learning English and evaluated the success of their chosen approaches. Students were also asked about their willingness to change their learning styles if they had failed to make progress. The interviews offered valuable insight into students' learning preferences, with most students showing rigidity in their learning styles.

*Keywords*: language learning styles, learner autonomy, individual differences, styles stretching

## Background and Motivation of the Research

The study examines how students at a Czech technical university learn English in an effort to address some of the issues which the author identi-

fied during her personal experiences of teaching at the Faculty of Electrical Engineering (FEEC) of the Brno University of Technology (BUT). Regardless of their level of English, the author observed that many students exhibited high levels of communication apprehension and a general unwillingness to communicate in addition to low accuracy levels in terms of grammar and vocabulary usage. More significantly, however, many students showed little motivation to improve their language skills, possibly as a result of previous failed attempts to make progress. Students also suffered from a low capacity for autonomous learning, with many unable to analyse their mistakes and learn from them; many seemed to prefer a limited range of learning approaches and were reluctant to explore other possibilities. These observations motivated the author to explore learning preferences of students in more detail and to determine the effectiveness of their chosen methods.

The research presented here aims to reflect the findings in teaching methodology and to tailor pedagogical approaches towards this type of student accordingly. Technical students are generally required to excel in logic and show an intuitive grasp of theory and procedural processes, with study programs placing a greater focus on abstract thinking and analysis. Given this academic background, it would be expected that technical students would apply a similar approach to the task of learning languages the same way but, as the results of our research revealed, the opposite is the case; students prefer learning languages in an intuitive and subconscious manner without placing a particular focus on grammatical rules. According to the Scopus database, no research has been carried out into the issue of the language learning approaches of technical students in the Czech context, and this indicates that there is a considerable research gap in this field.

## Introduction to Learning Styles

Learning styles form an inherent part of the approach of individual learner differences. If learners have an awareness of the style of learning which suits them best; they can manage and guide their learning process more consciously and efficiently and can also be more motivated to study as they take more responsibility for what and how they learn. This approach of self-directing their study processes can help students to develop and strengthen their capacity for autonomous learning.

Learning styles have been the subject of considerable research over the course of several decades, but conclusions on their efficacy have been mixed. They were initially considered to be one of the key factors in determining

whether or not students were successful in acquiring new language skills, and were generally believed to be crucial in overcoming the false beginner phase and mastering a language. Proponents of this theory argued that if teachers found out how their students preferred to learn and they adjusted their teaching styles accordingly, the majority of students would benefit greatly, with the learning process becoming easier, faster, more enjoyable and more efficient, thereby ensuring tangible progress. However, much of the research behind this belief proved to be flawed in many cases, and subsequent studies on the issue split into a variety of different approaches with varying degrees of scientific rigour and reliability. The search for the "ideal" inventory of learning styles can perhaps be compared to the search for the "ideal method of teaching"; methods which had once been feted as the most suitable were soon superseded by newly developed models which were themselves replaced in turn by the latest attempt to find the "Holy Grail" of the ultimate methodology. After a period of harsh criticism (in many cases, deservedly so) in which the very concept of learning styles was brought into question due to the lack of thorough methodology and poor research results (see Pashler et al., 2008; Kirschner & van Merriënboer, 2013; or Geake, 2008), a more diligent approach to the topic was adopted which led to a cautious renewal of interest in learning styles. The pendulum of the unpopularity of learning styles had already reached its trough and more recent research appears to have regained some scientifically convincing confidence in openly admitting past flaws and genuinely searching for methods which emphasise reliability rather than novelty.

## Definition of the Term

A plethora of definitions for learning styles have been offered since the emergence of the term, with most covering the above-mentioned fields of research and focusing on differences related to the personalities and cognitive abilities of individual learners (Ehrman & Leaver, 2003), variables in learning and teaching (Kolb, 1976; Entwistle, 1981), or personal and social behaviour (Sternberg, 1996; Rayner & Riding, 1997; see also Rayner, 2015). Reid uses general terms to describe the concept as "an individual's natural, habitual and preferred way(s) of absorbing, processing and retaining new information and skills" (Reid, 1995, p. 8), while Dörnyei (2005) notes that "[learning styles] are not firmly fixed ways of behaviour, but just tendencies and preferences more or less strong, which can be modified and extended according to various tasks and situations" (Dörnyei, 2005, p. 158). As Dunn and Dunn stated, "[l]earning style is a biologically and developmentally imposed set of characteristics that make the same teaching method wonderful for some and terrible for others" (Dunn & Dunn, 1979, p. 3).

> Learning styles are not dichotomous (black or white, present or absent). Learning styles generally operate on a continuum or multiple, intersecting continua. For example, a person might be more extroverted than introverted, more closure-oriented than open, or equally visual and auditory but with lesser kinesthetic and tactile involvement. Few, if any, people could be classified as having all or nothing in any of these categories. (Ehrman, 1996, p. 115)

The wide-ranging nature of these definitions reflects the vagueness of the term and also indicates the difficulties involved in approaching the topic. It also suggests a possible reason why research into learning styles has split into several directions since the period of critical revisionism.

## Current State of Research

Rayner (2015) categorised research on learning styles into four types, and we can generally state that this categorization is still valid today. The first type examines personality-based styles, based on the premise that learning styles are closely related to personality traits; many current studies combine this approach with specific personality inventory tools, such as the Myers-Briggs Type Indicator (Myers-Briggs, 1978) or Big Five (Costa & McCrae, 1992; see Komarraju et al., 2011; Siddiquei & Khalid, 2018; Khamal & Radhakrishnan, 2019; or Abuzeid et al., 2021). The second type focuses on cognition-centred styles, a field which Rayner sees as relatively stable as such studies are related to mental processing, memory and differences in perception (see Ubuz & Aydynier, 2019; Wang & Sanchez, 2022; Miller et al., 2011). This paper follows this particular approach, as do many other authors focusing on the issue of technical students learning English (see Synekop, 2020 or Nikolaeva & Synekop, 2020). The third type is a less stable field, as it relates cognitive styles relating to tasks and changing contexts, an approach which is perhaps closer to the concept of learning strategies or patterns. Recent studies have linked this type of style to, among others, contextually-determined cultural background (see Lemke-Westcott & Johnson, 2013; Armstrong & Li, 2017). The final type in Rayner's categorization is that of learning preferences affected by modality and learning experiences (e.g., see Lodge, Hansen, & Cottrell, 2015).

The COVID pandemic radically transformed the entire education sector and forced teachers to work under new and unprecedented conditions. The need to switch almost overnight to online learning methods forced teachers to adopt new approaches. Materials and modules for e-learning materials and modules existed prior to the pandemic, but they were not widely used, nor were they considered as a primary source of teaching. As a result, e-learning materials were designed with a greater focus on individual learning styles or personality

types (or a combination of both) in order to compensate for the lack of personal contact and the more individual nature of online teaching, and this specific approach became the subject of a relatively new type of study which examines the efficiency and quality of such materials and online learning experiences in general (see Scott-Monkhouse, 2023; Samonte et al., 2023; Alzain, 2022).

### Research in the Czech Republic

There is a relative lack of research examining learning styles in the Czech context, but the studies published to date mostly address the design of e-learning materials for foreign language instruction in distance studies (e.g., Šimonová, 2013; Šimonová & Poulová, 2012), while others focus on learning styles in science subjects such as physics, accounting and natural sciences (see Zajacová, 2016; Berková et al., 2020; Malčik & Miklošíková, 2017). The Scopus database suggests that there are currently no other Czech studies that focus on learning styles intended to improve language acquisition among students of technically oriented subjects.

## Autonomy of Teachers and Students

Many different aspects seem to play a role in learning, and teachers should always acknowledge the fact that students are individuals with different abilities, skills, potential and motivation. Individual differences (ID) are thus a crucial consideration in formulating teaching strategies as they are explicitly reflected in the ways in which students learn.

"ID constructs refer to dimensions of enduring personal characteristics that are assumed to apply to everybody and on which people differ by degree" (Dörnyei, 2005, p. 4). "It is related to some main processes in the field of second language acquisition (SLA), and it has been researched extensively in L2 studies, making the area one of the most thoroughly studied psychological aspects of SLA" (Dörnyei, 2005, p. 6). According to Simsek (2012),

> individual differences can be defined as personal characteristics that distinguish learners from each other in the teaching and learning processes. Learners are unique individuals who bring a critical set of variables to each learning situation, including delicate traits as indicators of their potential and the history of achievement as signs of previous accomplishments and predictors of future performance. (Simsek, 2012, p. 98)

A recognition of the differences between students is important in ensuring that they develop the skills to master their chosen discipline and to understand the techniques which are most effective for them; the aim should be to allow students to gain autonomy and gradually develop into fully independent learners. Holec (2000, p. 48), one of the earliest advocates of autonomy in language teaching, has defined autonomy as "the ability to take charge of one's learning," while Dickinson (1987, cited in Gardner & Miller, 1996, p. 6) accepts the definition of autonomy as a "situation in which the learner is responsible for all of the decisions concerned with his or her learning and the implementation of those decisions." These definitions of autonomy generally agree on the involvement of learners in taking greater responsibility for what they learn, how they learn, and when they learn.

In terms of the teacher's role in helping their students to develop autonomy, Nemethová (2020) says that students should be permitted to work in their own way; teachers should facilitate this by offering them a range of activities and tasks from which they can choose, by encouraging them to ask questions, and ensuring that they feel confident about seeking out alternatives that best suit their approach while still meeting requirements.

This approach provides the perfect grounds for studying students' preferred learning styles. By giving students the opportunity to choose their favoured learning style, it encourages them to think about and reflect on how they prefer to learn; they are then offered a variety of ways to access their chosen approach and can then receive feedback on how effective and efficient it is for them. Teachers should also propose some suggestions for improvement or adjustments if the student's chosen method does not allow them to learn effectively the way they prefer, a technique known as style stretching (see Tuan, 2011).

Griffiths (2015) mentions several authors who have discussed style stretching and the importance of style flexibility. For example, Little and Singleton (1990) claim that learning styles can be adapted with experience and training, while Cohen and Dörnyei (2002) even recommend encouraging learners to stretch their learning styles and be introduced to approaches in learning that they would not normally use. Dörnyei (2005) also suggests that the more styles a student employs, the more effective they can become as a learner. Oxford (2011) states that even if a learner has strong style preferences, they can still be altered or modified. Wong and Nunan (2011) also relate style flexibility to learning effectivity, and Cohen (2012) even actively recommends training by stretching the comfort zone of learning styles.

The research in learning styles complies with the so-called postmethod pedagogy as defined by Kumaravadivelu (2008, p. 87) which she describes as a "three-dimensional system consisting of three pedagogic parameters: particularity, practicality and possibility." The Parameter of Particularity states that "any postmethod pedagogy must be sensitive to a particular group of teachers

teaching a particular group of learners pursuing a particular set of goals within a particular institutional context embedded in a particular sociocultural milieu" (Kumaravadivelu, 2001, p. 138). This approach correlates with the need to consider the individual differences between teachers themselves, with each possessing their own set of beliefs and teaching styles. Teachers' approaches should therefore be adapted according to the needs and preferences of their particular group of learners.

The Parameter of Practicality "relates broadly to the relationship between theory and practice, and narrowly to the teacher's skill in monitoring his or her teaching effectiveness" (Kumaravadivelu, 2008, p. 172), thereby suggesting that teachers should apply their theoretical knowledge (or parts thereof) in practice according to the needs and capacities of specific groups of learners. This approach requires teachers to reflect upon themselves and their approaches, while constantly evaluating their teaching performance.

The Parameter of Possibility is also related to the awareness that learning a foreign language also involves students acquiring a different learner identity through their exposure to language ideologies. This is a phenomenon which teachers cannot ignore, and they must be able to react to possible problems and obstacles which students may face throughout the learning process. Kumaravadivelu (2008, p. 175) says that "language education provides its participants with challenges and opportunities for a continual quest for subjectivity and self-identity." Moreover, subjectivity, individual differences and personality are deeply rooted and reflected in an individual's preferred ways of learning, all of which can help an individual to create a foreign language identity through their chosen method of picking up elements of a language, whether conscious or unconscious.

## Research Objectives

The aim of the mixed-methods research into the English language learning styles of FEEC BUT students was to identify any similarities in the learning styles of the surveyed students. In the quantitative phase of the research, the Ehrman and Leaver Questionnaire (2003) was applied to identify commonly preferred dimensions of learning styles of English among technical students and to explore how flexible or rigid these preferences were or if any students exhibited strong preferences for particular approaches. In the subsequent qualitative phase of the study, students were invited to participate in interviews where they discussed the individual features of their English learning styles, including relevant examples, and assess the success of their chosen style. Another

question examined their willingness to change their learning styles if they had failed to make progress and their reasons for doing so. A semi-structured questionnaire was used for the interviews, with students only discussing the dimensions for which they had expressed a strong preference.

**Research Hypotheses**

**H1.** A high number of FEEC BUT students share some common preferences for learning styles.

**H2.** Students demonstrate flexibility in the case of at least two dimensions of learning styles.

**Research Questions**

**Q1.** Which dimensions of learning styles are most commonly preferred?

**Q2.** Are there any dimensions of learning styles in which students are particularly flexible or rigid? If so, what are they?

**Q3**. What are the reasons for the willingness or unwillingness to change their learning styles?

## Participants

The research group consisted of a sample group of FEEC BUT students, approximately half of whom were enrolled in bachelor's and the other half in master's study programs. The method of convenience sampling was applied. The majority of participants were males who ranged from 20 to 24 years of age. First-year students were not included in the research due to the high rate of student dropout after the first year of their studies, with this factor making the subsequent interviews impossible.

## Research Tools

### Ehrman and Leaver Construct (2003)

The questionnaire is based on the psychological personality typology determined by the Myers-Briggs Type Indicator questionnaire (1956) which enabled us to gain a fuller understanding of the participants' preferred dimensions of learning styles. The Ehrman and Leaver Construct (2003) (hereinafter referred to as the E+L questionnaire) assesses learning styles on a continuum between synoptic and ectenic poles. The synoptic pole relates to the subcon-

scious and implicit processing of information and is likely to perceive phenomena as wholes, unlike those reflecting the ectenic one, which prefer conscious control over the learning process and show a tendency to perceive phenomena as composites.

The questionnaire consists of sixty statements, thirty of which evaluate each of the two poles. Each pole is subdivided into ten dimensions, each of which is queried by three statements in the questionnaire. These statements are presented in a contrasting manner, and respondents select the statement which best reflects their thinking using a 9-point Likert scale ranging from a mild preference to a strong need.

The ten bipolar cognitive dimensions are as follows: field dependent vs. field independent, field sensitive vs. field insensitive, random vs. sequential, global vs. particular, inductive vs. deductive, synthetic vs. analytic, analogue vs. digital, concrete vs. abstract, levelling vs. sharpening, and impulsive vs. reflective.

The model also reflects the following theory of brain hemisphere functions developed by Seikel (2018):

The right hemisphere has the following functions:

– parallel processing,
– synthesis,
– providing linguistic context,
– perceiving data as a whole,
– comprehending and interpreting emotions.

The left hemisphere is responsible for the following functions:

– language reception,
– structure of language, mainly syntax, morphology and phonology,
– serial information processing,
– judgement,
– processing details,
– analysis and categorisation,
– memory-related operations.

Ehrman and Leaver, who developed the Construct, state that very few people prefer only one pole exclusively, with most people switching fluently from one pole to another and, by extension, moving between implicit and explicit learning. The implicit learning techniques of the synoptic pole are mostly governed by the right hemisphere, while explicit learning styles associated with the ectenic pole are dominated by the left hemisphere.

### Interview Questionnaire Based on the Ehrman and Leaver Construct (2003)

The questionnaire contained only those statements from the E+L questionnaire which referred to the dimensions for which students had shown a preference, more specifically the following questions:

1. What came to your mind when you read the questionnaire statements? Can you give any concrete examples of how you learn?
2. Would you find it difficult to learn in the styles indicated in the opposing statement in the questionnaire?
3. Did you experience learning success with your chosen style? Can you recall any difficulties you experienced using this style?
4. Would you be willing to change your style of learning if it offered the possibility of making progress? What would be the biggest obstacle for you in doing so?

## Analysis

### Quantitative Phase

The Czech version of the E+L questionnaire was successfully tested on a group of 150 students in 2020, but the responses obtained in this preliminary study are not included in the current research as it was only intended as a pilot study.

The quantitative phase was carried out in 2022, with a total of 400 participants filling out an online version of the Ehrman and Leaver questionnaire as part of their English classes. The questionnaire was adapted from the E+L questionnaire and featured 30 statements to which participants responded on a 10-grade Likert scale, with answers 1, 2, 3 and 7, 8, 9 showing their strong preference for either of the two poles and answer 5 meaning "I can do both well." For the purposes of the research, an additional answer 10 was added which means "I don't know, I cannot identify with either of the two statements" with the aim of identifying flexible students who are comfortable with both of the suggested approaches. Only 250 of the 400 questionnaires were selected for analysis, with any incomplete questionnaires or those featuring responses with answer 10 being excluded. The data was analysed as follows: the average score in each of the dimensions was calculated by summing the responses in the three questions representing each dimension. If the average score was lower than or equal to 4, the style was ranked on the left pole, that is, reflect-

ing synoptic dimensions. If the average score was greater than or equal to 6, the style was ranked on the right pole, that is, reflecting ectenic dimensions. Average scores ranging from 4.1 to 5.9 showed flexibility in the respective style. The style in a particular dimension was considered flexible if the percentage of flexibility among all students in the given dimension was greater than 30%. The full results are shown in Table 1. Mean, standard deviation and upper and lower quartiles were also calculated, and these are shown in Table 2.

**Table 1**

*Percentages of Responses in Individual Dimensions (N = 409)*

| Dimension | Average score | | |
|---|---|---|---|
| | avg. ≤ 4 % | avg. ≥6 % | avg. 4.1–5.9 % |
| D1 | | | |
| Field sensitive | 65 | | |
| Field insensitive | | 15 | |
| Flexible | | | 20 |
| D2 | | | |
| Field dependent | 27 | | |
| Field independent | | 26 | |
| Flexible | | | 47 |
| D3 | | | |
| Leveling | 36 | | |
| Sharpening | | 27 | |
| Flexible | | | 37 |
| D4 | | | |
| Global | 66 | | |
| Particular | | 8 | |
| Flexible | | | 26 |
| D5 | | | |
| Impulsive | 30 | | |
| Reflective | | 40 | |
| Flexible | | | 30 |

*Table 1 continued*

| Dimension | Average score | | |
|---|---|---|---|
| | avg. ≤ 4 % | avg. ≥6 % | avg. 4.1–5.9 % |
| D6 | | | |
| Synthetic | 42 | | |
| Analytic | | 28 | |
| Flexible | | | 30 |
| D7 | | | |
| Analogue | 25 | | |
| Digital | | 37 | |
| Flexible | | | 38 |
| D8 | | | |
| Concrete | 78 | | |
| Abstract | | 5 | |
| Flexible | | | 17 |
| D9 | | | |
| Random | 25 | | |
| Sequential | | 47 | |
| Flexible | | | 28 |
| D10 | | | |
| Inductive | 27 | | |
| Deductive | | 33 | |
| Flexible | | | 40 |

**Table 2**

*Means and Standard Deviations of the Likert-scale Responses*

| Dimension | Basic characteristics of the dataset | | | | |
|---|---|---|---|---|---|
| | Mean | SD | UQ | Median | LQ |
| D1 | 3.56 | 1.97 | 4.67 | 3 | 2 |
| D2 | 4.91 | 1.30 | 6 | 5 | 4 |
| D3 | 4.87 | 1.56 | 6 | 4.67 | 4 |
| D4 | 3.58 | 1.51 | 4.67 | 3.50 | 2.33 |

*Table 2 continued*

| Dimension | Basic characteristics of the dataset | | | | |
|---|---|---|---|---|---|
| | Mean | SD | UQ | Median | LQ |
| D5 | 5.24 | 1.92 | 6.67 | 5 | 3.67 |
| D6 | 4.70 | 1.76 | 6 | 4.67 | 3.33 |
| D7 | 5.32 | 1.64 | 6.33 | 5.33 | 4 |
| D8 | 3.23 | 1.44 | 4 | 3 | 2.33 |
| D9 | 5.56 | 1.88 | 7 | 5.67 | 4 |
| D10 | 5.14 | 1.66 | 6.33 | 5 | 4 |

### *Quantitative Phase Results*

The analysis revealed that there were five preferred learning styles, three of which lay on the synoptic pole and two on the ectenic pole. The synoptic learning styles, namely field sensitive, global and concrete approaches, were found to be more popular than the two ectenic learning styles, the reflective and sequential approaches. The analysis confirmed H1, as the results show that a high percentage of students favoured each of the preferred dimensions. The field-sensitive learning style was favoured by 65% of the students, the global learning style by 66%, the concrete learning style by 78%, the reflective learning style by 40% and the sequential learning style by 47%. The results also showed flexibility in four dimensions, namely field dependence/independence, levelling/sharpening, analogue/digital, and inductive/deductive, a finding which confirms H2.

## Qualitative Phase

The qualitative phase consisted of interviews conducted with ten students who had participated in the questionnaire phase. A total of 36 students of the 250 respondents were selected for interviews based on the combination of their learning styles, but only ten of them were willing to participate. The interviews were conducted online using MS Teams software in order to allow students to feel comfortable while being interviewed. The interviews were intended to be diagnostic, with the interviewers applying the techniques of paraphrasing, interpretation and rephrasing. The interviewed students showed a very strong preference for at least two of the commonly shared learning styles dimensions, and as a consequence, the number of students who provided answers to individual dimensions differed. In summary, the following dimensions were discussed by the interviewed students: the field sensitivity dimension (by ten students), the

global learning style (six students), the concrete learning style (eight students), the reflective learning style (one student), and the sequential learning style (two students) (see Table 3).

**Table 3**

*Table of Students Selected for Interview Based on Their Preferred Learning Styles*

| Learning style | | Students' preferences | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | S1 | S2 | S3 | S4 | S5 | S6 | S7 | S8 | S9 | S10 |
| Field sensitive | 10 | x | x | x | x | x | x | x | x | x | x |
| Global | 6 | | x | x | x | x | | | | x | x |
| Concrete | 8 | x | x | x | x | x | x | x | x | | |
| Sequential | 2 | | | | | | x | | | x | |
| Reflective | 1 | | | | | | | | | | x |

## *Qualitative Phase Results*

In this section, we will examine the styles which were discussed by the interviewees in more detail:

- **field-sensitive (synoptic)**

A field-sensitive person can be described as someone who prefers to work with new material in context, such as in stories and articles or, at the very least, in sentences. They often pick up new words or ideas in a haphazard manner, without planning in advance. Students mentioned listening as the main input channel in this style, with some stating that they managed to pick up new words, phrases, and grammatical structures without actually realizing it ("I can recognize a mistake because it sounds strange," "I heard it said this way").

Most students claimed that they learned successfully using this approach; they said that they remembered the acquired structures later on, but that this was conditioned by the frequency at which they were exposed to it or by the context or situation in which they encountered the structures. However, students were not capable of learning some language skills in this style, in particular those of grammar, accuracy of translation and meaning and irregular verb forms. When asked about their willingness to change their way of learning, few students said that they would be willing to change, but most of them also reported that they had already tried and it had not worked for various reasons (see Table 4).

Their answers thus perfectly matched the description of the dimension in the questionnaire manual; the results demonstrate that most are rigid in their preferred way of learning, offering little or no space for change. In terms of the pedagogical implications arising from the students' answers, we can see that other ways of teaching are necessary in order to improve their accuracy and precision in meaning.

**Table 4**

*Characteristics of Learning in Field Sensitivity Dimension*

| Dimension Field Sensitive (*N* = 10) | | | | |
|---|---|---|---|---|
| Learning by this style | Drawbacks | Learning in an opposite style | Flexibility | Pedagogical implications |
| Context: mostly audio, less frequently written. Acquired items: words, phrases, gram. structures. Success in learning such structures conditioned by the frequency of exposure. | Inability to learn grammar correctly. Low accuracy in translation and lexical production. | Tried drill, memorising vocabulary and gram. rules in elementary and secondary education, without success. Hate memorising tend to forget memorized structures soon. Never really understood grammar from gram. rules. | Rigid | Alternative ways of teaching accuracy in meaning and grammar. |

- **global (synoptic)**

People who learn globally tend to adopt a "big picture" perspective; seeing the "forest rather than the trees," they typically start with the main points and only come to the details at a later point. Students who favoured this approach described understanding the meaning and the plot of a text/video first; only then could they focus on details, such as form and usage.

Learners using this approach also stated that they had difficulty learning details this way and that they struggled to discern differences in the meanings of words. When asked about their willingness to learn in an opposite way, most students replied that it would be possible, but they would find it unnatural and would have to skip a lot of information in order to see how things are connected into a meaningful unit (see Table 5 for more details).

The pedagogical implications of these views are clear. There is no need to force such learners to adopt an opposite way of learning, since context as a determiner of meaning is their primary means of understanding and learning, and a subsequent focus on the details is a frequent and normal practice in English lessons.

**Table 5**

*Characteristics of Learning in Global Dimension*

| Dimension Global (*N* = 6) | | | | |
|---|---|---|---|---|
| Learning by this style | Drawbacks | Learning in an opposite style | Flexibility | Pedagogical implications |
| Learning tenses: first they need to understand the context, only then can they identify and concentrate on grammatical forms.<br>Text, video: first story, only after repeated input do they notice details, grammar, vocabulary.<br>New word: first meaning, then form.<br>Success in learning. | Difficulty in finding and understanding details, differences in meaning.<br>Difficulty to translate a sentence without context.<br>Difficulty to remember shades of lexical meaning and exceptions in grammatical rules. | Probably possible, but unnatural.<br>Few students tried. | Rigid | Focus on context as the determiner of meaning and draw attention to details and their form. |

- **concrete (synoptic)**

People favouring this style need to interact with the world directly in order to acquire new knowledge and learn through practical application, especially if they can touch, see or hear the new information. The students offered various examples of learning through the manipulation of objects and by using their senses for both concentrating and relieving stress (see Table 6 for more details).

However, students stated that they were unable to learn theory using this style. They have difficulty memorising things, so they write cheat notes and highlight important information, with this approach allowing them to subsequently recall the location of the important information visually.

Learning by any other method would likely be impossible for such students as they would miss the application of the theory and, as was noted above, they have difficulty learning information by heart. Some students also mentioned that they are unable to focus on form alone and that they would miss seeing the continuity with the previous knowledge they have gained.

It can therefore be concluded that students preferring this style tend to be quite rigid in their preferences and that it would not be possible to shift their abilities to the other pole.

**Table 6**

*Characteristics of Learning in Concrete Dimension*

| Dimension Concrete (*N* = 8) | | | | |
|---|---|---|---|---|
| Learning by this style | Drawbacks | Learning in an opposite style | Flexibility | Pedagogical implications |
| Learning through manipulating things, trial and error, programming, experimenting, testing equipment before reading a manual.<br>Employing senses:<br>Listening: music (to create context, shut the outer world).<br>Haptic and kinaesthetic experience: walking, changing postures, squeezing a rubber, fiddling with a pen (to concentrate better, to cope with stress).<br>Inductive learning (first try out, then deduce theory).<br>Success in learning. | Cannot learn theory, difficulty memorising things. | Impossible, they would miss the application.<br>Strategies to learn theory: writing cheaters, highlighting important information. | Rigid | No need to change it. |

- **sequential (ectenic)**

Students favouring this approach learn most effectively when they are provided with a sequence of steps that they can follow such as textbooks and lesson plans. Many students at FEEC BUT appreciate learning approaches that offer systematized and hierarchized information and the logical structure of a textbook. If they are forced to learn in an opposite manner, that is, by developing their own system, they find it difficult to differentiate between important and unimportant information. Nevertheless, if they were certain that this approach could offer better results, these students would adopt the opposite style even if took them more time to learn (see Table 7).

Thus, it can be concluded that students favouring this approach are flexible and that there are some grounds for persuading them to try another learning style. Students who have the opportunity to organise things in their own way employ creativity which can help them to remember new information more effectively.

**Table 7**

*Characteristics of Learning in Sequential Dimension*

| Dimension Sequential (*N* = 2) | | | | |
|---|---|---|---|---|
| Learning by this style | Drawbacks | Learning in an opposite style | Flexibility | Pedagogical implications |
| Appreciate structured chapters in textbooks. Hierarchy, order and system in new information. Topics linked logically one to another. | Difficulty in distinguishing importance of information, in prioritizing information on their own. | Difficult, but feasible. | Flexible | Enhancing students´ creativity. Encouraging them to reorganize information in their own way. |

- **reflective (ectenic)**

This style can be characterised by the need to think things through before applying the newly acquired skills or information. Only one student mentioned this learning style in the interviews (see Table 8), stating that he feels nervous or uncomfortable when asked to answer a question without preparation and would be far happier if he had some time to think before giving an answer. When asked to consider learning in an opposite way, such as by studying in advance in the comfort of his home, he said that this would be feasible for him; he also predicted that it would make him feel more confident and secure and more willing to react and respond more quickly.

This suggests that students would benefit from dedicating time to home preparation rather than just following their teachers' instructions. The approach would also help students feel more confident in learning and to react faster and more spontaneously in class.

**Table 8**

*Characteristics of Learning in Reflective Dimension*

| Dimension Reflective (*N* = 1) | | | | |
|---|---|---|---|---|
| Learning by this style | Drawbacks | Learning in an opposite style | Flexibility | Pedagogical implications |
| Need time to think about the answer in order to form a reply confidently. | Slow reactions, feelings of uncertainty and uneasiness if expected to react fast and spontaneously. | Faster reactions would require a familiarity with the topic and home preparation. | Flexible | Enhancing home preparation and allowing enough time for preparation in class. |

## Conclusion

The findings of the research have offered some valuable information on the ways in which students learn English. The results of the quantitative research show that the vast majority of students (65%, 66% and 78% in field-sensitive, global and concrete learning styles, respectively) prefer to learn implicitly and subconsciously, without paying too much attention to form and accuracy. Their approach to learning languages can be described as somewhat unsystematic, with students picking up elements of language in a random and osmotic manner based on the immediate content they came across. Students also showed a preference for learning through practical application prior to studying the theory, even if this approach requires a considerable amount of trial and error. Learning through the senses is an essential part of students' learning experience, with many expressing enthusiasm for hands-on experience. Many also stated their need to see the "big picture" before focusing on details. The results also suggest that students are not very spontaneous in their reactions (40% of the students favoured a reflective learning style), as they prefer to have some time to think about their answers in advance and to consider systematically prepared materials.

The more detailed qualitative data about students' learning styles was based on a relatively small sample of ten students who participated in the interviews, and the low participation of students in this element of the study is perhaps the biggest limitation of the research. Nevertheless, the interviews provided interesting insights into individuals' ways of learning. Although the small sample size makes it difficult to draw generalised conclusions on students' approaches, several ideas emerged in the discussions on individual learning styles. The strong preference for field sensitivity noted by students can result in low accuracy and precision in grammar and a lack of understanding of the ambiguities of meaning in vocabulary. Similarly, a strong preference for learning globally can imply a limited ability to focus on details before fully understanding the context and scanning for a particular piece of information. This approach often requires a longer time because students must first grasp the whole meaning before moving on to distinguish the components of a statement. A high preference for learning concretely suggests that students struggle with memorising information, and this could also imply they may have problems with concentration, with six out of eight interviewees stating that they need to use their physical senses to stave off their discomfort before memorising theory. A preference for sequential learning styles may also indicate a low willingness to be creative and develop their own systems, but this is something that can be learnt in order to make progress. Lastly, the preference for reflective styles noted by many students may imply a low confidence in their own abilities, but this could

be strengthened or even avoided with careful home preparation. In conclusion, the qualitative phase results brought a number of interesting findings which should be considered and reflected upon in the formulation of new pedagogical approaches tailored to this specific group of learners.

The research also explored the similarities which students share in their learning styles and preferences and the way in which these styles influence their success or failure in learning different language skills and abilities. Once again, while the quantitative section covered a relatively wide range of student preferences for learning, the small sample involved in the qualitative part makes it hard to draw any substantiated conclusions. Nonetheless, the interviews still suggested various implications of the reasons why students made progress or failed to do so and also raised questions about the ways in which students can assess their language skills to help them learn more effectively. Teachers can play a key role here by helping students analyse what approach may be beneficial to them and identify potential problems and obstacles they might face if they use the same learning styles in all tasks. It would be reasonable to discuss the benefits and the drawbacks of individual learning styles with students and to show them the importance of using a variety of learning styles to acquire a wider range of skills. This is where the so-called stretching of styles comes into play; a greater willingness to step out of one's comfort zone and adapt to new demands might turn out to be one of the key personality features determining the success or failure of achieving one's goals, but this is a conjecture which would require further research.

## References

Abouzeid, E., Fouad, S., Wasfy, N. F., Alkhadragy, R., Hefny, M., & Kamal, D. (2021). Influence of personality traits and learning styles on undergraduate medical students' academic achievement. *Advances in Medical Education and Practice, 21*, 769–777.

Alzain, E. (2022). Online EFL learning experience in Saudi universities during COVID-19 pandemic. *International Journal of English Language and Literature Studies, 11*(3), 109–125. https://doi.org/10.55493/5019.v11i3.4597

Armstrong, S. J., & Li, Y. (2017). A study of Anglo expatriate managers' learning, knowledge acquisition, and adjustment in multinational companies in China. *Academy of Management Learning and Education, 16*, 1–22.

Berková, K., Borůvková, J., Frendlovská, D., & Krpálek, P. (2020). Learning styles of Czech college and university students in accounting education. *Universal Journal of Education Research, 8*(8), 3459–3468.

Cohen, A. (2012). Strategies: The interface of styles, strategies and motivation on tasks. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Language learning psychology: Research, theory and pedagogy* (pp. 136–150). Palgrave Macmillan.

Cohen, A., & Dörnyei, Z. (2002). Focus on the language learner: Motivation, styles and strategies. In N. Schmitt (Ed.), *An introduction to applied linguistics* (pp. 170–190). Edward Arnold.

Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders, 6*(4), 343–355.

Dickinson, L. (1996). *Autonomy and motivation: A literature review*. Gardner & Miller.

Dörnyei, Z. (2005). *The psychology of the language learner: Individual differences in second language acquisition*. Taylor & Francis Inc.

Dunn, R., Dunn, K., & Price, G. (1979). *Learning style inventory.* Price Systems.

Ehrman, M. (1996). *Second language learning difficulties: Looking beneath the surface*. Sage.

Ehrman, M. E., & Leaver, B. L. (2003). Cognitive styles in the service of language learning. *System, 31*, 391–415.

Entwistle, N. J. (1981). *Styles of teaching and learning: An integrated outline of educational psychology for students, teachers and lecturers*. Wiley.

Geake, J. (2008). Neuromythologies in education. *Educational Research, 50*, 123–133.

Griffiths, C., & İnceçay, G. (2016). Styles and style-stretching: How are they related to successful learning? *Journal of Psycholinguistic Research, 45*, 599–613. https://doi.org/10.1007/s10936-015-9366-2

Griffiths, C., & Soruç, A. (2020). *Individual differences in language learning.* Palgrave Macmillan. https://doi.org/10.1007/978-3-030-52900-0_7

Holec, H. (2001). *Autonomy in foreign language learning.* Pergamon.

Kamal, A., & Radhakrishnan, S. (2019). Individual learning preferences based on personality traits in an e-learning scenario. *Education and Information Technologies, 1*(24), 407–435.

Kirschner, P. A., & van Merriënboer, J. J. G. (2013). Do learners really know best? Urban legends in education. *Educational Psychologist, 48*, 169–183.

Kolb, D. A. (1976). *Learning style inventory: Technical manual*. Prentice Hall.

Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences, 51*(4), 472–477.

Kumaravadivelu, B. (2006). *Understanding language teaching: From method to postmethod.* Routledge.

Lemke-Westcott, T., & Johnson, B. (2013). When culture and learning styles matter: A Canadian university with Middle-Eastern students. *Journal of Research in International Education*, *12*(1), 66–84.

Little, D., & Singleton, D. (1990). Cognitive style and learning approach. In R. Duda & P. Riley (Eds.), *Learning styles* (pp. 11–19). University of Nancy.

Lodge, J. M., Hansen, L., & Cottrell, D. (2016). Modality preference and learning style theories: Rethinking the role of sensory modality in learning, *Learning: Research and Practice, 2*(1), 4–17.

Malcik, M., & Miklosikova, M. (2017). Learning styles of students as a factor affecting pedagogical activities of a university teacher. *International Journal of Emerging Technologies in Learning (iJET)*, *12*(02), 210–218.

Miller, M. B., Donovan, C. L., Bennett, C. M., Aminoff, E. M., & Mayer, R. E. (2011). Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage, 59*(1), 83–93.

Myers, I. B., & Briggs, K. (1976). *The Myers-Briggs Type Indicator, Farm G*. Consulting Psychologists Press.

Némethová, I. (2020). Autonomy and motivation in language learning and teaching. *MUNI Journal CASALC, 10*(1). https://doi.org/10.5817/CASALC2020-1-12

Nikolaeva, S., & Synekop, O. (2020). Motivational aspect in student's language learning style in differentiated instruction of English for Specific Purposes. *Revista romaneasca pentru educatie multidimensionala—Journal for Multidimensional Education, 12*(2), 169–182.

Oxford, R. (2011). *Teaching and researching language learning strategies*. Pearson Longman.

Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2008). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest 9,* 105–119.

Rayner, S. G. (2015). Cognitive styles and learning styles. In J. D. Wright (Ed.), *International Encyclopedia of the Social and Behavioral Sciences*, Second Edition, pp. 110–117.

Rayner, S., & Riding, R. (1997). Towards a categorisation of cognitive styles and learning styles. *Education Psychology, 17*, 5–28.

Reid, J. M. (1995). *Learning styles in the ESL/EF classroom*. Heinle and Heinle Publishers.

Reinert, H. (1976). One picture is worth a thousand words? Not necessarily. *Modern Language Journal*, *60*, 160–168.

Richardson, J. (2011). Approaches to studying, conceptions of learning and learning styles in higher education. *Learning and Individual Differences, 21*(3), 288–293.

Samonte, M. J., Acuña, G. E. O., Alvarez, L. A. Z. & Miraflores, J. M. (2023). A personality-based virtual tutor for adaptive online learning system. *International Journal of Information and Education Technology, 13*(6), 889–905.

Scott-Monkhouse, A. (2023). From face-to-face tuition to online classes: 'Re-styling' a course of English for academic purposes. *Language Learning in Higher Education, 13*(2), 497–516. https://doi.org/10.1515/cercles-2023-2026

Seikel, J. (2018). An attentional view of right hemisphere dysfunction. *Clinical Archives of Communication Disorders, 3*, 76–88.

Siddiquei, N., & Khalid, R. (2018). The relationship between personality traits, learning styles and academic performance of e-learners. *Open Praxis, 10*(3), 249–263.

Simsek, A. (2012). Individual differences. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning.* Springer.

Sternberg, R. J. (1996). Styles of thinking. In B. P. Baltes & U. M. Staudinger (Eds.), *Interactive minds* (pp. 347–365). Cambridge University Press.

Synekop, O. (2020). Regulation aspect of learners' language learning style in differentiated ESP instruction. *Advanced Education, 15*, 25–31.

Šimonová, I. (2013). Learning styles in foreign language instruction. In *Proceedings of the International Conference on Interactive Collaborative Learning*, Kazan, Russia, 180–189.

Šimonová, I., & Poulová, P. (2012). Students' learning preferences in foreign language instruction: Comparative research. In *Proceedings of the 2012 IEEE Global Engineering Education Conference (EDUCON)* (pp. 1–7). Marrakech, Morocco, pp. 1–7.

Tuan, L. (2011). Matching and stretching learners' learning styles. *Journal of Language Teaching and Research*, *2*(2), 285–294.

Ubuz, B., & Aydınyer, Y. (2019). Project-based geometry learning: Knowledge and attitude of field-dependent/independent cognitive style students. *The Journal of Educational Research, 112*(3), 285–300.

Wang, A., & Pellicer-Sánchez, A. (2022). Incidental vocabulary learning from bilingual subtitled viewing: An eye-tracking study. In *Language Learning: A Journal of Research in Language Studies, 72*(3), 765–805.

Wong, L., & Nunan, D. (2011). The learning styles and strategies of effective language learners. *System*, *39*(2), 144–163.

Zajacová, B. (2016). Research of learning styles—Getting to know students' individualities. *Anthropologist, 2*(1), 35–42.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/22
https://doi.org/10.31261/TAPSLA.15946

CC BY SA

**Zdeňka Neumanová**
https://orcid.org/0000-0002-1234-4885
Department of English and American Studies
Masaryk University, Brno, Czechia

# An Investigation of Complexity, Accuracy, and Fluency in the Speech of EFL Learners

Abstract

This study investigates the relationship between several quantitative measures of L2 speech complexity, accuracy, and fluency (CAF) by comparing the oral productions of English L2 learners at different proficiency levels. Forty English as a Foreign Language (EFL) learners of varied proficiency levels performed a simple oral task. The performance of the L2 learners was analyzed regarding linguistic complexity, accuracy, and fluency to investigate the interplay of the CAF dimensions and how language proficiency levels (A2, B1, and B2) relate to these dimensions in the speech of EFL learners.

This study confirms the prevailing view that the three CAF dimensions are interconnected and that the CAF triad is a useful and valid way to investigate and describe L2 performance development. The results indicate that all three dimensions of CAF significantly predict L2 learners' oral proficiency, that is, from the developmental perspective, L2 learners' output is produced with higher complexity, accuracy, and fluency as they progress to a higher proficiency level. Moreover, the current paper discusses ways to measure CAF. Specific measures of complexity (index of developmental levels), accuracy (syntactic error rate), and fluency (articulation rate, and dysfluency rate) were identified as effective discriminators between proficiency levels.

*Keywords*: accuracy, complexity, fluency, L2, proficiency

During the course of the current study, the focus will be on three proficiency dimensions: complexity, accuracy, and fluency (CAF), and their interplay at the group level. The purpose of this study is to investigate how complexity, accuracy, and fluency interact in English learners' second language (henceforth referred to as L2) development. Motivated by one of the principal questions in second language acquisition (SLA) research relating to the nature of linguistic changes taking place in the L2 system of the learners as they become more

proficient, the current study investigates the linguistic dimensions underlying complexity, accuracy, and fluency in the context of L2 speech produced by 40 Czech-speaking learners of English. It aims to resolve the multifaceted processes that underlie language proficiency, contributing to its broader understanding in the L2 context. L2 speech production samples were analyzed as manifestations of what Czech EFL learners are/are not capable of doing in their development, and provide a clear picture of the underlying L2 systems, which cannot be studied directly. By and large, empirical evidence gathered for the current study is relevant from a descriptive point of view and can contribute to the formulation of explanatory hypotheses regarding L2 development.

Moreover, being able to objectively measure progress, the study will examine the relationship between language proficiency levels and complexity, accuracy, and fluency and inform language teaching practices and curriculum development.

A methodological challenge in CAF research is, however, the appropriate selection of CAF measures. Michael (2017) suggests that researchers use some of the measures employed in previous studies to ensure comparability with previous findings and that these measures are supplemented by context-specific measures that take into account the specific characteristics of the research context. Therefore, the current study also aims to examine which specific measures of complexity, accuracy, and fluency are best able to discriminate between proficiency levels.

Furthermore, even though I accept the status of complexity, accuracy, and fluency as distinct dimensions of L2 performance and proficiency, I do not exclude the fact that they can be interconnected and that they may somehow interact in L2 production. Moreover, if we only examine the dimensions one by one, we miss their interaction and the fact that the way they interact changes with time as well (Larsen-Freeman, 2009, p. 582). For this reason, the present explanatory study also aims to broaden our understanding of the interplay of complexity, accuracy, and fluency in the speech of young adults studying at a university.

In this article, I first provide a theoretical and methodological overview of previous research and the CAF framework. This is followed by a description of research procedures and a presentation of the findings. Next, the results of the current research are discussed. The article concludes by highlighting the underlying relationship between the CAF dimensions and L2 proficiency and outlining future directions for research.

To address the aforementioned conceptual and methodological challenges, the present study addresses the following research questions:

RQ1: Can all three dimensions—complexity, accuracy, and fluency—significantly predict L2 learners' oral proficiency?

RQ2: Which measures are best able to discriminate between proficiency levels?

RQ3: What are the interactions between complexity, accuracy, and fluency?

## Theoretical Part

In this section, I first present the individual aspects that have traditionally been described, albeit to varying degrees, in the literature. I then turn to the more recently uncovered aspects of the CAF triad.

### Proficiency

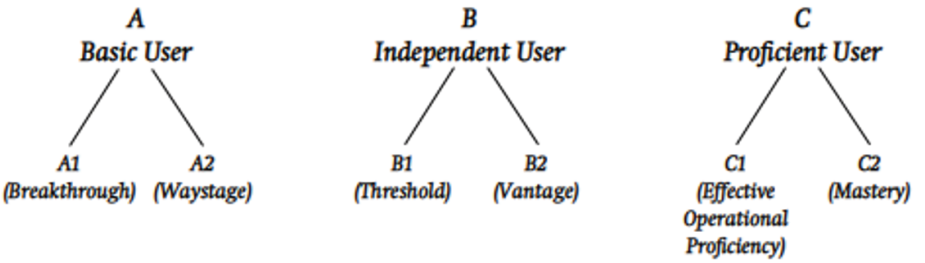
In the fields of SLA and foreign language learning (FLL), measuring L2 proficiency, defined as the ability to draw on and use competence in different tasks (Taylor, 1988, p. 166; R. Ellis, 2008, p. 976), is one of the key concerns, where the three dimensions of the CAF triad are widely recognized as the key components. Ortega (2014, p. 197) argues that "mastery and accuracy, that is, teleological arriving to isomorphic conformity with idealized native speaker norms," is the only thinkable way of defining linguistic and developmental success.

Proficiency has been interpreted in various ways in L2 research (for a discussion of how to define L2 proficiency, see Hulstijn, 2011). It is usually defined rather implicitly and its operationalisation differs across studies (Ortega, 2003). Studies have relied on institutional criteria which serve as a guarantee of EFL learners' proficiency (Gráf, 2015), different exam types, such as the Key English Test (KET) (De Felice & Pulman, 2009), or operationalised proficiency in terms of CEFR levels (Gyllstad et al., 2014).

The Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2001a) is an established benchmark for language competence (Jones & Saville, 2009). It comprises six levels of competence (A1 to C2, as shown in Figure 1), which have become a common currency in language education, prevalent in curricula, syllabuses, textbooks, and teacher training courses (Anderson, 2007, p. 660).

**Figure 1**

*The CEFR's Initial Division into Three Broad Levels—A, B and C (Council of Europe, 2001, p. 23)*




The CEFR has established statements defining what is required for each stage within the framework and is based on performance-based "can-do statements," or "Reference Level Descriptors," evolved from the collective judgments of a body of experts (Van Ek & Trim, 1991a, 1991b, 2001), as depicted quite clearly in Figure 2.

**Figure 2**

*An Illustrative Scale for Grammatical Accuracy (Council of Europe, 2001, p. 114)*

| | GRAMMATICAL ACCURACY |
|---|---|
| C2 | *Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).* |
| C1 | *Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.* |
| B2 | *Good grammatical control; occasional 'slips' or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.* |
| | *Shows a relatively high degree of grammatical control. Does not make mistakes which lead to misunderstanding.* |
| B1 | *Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influence. Errors occur, but it is clear what he/she is trying to express.* |
| | *Uses reasonably accurately a repertoire of frequently used 'routines' and patterns associated with more predictable situations.* |
| A2 | *Uses some simple structures correctly, but still systematically makes basic mistakes – for example tends to mix up tenses and forget to mark agreement; nevertheless, it is usually clear what he/she is trying to say.* |
| A1 | *Shows only limited control of a few simple grammatical structures and sentence patterns in a learnt repertoire.* |

For empirical detail about levels of competence in learner English, see the English Profile (Harrison & Barker (2015), for example, introduce the EP program and discuss its latest findings). Moreover, its sub-project, The English Grammar Profile (EGP) resource, is a database of over 1,200 empirically derived statements and provides great detail on learner grammar competence.

## CAF

Many researchers believe in the multi-componential nature of L2 performance, which can be captured by the notions of complexity, accuracy, and fluency (Ellis, 2003; Ellis & Barkhuizen, 2005). According to Skehan (2009), successful performance has been characterized as containing: (1) more advanced language, leading to complexity; (2) a concern to avoid errors, leading to higher accuracy if achieved; and (3) the capacity to produce speech at a normal rate and without interruptions, resulting in greater fluency (p. 510).

While the early working definitions of CAF are essentially used—complexity related to the size, richness, and diversity of L2 performance; accuracy as the measure for error-free language use; and fluency referring to the smooth production of speech with a limited number of hesitations and pauses (Michel, 2017)—many studies have featured CAF as variables that reflect the effect of other factors on language production (Housen et al., 2012, p. 2). These studies have included factors such as age (Mora, 2006), task design (Robinson et al., 2009), task complexity (Robinson, 2011), pre-task planning (Pang & Skehan, 2014), or the relationships among the CAF measures and learner oral proficiency (Miyamoto, 2019). Some studies have used CAF to examine longitudinal learner trajectories (Gunnarsson, 2012; Ferrari, 2012; Hokamura, 2018).

### *Complexity*

Housen et al. (2012, p. 4) describe complexity as a palimpsest, a term that has acquired several closely related meanings in the context of SLA. For example, Ellis and Barkhuizen (2005, p. 139) define complexity as "the use of more challenging and difficult language," while Biber et al. (2011, p. 6) refer to "the more advanced grammatical structures that students exhibit as they progress in their language proficiencies." The second definition highlights the ability to employ a range of sophisticated structures and lexical items, defining complexity based on empirical observations of L2 production. Although Bulté and Housen (2012) propose a comprehensive typology of subdimensions of linguistic complexity, the current study adheres to an objective and quantitative definition of complexity. Manifestations of complexity in L2 production include syntactic structures and grammatical morphemes (for more details, see: Data and Methods).

### *Accuracy*

Accuracy is often defined as the ability to produce error-free language (Ellis, 2008; Housen & Kuiken, 2009; Polio & Shea, 2014) and is largely associated with learner's linguistic knowledge representations. Palloti (2009, p. 592) mentions that "accuracy is perhaps the simplest and most internally coherent construct, referring to the degree of conformity to certain norms."

Although researchers often criticize the study of accuracy and mention that it is difficult to define an error (Gilquin & De Cock, 2011, p. 142), the existing definitions seem to concur in defining errors as deviations from a particular norm (Wolfe-Quintero et al., 1998). Straightforward though this general characterization seems, the difficulty seems to lie in selecting the criteria for evaluating accuracy and identifying errors, including the choice of the appropriate norm, that is, whether the criteria should be tuned to prescriptive standard norms or non-standard usages acceptable in some social contexts as well (Ellis, 2008; Polio, 1997).

The fact that the same language may have several normative standards or that raters might not always agree on what is accurate (Kuiken & Vedder, 2014) adds another layer to the discussion on the characterization and measures of accuracy. Moreover, "even if there was agreement regarding the norm, there remains the question of how 'far away' a deviation from this chosen norm is" (Michel, 2017, p. 9). All in all, "anyone who has worked on assessing accuracy in L2 data will know this only too well; some degree of personal judgment has to be invoked occasionally" (Foster & Wigglesworth, 2016, p. 112).

To avoid subjectivity, the current study focused on syntactic errors identified and tagged by two raters following Granger et al.'s (2022) error tagging manual and checked for inter-rater reliability (see: Data and Methods). Moreover, accuracy was gauged using a global measure (an error rate, i.e., number of errors per 100 words) for the purpose of this paper.

Having explored the intricate natures of language complexity and accuracy, which are both largely associated with learners' linguistic knowledge representations, we now turn to the third element of the triad: fluency, which is a performance phenomenon representing the outcome of psycholinguistic processing (Lennon, 1990).

### *Fluency*

Language researchers have frequently analyzed oral production data to determine which linguistic phenomena contribute to L2 speech fluency (see, e.g., Kormos & Dénes, 2004), mainly because L2 speakers are often concerned with maintaining fluency since failure to do so can lead to a loss of a listener's attention and their face (Lennon, 2000).

Current research suggests that speech fluency is a multi-componential construct with various sub-dimensions, such as speed fluency, breakdown fluency, or repair fluency (Tavakoli & Skehan, 2005). According to Wood (2012, p. 9), fluency is often used as a synonym for the effective spoken use of a language, while Ellis and Barkhuizen (2005, p. 139) define fluency as "the production of language in real time without undue pausing or hesitation." Skehan (2009) adds that fluency is the capacity to produce speech at a normal rate and without interruption, and Segalowitz (2010) distinguishes between cognitive (the smoothness of the underlying processes), utterance (acoustically measurable aspects of performance), and perceived (the speaker's fluency impression on the hearer) facets of fluency. What these definitions emphasize is that we associate fluency not only with a quantifiable dimension for describing language performance but also with an impression made on the hearer. However, many conversational traits arise from the fact that conversation is typically spontaneous, that is, it is characterized by what has been called normal dysfluency, such as pauses, hesitators (er, um), and repeats. These impair a speaker's turn when the planning needs to catch up (Biber et al., 1999, p. 1048).

The current study examined some of the most typical characteristic features of spoken discourse (Biber et al., 1999; Wood, 2012), namely: silent pauses, filled pauses, repair, and repeats.

## *CAF at Different Proficiency Levels*

The current article investigates complexity, accuracy, and fluency at different levels of CEFR. Larsen-Freeman (2009, p. 582) points out that examining these dimensions individually overlooks their interaction, which can change over time.

Some notable research focusing on the relationship between L2 proficiency and the CAF constructs sheds light on their interconnectedness and mutual influence. Some significant studies analyzed complexity and accuracy development (Polat & Kim, 2013; Spoelman & Verspoor, 2010, to name only a few), while other studies looked at the development of all three dimensions from a dynamic point of view. Larsen-Freeman (2006), for example, observed five Chinese adult learners over a six-month period. The average results showed progress in development in every dimension (although with great variation between subjects). Moreover, Iwashita et al. (2008) argue that features from grammatical accuracy and complexity, vocabulary, pronunciation, and fluency help distinguish overall levels of performance, with particular features of vocabulary and fluency having the strongest impact.

Finally, studies on information processing theory (Robinson, 2005; DeKeyser, 2007) and intra-individual variability (Spoelman & Verspoor, 2010) highlight

the non-linear nature of L2 development and the complex interaction between accuracy and complexity. Barrot and Gabinete (2021) suggest that complexity, accuracy, and fluency in writing are influenced not only by proficiency level but also by learners' L1 background, while Kowal (2018) found interplay between all three constructs.

Overall, these studies suggest correlations between and among the L2 proficiency and the components of complexity, accuracy, and fluency.

## Data and Methods

The present study adopts a quasi-longitudinal design and focuses on the analysis of speech produced by Czech university learners of English.

The analyzed corpus consists of transcribed interviews with 40 first-year university learners of English (recorded over a period spanning from 2020 to 2021) who were approximately 20 years old on average (SD = 1.4).

L2 speech production samples analyzed in the current study cover a range of proficiency levels, falling within the A2, B1, and B2[1] levels, which were assessed using The Oxford Placement Test (OPT) before the individual interviews.

**Table 1**

*Descriptive Statistics of the Transcribed Oral Performances*

| Category | Word count | | |
|---|---|---|---|
| | A2 | B1 | B2 |
| Minimum | 467 | 416 | 746 |
| Maximum | 1091 | 1305 | 1571 |
| Mean | 724 | 800 | 1014 |
| Median | 723 | 777 | 922 |
| SD | 194 | 258 | 272 |

The interviews, each of which took approximately 15 minutes, consisted of three parts: a topic discussion, a picture description, and free conversation. First, the respondents were asked to choose a topic[2] and speak about it for three minutes without any interruptions. The first section of the oral interview was stimulated by the interviewer's question (*What topic have you chosen?*) when

[1] A2 (13 learners), B1 (15 learners), B2 (12 learners).

[2] Topic 1: A film you have seen/ a book you have read and think is particularly good/bad. Topic 2: A place or a country you have visited and liked.

the recording was started.[3] The learners were supposed to produce an extensive response. The second task was based on picture description, while the last task (free conversation) started with the students introducing themselves to enable them to talk about something familiar. Then, the interviewer posed topical scripted questions (e.g., *What can you tell me about your family?* or *Do you think English will be useful for you in the future?*), which were mostly concerned with familiar topics and with learning English. It should go without saying that some of the benefits of this task type can be re-constructed as weaknesses with lower-level learners. The process was rather free-flowing and indeterminate with talkative and accurate learners while less talkative (and less accurate) learners were often guided by prescripted questions.

In the present study, task performances were transcribed and analyzed, focusing on measures of complexity, accuracy, and fluency. To ensure comparability with previous research and address the research objectives, several general measures were assigned to each construct, taking into account their validity based on relevant studies (Bosker et al., 2013; Mostafa et al., 2020; Noris & Ortega, 2009; Polio & Shea, 2014; or Shungo & Kormos, 2020), and research objectives.

The study examined both lexical and syntactic complexity. Following Shungo and Kormos (2020), lexical complexity was measured as lexical density, which refers to the proportion of content words to the total words produced. This measure was computed using LexTutor (Cobb, 2011). In addition to lexical complexity, syntactic complexity was assessed using the Index of Developmental Levels (henceforth referred to as IDL), following the approach of Mostafa et al. (2020). They argue that IDL is a better predictor of L2 oral proficiency than widely-used structural complexity measures (mean length of AS-unit and subordination measure). By employing this measure of complexity, the study aims to examine the learners' ability to use advanced syntactic structures in their speech production. Table 2 likely presents the details of the IDL calculations for different morphemes and syntactic forms.

Following Mostafa et al. (2020), different weights were assigned to forms based on their developmental levels. Therefore, in the score sheets, 1 or 2 points (representing single or double occurrences of a form) were multiplied by the respective developmental levels of those forms. In other words, the EFL learners' total IDL scores increased as they produced more developmentally advanced target forms.

[3] Utilizing the built-in recording capabilities of a smartphone and a computer.

**Table 2**

*Developmental Levels and Stages for the Acquisition of the English Morphemes, Negation, Questions, and Relative Clauses*

| Developmental levels for English morphemes | Developmental stages for the acquisition of English negation | Developmental levels for English questions | Acquisitional Order of Relative Clauses |
|---|---|---|---|
| Morpheme (level) | Item (level) | Item (level) | Relative clause type (level) |
| Ing (1)<br>Plural -s (1)<br>Be copula (1)<br>Be auxiliary (2)<br>a/the (2)<br>Irregular past (3)<br>Regular past -ed (4)<br>Third person -s (4)<br>Possessive -s (4) | Preverbal negation with no/not (1)<br>Preverbal negation with don't (2)<br>Postverbal negation in restricted contexts (3)<br>Postverbal negation in all contexts (4) | Words and fragments with rising intonation (1)<br>Canonical word order with rising intonation (2)<br>Fronting of a questioning element (3)<br>Inversion in two restricted contexts (4)<br>Inversion expands to full range of target-like contexts (5)<br>Negative questions/Question tags/ Questions in embedded clauses (6) | Subject (1)<br>Direct object (2)<br>Indirect object (3)<br>Object of preposition (4)<br>Genitive (5)<br>Object of comparison (6) |

*Note*. Separate IDL score sheets for morphemes, questions, negations, and relative clauses were developed for each participant. The participants received 1 point for producing a target form once, and they received 2 points for producing that form twice. For each type of form, the two examples had to be sufficiently dissimilar to be awarded the full points. They did not receive any additional point for producing a target form more than twice. In counting the IDL scores, grammatically acceptable forms were considered.

Regarding accuracy, a general measure known as the syntactic error rate (SER, i.e., the average number of syntactic errors per 100 words) was used. This measure was chosen for its comparability with previous research and its validity. A British English native speaker was trained to identify all syntactic errors (errors that contravene general rules of English grammar), which were categorized into sub-categories including determiners, articles, nouns, pronouns, adjectives, adverbs, verbs, word class, as described in Granger et al. (2022). To illustrate some of the syntactic errors observed, several examples were selected from an interview. Among other errors made by learner TT05, there were errors in the use of articles (GA), noun number (GNN), verb tense (GVT), verb number (GVN), personal pronouns (GPP), and errors involving independent prepositions (LSPR). Examples (1) to (4) demonstrate these errors.

1) ... and sight-seeing it's <GA corr="a very nice place"> very nice place </GA> with mountains and Alps it's <GA corr="a very good place"> very good place </GA> for young people maybe and old people... TT05
2) <B> Four <GNN corr="months"> month </GNN> in <name of a town> in <GA corr="a hotel"> hotel </GA> </B> TT05

3) ... I <GVT corr="visited"> visit </GVT> with my family and my family <GVN corr="don't"> doesn't </GVN> speak English and I must translate everything TT05
4) ...and <GPP corr="the people"> they people </GPP> are <LSPR corr="from other"> other </GPP> other <GNN corr="nations"> nation </GNN> I can see two girls and two boys... TT05

To assess the reliability of the accuracy measure, the author of this study examined a randomly selected 20% of the data. The inter-rater reliability, measured using Cohen's kappa, was found to be 0.92, indicating a high level of agreement (96.23%). This level of agreement can be considered excellent (Fleiss et al., 2013).

In the current study, no distinction was made between errors and mistakes (Ellis, 1997).[4] Any deviations from the norms of English grammar in terms of syntax were counted as errors. Additionally, omitted grammatical forms were counted as errors. However, self-corrections made by EFL learners after committing an error were not counted.

Finally, fluency, despite its multifaceted nature and a myriad of definitions, is defined in this study as "the production of language in real time without undue pausing or hesitation" (Ellis & Barkhuizen, 2005, p. 139). Drawing from Shungo and Kormos (2020), three subcomponents of utterance fluency were identified: speed, breakdown, and repair fluency. Unfilled pauses, defined by Bosker et al. (2013) as silence longer than 250 milliseconds, were manually coded by the researcher using Praat (Boersma & Weenink, 2012), marking the boundaries of clauses and pauses.

Speed fluency was assessed using the articulation rate (henceforth referred to as AR), calculated as the mean number of words per second, and divided by the total speech duration excluding pauses. Breakdown fluency was measured through the filled-pause ratio (FPR), which is the total number of filled pauses (e.g., ah, eh) divided by the total number of words. Repair fluency was determined by the dysfluency rate (DR), calculated as the mean number of dysfluencies (unfilled and filled pauses, repeats, and repairs) per minute, divided by the total speech duration (including pauses).

Table 3 provides a comprehensive overview of the specific measures of complexity, accuracy, and fluency.

[4] According to Ellis (1997), **errors** reflect gaps in the learners' linguistic knowledge (i.e., they do not know the correct form) while **mistakes** are occasional lapses in the learners' performance (which happens when the learners are not able to perform what they know).

**Table 3**

*CAF Measurements*

| Measure | Calculation |
|---|---|
| *Syntactic complexity* | |
| Index of Developmental Levels (IDL) | separate IDLs |
| *Lexical complexity* | |
| Lexical density (LD) | content words/total words produced |
| *Accuracy* | |
| Syntactic error rate (SER) | syntactic errors/100words |
| *Fluency* | |
| Articulation rate (AR) | words per second/total speech duration |
| Filled pause ratio (FPR) | filled pauses/words |
| Dysfluency rate (DR) | dysfluencies per minute/total speech duration |

To describe L2 speech production samples, descriptive statistics of linguistic measures were summarized in tables (see sections: Complexity in the Speech of EFL Learners; Accuracy in the Speech of EFL Learners; Fluency in the Speech of EFL Learners). The Shapiro-Wilk normality test suggested whether linguistic measures were normally distributed and where the Shapiro-Wilk normality test suggested that a linguistic measure was not normally distributed; nonparametric statistical tests were selected to correlate EFL learners' proficiency with specific linguistic measures. ANOVAs were performed to identify differences in the CAF measures among proficiency levels.

In the following section, I will provide an overview of the quantitative differences observed among the scrutinized proficiency levels.

## Results

To address the first and second research questions, which investigate the predictive nature of complexity, accuracy, and fluency on L2 learners' oral proficiency, as well as the identification of the most discriminating measures between proficiency levels, the present study examined the relationship between these three dimensions and the learners' proficiency levels.

Initially, a comprehensive analysis was conducted, scrutinizing complexity, accuracy, and fluency individually in relation to the proficiency level of EFL learners. The Shapiro-Wilk normality tests indicated that three measures (IDL, FPR, and AR) were not normally distributed. As a result, nonparametric statistical tests were selected to assess the correlation between the learners' proficiency levels and the three measures. ANOVAs were performed to identify differences in the CAF measures among proficiency levels.

Furthermore, Pearson correlation analyses and Spearman correlation analyses were carried out between measures of complexity (LD and IDLs), accuracy (SER), and fluency (AR, FPR, and DR).

These analyses were conducted both overall and within each performance.

The following sections of the paper present descriptive and inferential statistics, as well as correlation analyses.

## Complexity in the Speech of EFL Learners

As previously mentioned, lexical and syntactic complexity were assessed using lexical density (the ratio of content words to the total words produced, computed with LexTutor, Cobb, 2011) and an index based on the developmental levels of L2 morphological and syntactic forms (Index of Developmental Levels, IDL).

Before presenting the results, descriptive statistics for IDL and LD are provided in Table 4. Unlike IDL, which emerged as a significant positive predictor of L2 oral proficiency, lexical density did not prove to be a reliable predictor of proficiency. Specifically, B2-level learners exhibited significantly higher IDL scores compared to A2 and B1 levels, indicating that IDL scores increase as learners' proficiency levels improve.

**Table 4**

*Descriptive Statistics—Complexity*

| Coefficient | CEFR level | Mean | SD |
|---|---|---|---|
| LD | A2 | 0.485 | 0.03 |
| | B1 | 0.457 | 0.02 |
| | B2 | 0.463 | 0.03 |
| IDL | A2 | 26.31 | 11.10 |
| | B1 | 46.13 | 18.52 |
| | B2 | 60.25 | 12.02 |

Pairwise comparisons were conducted to detect differences between proficiency levels. The results indicated a significant overall effect on the proficiency level for IDL, as shown in Table 5.

**Table 5**

*Pairwise Comparisons of Proficiency Levels*

| CEFR levels | IDL | |
|---|---|---|
| | Test | Sig. |
| A2–B1 | –11.846 | 0.007 |
| A2–B2 | –21.346 | 0.000 |
| B1–B2 | –9.500 | 0.036 |

All in all, the results proved that L2 learners with higher proficiency levels used more difficult morphological and syntactic structures.

## Accuracy in the Speech of EFL Learners

The corpus of 40 transcribed interviews was analyzed to identify errors that involved the violation of specific grammatical rules[5] (Granger et. al., 2022). Accuracy was measured using a general metric called syntactic error rate, which represents the mean number of syntactic errors per 100 words. For the corresponding data, refer to Table 6.

**Table 6**

*Descriptive Statistics—Accuracy*

| Proficiency level | Mean | SD |
|---|---|---|
| A2 | 9.16 | 2.64 |
| B1 | 6.91 | 2.31 |
| B2 | 4.53 | 1.54 |
| Total | 6.93 | 2.85 |

The results of a detailed syntactic error analysis in the speech of Czech EFL learners revealed that by far the most frequent error type was errors in the use of articles, accounting for 30.8% of all errors. Verb tense was the second most frequent category (13.1%), followed by errors involving independent prepositions (10.4%). Interestingly, there did not appear to be significant differences in error types among the three proficiency levels, which shows that the use

[5] In their Error Tagging Manual, Granger et. al. (2022, p. 4) make an important distinction between errors (the breaking of a specific linguistic rule) and infelicities (instances of non-erroneous, but odd-sounding, language). Infelicities were not taken into account in the analysis.

of articles and verb tense is highly problematic even as the learners progress to a higher proficiency level.

Unfortunately, Granger et al. (2022) do not divide errors in the use of articles into subgroups. For this reason, it is rather difficult to determine which aspect of their use is the most problematic for Czech EFL learners. A more detailed analysis, however, revealed that omission of *a/an* in obligatory contexts accounts for 54.1% of all article errors, and that omission of *the* in obligatory contexts accounts for another 23.8% of all errors in the category of articles. Such a result seems to suggest that article omission in English might be the result of L1 transfer as Czech does not have articles.

As mentioned above, the analysis of the most frequent error types indicated similar frequencies across proficiency levels. Articles proved to be the most challenging grammatical feature for the majority of Czech EFL learners, including those in the A2–B2 proficiency groups. This was followed by verb tense and prepositions, which were found to be common sources of errors for Czech EFL learners across all three proficiency levels. For details on the types of errors made by advanced EFL learners, refer to Gráf (2015).

Moreover, to identify differences between proficiency levels, ANOVAs and pairwise comparisons were conducted. The results demonstrated a significant correlation between EFL learners' proficiency levels and syntactic error rate, indicating that B2 learners exhibited the highest level of accuracy, while A2 learners had the least accurate speech production. Refer to Table 7 for detailed information.

**Table 7**

*Multiple Comparisons, LSD Test*

| Proficiency level | | Mean difference | Sig. |
|---|---|---|---|
| A2 | B1 | 2.2482 | 0.011 |
| | B2 | 4.6282 | 0.000 |
| B1 | A2 | –2.2482 | 0.011 |
| | B2 | 2.3800 | 0.009 |
| B2 | A2 | –4.6282 | 0.000 |
| | B1 | –2.3800 | 0.009 |

The current study proposes that syntactic accuracy on the part of EFL speakers could serve as a reliable predictor of language proficiency. The findings indicate a strong association between EFL learners' proficiency level and grammatical accuracy, which is significant for several reasons: (1) language assessors can use syntactic accuracy as a valuable indicator when evaluating the

proficiency of EFL learners, (2) educators may focus on designing instructional materials and activities that specifically target syntactic skills, (3) teachers can use this information to target their instruction to address syntactic challenges commonly faced by learners at different proficiency levels (for example, articles, or verb tense), (4) the finding encourages further investigation into the specific mechanisms through which syntactic skills impact overall proficiency.

## Fluency in the Speech of EFL Learners

In this section of the paper, the analysis focused on examining the relationship between proficiency level and fluency in the speech of Czech EFL learners. Fluency was measured using three indicators: articulation rate (AR), filled pause ratio (FPR), and dysfluency rate (DR).

The statistical analysis revealed that both articulation rate ($p < 0.012$) and dysfluency rate ($p < 0.019$) exhibited a significant relationship with proficiency levels. Specifically, the results indicated that B2-level learners were statistically more fluent than A2 and B1 learners. Detailed information can be found in Table 8.

**Table 8**

*Pairwise Comparisons*

| Proficiency levels | Test | Sig. |
|---|---|---|
| A2–B1 | –4.503 | 0.307 |
| A2–B2 | –13.603 | 0.003 |
| B1–B2 | –9.100 | 0.043 |

Regarding the dysfluency rate, a significant difference was observed between A2 and B2 levels ($p < 0.005$). This indicates that B2-level learners exhibit a higher level of fluency compared to A2-level learners. For more information on fluency in the speech of EFL learners, see, for example, Huang and Gráf (2020), who compared speech rates and the frequency and location of unfilled pauses in the speech of native English speakers and learners of English. They found significant differences showing that between B2 and C1 levels the growth of proficiency is accompanied by an increase in speech rate and a decrease in the frequency of pausing, particularly within clauses and within constituents (Huang & Gráf, 2020, p. 57).

In summary, the findings indicate that all three dimensions of complexity, accuracy, and fluency (measured as SER, IDL, AR, and DR) significantly predict L2 learners' oral proficiency. Each dimension demonstrated at least

one coefficient that established a statistically significant relationship with EFL learners' proficiency levels. Therefore, the coefficients for syntactic error rate, index of developmental levels, articulation rate, and dysfluency rate can be considered as the measures that are most effective in discriminating between proficiency levels.

### Interactions between Complexity, Accuracy, and Fluency

To address the third research question, which focused on the interactions between complexity, accuracy, and fluency, several correlation analyses were conducted separately for the measures of the three dimensions.

First, the correlation analysis between complexity (measured as IDL) and fluency (measured as AR) revealed a positive correlation between IDL and AR for the A2 proficiency level ($R = 0.574$, $p = 0.040$) and B1 proficiency level ($R = 0.651$, $p = 0.009$). This suggests that as learners used more complex morphological and syntactic structures, their oral production was also more fluent.

Second, negative correlations were observed between IDL and FPR for A2 ($R = –0.617$, $p = 0.025$) and B1 ($R = –0.527$, $p = 0.043$) proficiency levels, as well as between AR and FPR for the A2 ($R = –0.728$, $p = 0.005$) and B1 ($R = –0.692$, $p = 0.004$) proficiency levels. These findings indicate that as EFL learners produced more complex speech, their fluency increased while the filled pause ratio decreased.

In conclusion, there were no statistically significant correlations found between specific measures of complexity, accuracy, and fluency for B2-level learners, indicating that they, and more importantly all EFL learners across proficiency levels, can focus on all three dimensions simultaneously. Consequently, the findings confirm the interplay of complexity, accuracy, and fluency as interconnected systems. The development of these proficiency dimensions occurs concurrently.

## Conclusions

The current research proposed measures for evaluating L2 learner oral performance and examined whether these predict L2 oral proficiency. It indicates that specific measures (for example, Index of Developmental Levels, based on developmental levels of L2 forms) better discriminate between learners of varied proficiency levels and provide empirical evidence that in oral performances, learners of higher proficiency use significantly more L2 forms that belong

to higher levels in developmental sequences than lower proficiency levels learners. As such, the current study provides empirical support "linking performance measures to the use of developmentally more advanced language" (Lambert & Kormos, 2014), which may have pedagogical significance. ESL teachers can, for example, gain insight into the current developmental levels of their learners for various L2 morpho-syntactic forms (by considering a measure such as IDL), or provide feedback on weak points of EFL learners, such as the use of articles and tenses, by considering syntactic error rate (SER) and examining the most frequent error types. These might be hard to detect for the speed with which learners produce speech, or might be wreathed in sophisticated avoidance strategies. Furthermore, the division of the respondents into proficiency levels helped identify the most problematic areas as those that occur throughout all of the proficiency bands.

Additionally, the analysis of accuracy highlights the practical significance of studying learner errors and has important implications for error correction, because understanding the nature of errors is crucial for effective error correction in language teaching (see, for example, Scrivener, 2005). Teachers need to have a thorough understanding of the types of errors made by learners to provide targeted and effective feedback. By identifying common error patterns and focusing on specific areas of difficulty, teachers can tailor their instruction and help learners improve their accuracy in language production.

Moreover, the current study contributes to the existing knowledge on the interplay between complexity, accuracy, and fluency in learner language by confirming the prevailing view that these three dimensions are interconnected. Furthermore, it demonstrates that complexity, accuracy, and fluency are influenced by proficiency level, aligning with the findings of previous research (Barrot & Gabinete, 2021). By and large, from the developmental perspective, it shows that L2 learners' output is produced with higher complexity, accuracy, and fluency when they progress to a higher proficiency level, that is, L2 learners with higher oral proficiency are more adept at using more advanced L2 forms, with fewer errors and dysfluencies. This suggests that as L2 learners develop proficiency, the structures they employ become more accurate, their speech becomes more fluent, and they tend to utilize L2 forms that belong to developmentally higher levels.

However, average group trajectories may not correspond to the developmental trajectory of a single individual subject (Larsen-Freeman, 2006). Thus, it is necessary to look at specific scores obtained by L2 learners individually and over more data collection points to reveal whether complexity, accuracy, and fluency follow separate developmental trajectories or not.

Finally, it is important to recognize that the components of the CAF triad are not isolated but interact with each other, as suggested by Larsen-Freeman (2006). The interplay between the three dimensions is multivariate and dynamic

(Spoelman & Verspoor, 2010, p. 547), meaning that changes in one dimension can affect the others. This highlights the complexity of language acquisition and the need for comprehensive research. To gain a deeper understanding of these relationships is necessary to conduct longitudinal and non-linear CAF research with a focus on difference and variation (Larsen-Freeman, 2009). For this reason, a more in-depth analysis of learner data, for example, a longitudinal one, would certainly reveal more interesting regularities.

## References


Anderson, C. (2007). The CEFR and the need for more research. *Modern Language Journal*, *91*(4), 659–663.

Baddeley, A. (2007). *Working memory, thought, and action*. Oxford University Press.

Barrot, J., & Gabinete, M. K. (2021). Complexity, accuracy, and fluency in the argumentative writing of ESL and EFL learners. *International Review of Applied Linguistics in Language Teaching, 59*(2), 209–232.

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson.

Biber, D., & Conrad, S. (2001). Quantitative corpus-based research: Much more than bean counting. *TESOL Quarterly, 35*(2), 331–336.

Biber, D., Gray, B., & Poonpon, K. (2011). Should we use characteristics of conversation to measure grammatical complexity in L2 writing development? *TESOL Quarterly, 45*(1), 5–35.

Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing, 30*, 159–175.

Bulté, B., & Housen, A. (2012). Defining and operationalising L2 complexity. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 23–46). John Benjamins.

De Felice, R., & Pulman, S. (2009). Automatic detection of preposition errors in learner writing. *CALICO Journal, 26*(3), 512–528.

DeKeyser, R. (2007). Skill acquisition theory. In B. Van Patten & J. Williams (Eds.), *Theories in Second Language Acquisition: An Introduction* (pp. 94–112). Lawrence Erlbaum.

Ellis, R. (1994). *The study of second language acquisition*. Oxford University Press.

Ellis, R. (1997). *Second Language Acquisition*. Oxford University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2008). *The Study of Second Language Acquisition (2nd edition*). Oxford University Press.

Ellis, R., & Barkhuizen, G. (2005). *Analyzing learner language*. Oxford University Press.

Ferrari, S. (2012). A longitudinal study of complexity, accuracy, and fluency variation in second language development. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy, and fluency in SLA* (pp. 277–298). John Benjamins Publishing Company.

Fleiss, J. L., Levin, B., & Paik, M. C. (2013). *Statistical methods for rates and proportions*. John Wiley & Sons.

Foster, P., & Wigglesworth, G. (2016). Capturing accuracy in second language performance: The case for a weighted clause ratio. *Annual Review of Applied Linguistics, 36,* 98–116.

Gilquin, G., & De Cock, S. (2011). Errors and disfluencies in spoken corpora: Setting the Scene. *International Journal of Corpus Linguistics, 16*(2), 141–172.

Gráf, T. (2015). *Accuracy and fluency in the speech of the advanced learner of English*. Univerzita Karlova v Praze.

Granger, S., Swallow, H., & Thewissen, J. (2022). *The Louvain Error Tagging Manual. Version 2.0.* CECL Papers 4. Centre for English Corpus Linguistics: Louvain-la-Neuve.

Gunnarsson, C. (2012). The development of complexity, accuracy, and fluency in the written production of L2 French. In A. Housen, F. Kuiken, & I. Vedder (Eds.), *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy, and Fluency in SLA* (pp. 247–276). John Benjamins Publishing Company.

Gyllstad, H., Granfeldt, J., Bernardini, P., & Källkvist, M. (2014). Linguistic correlates to communicative proficiency levels of the CEFR. The case of syntactic complexity in written L2 English, L3 French and L4 Italian. *EUROSLA Yearbook, 14*, 1–30.

Harrison, J., & Barker, F. (2015). *English profile studies: English profile in practice*. Cambridge University Press.

Hokamura, M. (2018). The dynamics of complexity, accuracy, and fluency: A longitudinal case study of Japanese learners' English writing. *JALT Journal, 40*(1), 23–46.

Housen, A., & Kuiken, F. (2009). Complexity, accuracy, and fluency in Second Language Acquisition. *Applied Linguistics*, *30*(4), 461–473.

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 performance and proficiency: Complexity, accuracy, and fluency in SLA*. John Benjamins Publishing Company.

Huang, L.-F., & Gráf, T. (2020). Speech rate and pausing in English: Comparing learners at different levels of proficiency with native speakers. *Taiwan Journal of TESOL, 17*(1), 57–86.

Hulstijn, J. H. (2011). Language proficiency in native and nonnative speakers: An agenda for research and suggestions for second-language assessment. *Language Assessment Quarterly, 8*(3), 229–249.

Iwashita, N., Brown, A., McNamara, T., & O'Hagan, S. (2008). Assessed levels of second language speaking proficiency: How distinct? *Applied Linguistics, 29*(1), 24–49.

Jones, N., & Saville, N. (2009). European language policy: Assessment, learning, and the CEFR. *Annual Review of Applied Linguistics, 29*, 51–63.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*(2), 145–164.

Kowal, I. (2018). *The dynamics of complexity, accuracy, and fluency in second language development*. Jagiellonian University Press.

Kuiken, F., & Vedder, I. (2014). Rating written performance: What do raters do and why? *Language Testing, 31*(3), 329–348.

Lambert, C., & Kormos, J. (2014). Complexity, accuracy, and fluency in task based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics, 35*(5), 607–614.

Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of the Chinese learners of English. *Applied Linguistics, 27*(4), 590–619.

Larsen-Freeman, D. (2009). Adjusting expectations: The study of complexity, accuracy, and fluency in second language acquisition. *Applied Linguistics, 30*(4), 579–589.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*(3), 387–417.

Lennon, P. (2000). The lexical element in spoken second language fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25–42). University of Michigan Press.

Michel, M. C. (2017). Complexity, accuracy, and fluency in L2 production. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 50–68). Taylor & Francis.

Miyamoto, M. (2019). *Capturing L2 oral proficiency with CAF measures as predictors of the ACTFL OPI rating*. [Doctoral dissertation, Purdue University, School of Languages and Cultures, West Lafayette, Indiana, 2019].

Mora, J. C. (2006). Age Effect on Oral Fluency Development. In C. Muñoz (Ed.), *Age and the Rate of Foreign Language Learning* (pp. 65–88). Multilingual Matters.

Mostafa, T., Kim, Y., & Friginal, E. (2020). Examining a developmentally based measure of L2 oral performances: Does it predict L2 learners' oral proficiency? *System, 89*, 1–16.

Nation, I. S. P. (2009a). *Teaching ESL/EFL reading and writing*. Routledge.

Nation, I. S. P. (2009b). *Teaching ESL/EFL listening and speaking*. Routledge.

Nation, I. S. P. (2011). Research into practice: Vocabulary. *Language Teaching*, *44*(4), 1–11.

Norris, J. M., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied Linguistics, 30*(4), 555–578.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492–518.

Ortega, L. (2014). Trying out theories on interlanguage: Description and explanation over 40 years of L2 negation research. In Z. Han & E. Tarone (Eds.), *Interlanguage: Forty years later* (pp. 173–202). John Benjamins.

Palloti, G. (2009). CAF: Defining, refining, and differentiating constructs. *Applied Linguistics, 30*(4), 590–601.

Pang F., & Skehan, P. (2014). What do learners do when they plan? In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 95–128). John Benjamins.

Polat, B., & Kim, Y. (2013). Dynamics of complexity and accuracy: A longitudinal case study of advanced untutored development. *Applied Linguistics, (35)*2, 184–207.

Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*, 101–143.

Polio, C., & Shea, M. C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing, 26*, 10–27.

Porcino, M. C., & Finardi, K. R. (2012). The tension between accuracy and fluency of L2 speech: Evidence from communicative tasks. *The ESPecialist, 33*(1), 25–44.

Robinson, P. (2003). The cognition hypothesis, task design, and adult task-based language learning. *Second Language Studies, 21*(2), 45–107.

Robinson, P. (2005). Aptitude and second language acquisition. *Annual Review of Applied Linguistics, 25*, 46–73.

Robinson, P. (2009). Syllabus design. In M. H. Long & C. J. Doughty (Eds.), *The Handbook of Language Teaching* (pp. 294–310). Blackwell Publishing.

Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3–38). John Benjamins.

Rossiter, M. J., Derwing, T. M., Manimtim, L. G., & Thomson, R. I. (2010). Oral fluency: The neglected component in the communicative language classroom. *The Canadian Modern Language Review*, *66*(4), 583–606.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Scrivener, J. (2005). *Learning teaching: A guidebook for English language teachers*. MacMillan Education.

Shungo, S., & Kormos, J. (2020). Linguistic dimensions of comprehensibility and perceived fluency: An investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition, 42*(1), 143–167.

Skehan, P. (2009). Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics, 30*(4), 510–532.

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

Spoelman, M., & Verspoor, M. (2010). Dynamic patterns in the development of accuracy and complexity: A longitudinal case study in the acquisition of Finnish. *Applied Linguistics, 31*(4), 532–553.

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–273). Benjamins.

Taylor, D. (1988). The meaning and use of the term 'competence' in linguistics and applied linguistics. *Applied Linguistics, 9*, 148–168.

Van Ek, J. A., & Trim, J. L. M. (1991a). *Waystage 1990*. Council of Europe/Cambridge University Press.

Van Ek, J. A., & Trim, J. L. M. (1991b). *Threshold 1990*. Council of Europe/Cambridge University Press.

Van Ek, J. A., & Trim, J. L. M. (2001). *Vantage*. Council of Europe/Cambridge University Press.

Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy, and complexity*. University of Hawaii, Second Language Teaching and Curriculum Center.

Wood, D. (2012). *Formulaic language and second language speech fluency: Background, evidence and classroom applications*. Continuum.

Zhang, M. (2018). Collaborative writing in the EFL classroom: The effects of L1 and L2 use. *System, 76*, 1–12.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/19
https://doi.org/10.31261/TAPSLA.16594

CC BY SA

**Eihab Abu-Rabiah**
Kaye Academic College of Education,
Ben-Gurion University of the Negev, Israel
https://orcid.org/0000-0002-8837-1089

# The Productive Vocabulary Size of Second Language Learners upon Entry into Higher Education

Abstract

Extensive vocabulary acquisition is a cornerstone of second language (L2) proficiency, directly influencing both receptive and productive language skills. However, research on the productive vocabulary size of L2 learners transitioning to higher education, particularly their mastery of high-frequency words, remains limited. This study investigated the productive Hebrew vocabulary size and frequency distribution of Arabic-speaking learners entering higher education programs where Hebrew is the primary language of instruction.

The research employed a corpus-driven approach, analyzing 156 Hebrew-language argumentative essays (18,054 orthographic words) written by native Arabic-speaking students during a college entrance examination. Automated tools were used to add contextual vocalization and disambiguate homographs, followed by manual annotation mapping each word to its corresponding lemma. The identified lemmas were then compared to established written and spoken Hebrew frequency lists. This process aimed to chart the vocabulary profile of the research population.

The study determined that learners had a productive vocabulary size of approximately 1,000 lemmas, despite completing over 1,000 hours of formal L2 instruction. A comparison with established written and spoken Hebrew frequency lists indicated that 50% of the identified lemmas fell within the 1,000 most frequent Hebrew lemmas. Additionally, the learners exhibited a typical vocabulary profile, employing more lemmas from the 1k frequency band (the 1,000 most frequent words) than from the 2k frequency band (words ranked 1,001–2,000). Similarly, their use of lemmas from the 2k band exceeded that of the 3k band (words ranked 2,001–3,000), which in turn surpassed their use of lemmas from the 4k band (words ranked 3,001–4,000). These findings highlight the learners' significant reliance on high-frequency vocabulary in L2 writing, emphasizing the need for targeted academic vocabulary instruction as they transition to higher education.

*Keywords*: second language acquisition, writing assessment, vocabulary assessment, frequency distribution, Zipf's law, Arabic-speaking learners, Hebrew frequency lists

# Literature Review

## Vocabulary Knowledge

Vocabulary knowledge serves as the fundamental underpinning of language, and, without it, both language comprehension and production become unattainable (Milton & Treffers-Daller, 2013). As aptly noted by Wilkins (Wilkins, 1972, p. 111), "without grammar, very little can be conveyed, [but] without vocabulary, nothing can be conveyed." Vocabulary knowledge, or word knowledge, can be categorized into two primary types: receptive (or passive) knowledge and productive (or active) knowledge. Receptive knowledge pertains to words recognized when heard or read, surpassing the less extensive productive knowledge, which involves words recalled for use in speech or writing (Milton, 2009).

Words exhibit variations in multiple facets, including their sounds, letters, meanings, length, inflections, derivations, and more. However, the most pivotal aspect of these variances is the frequency with which words occur. Frequency plays a pivotal role since it determines which words learners are likely to encounter and how frequently they encounter them (Milton, 2009). Furthermore, words that are more frequent are typically acquired earlier (Ambridge et al., 2015). It is evident that certain words occur more frequently than others, requiring little knowledge about language (Schmitt & Schmitt, 2020).

## Vocabulary Size

To gauge vocabulary knowledge and explore differences in word frequencies, one common approach is measuring vocabulary size or breadth. Vocabulary size signifies the number of words an individual is familiar with (Schmitt, 2014), making it a fundamental aspect of vocabulary research. Vocabulary size can be evaluated receptively and productively. Receptive vocabulary size is often evaluated through tests that assess knowledge of a sample of words (Nation & Coxhead, 2021). An example of such a tool is the Vocabulary Knowledge Scale, which uses a 5-point self-report system. The VKS allows learners to indicate their level of familiarity with specific vocabulary items, ranging from not recognizing the word at all to being able to use it confidently in a sentence (Wesche & Paribakht, 1996). Productive vocabulary size is assessable through written compositions or an individual's language output (Nation & Coxhead, 2021). It's essential to recognize that not all known words may be utilized, and longer language productions tend to incorporate a greater variety of words (Richards & Malvern, 1997; Nation & Coxhead, 2021). Studies, such as those by Pan et al. (2005) and Rowe et al. (2012), focused on written language, counting

distinct words. For example, Shakespeare's vocabulary exceeded 20,000 words, determined by scrutinizing word forms in his works (Craig, 2011). In spoken language, numerous studies transcribed and tallied words from recorded language speaker output (Nation & Coxhead, 2021).

Assessing students' receptive vocabulary size offers teachers insights into their text and listening comprehension capabilities. Conversely, understanding their productive vocabulary size provides an indication of their speaking and writing proficiency (Webb, 2008). Research consistently shows that learners' receptive vocabulary size tends to be larger than their productive vocabulary size (Fan, 2000; Webb, 2008). Studies report this ratio can vary between 50% and 80% (Milton, 2009; Nizonkiza, 2016). In general, receptive vocabulary size can be a good indicator of productive vocabulary size (Webb, 2008). It is important to note, however, that this receptive-productive relationship may be moderated by the specific language learning environment (Webb, 2008). In foreign language contexts with limited opportunities for naturalistic L2 use outside the classroom, the receptive-productive ratio could differ from patterns observed in second language immersion settings where learners have greater access to the L2 in daily life.

Examining the vocabulary size of native English speakers reveals an average vocabulary size of 10,000 to 11,000 word families among educated adult L1 English speakers (Milton & Treffers-Daller, 2013; Brysbaert et al., 2016). Generally, it is agreed that a substantial vocabulary is requisite for English language functionality, with estimates spanning from 2,000 to 10,000 word families (Milton & Treffers-Daller, 2013; Van Zeeland & Schmitt, 2013). For everyday spoken communication in English, 2,000 to 3,000 word families suffice (Van Zeeland & Schmitt, 2013), while 6,000 to 7,000 word families are needed for watching movies and TV (Webb & Rodgers, 2009a, 2009b), and 8,000 to 9,000 word families for extensive reading (Nation, 2006).

In contrast, second language learners typically possess more limited vocabularies (Laufer & Yano, 2001). Receptively, these learners typically grasp 2,000 to 4,000 word families, despite over 1,000 hours of instruction (Laufer, 2017). However, research also suggests that L2 learners at a university level can achieve larger vocabularies. For instance, Lenko-Szymanska (2002) found that Polish learners of English possessed a passive vocabulary knowledge of 8,236 to 9,218 words (first-year vs. fourth-year). Overall, vocabulary size measurements are employed as general indicators of language proficiency and are strongly correlated with proficiency in reading, writing, listening, and speaking skills (Llach & Gallego, 2009; Milton et al., 2010; Mairano & Santiago, 2020).

### Importance of the Current Study

Vocabulary size is a cornerstone of fluency in any language. However, a critical gap exists in SLA research: we lack robust data on the productive vocabulary size of learners entering higher education, where the target language (L2) becomes the primary mode of instruction, particularly concerning high-frequency words. This study aims to bridge this gap by investigating the productive Hebrew vocabulary size of Arabic-speaking learners about to enter higher education. By analyzing written essays through a corpus-based approach, the research will explore vocabulary distribution and identify frequently used words. These findings, compared to existing Hebrew frequency lists, will provide a comprehensive picture of the learners' productive vocabulary landscape. By shedding light on the vocabulary size and characteristics of this specific population, this research holds significant value for educators and researchers in SLA. It can inform curriculum development, instructional practices, and ultimately bridge the gap in understanding L2 learners' preparedness for higher education programs.

## Methodology

### Research Questions

The present study is guided by the following main research questions:

1. Does the rank-frequency distribution of words in a corpus of L2 Hebrew writing produced by Arabic-speaking learners conform to Zipf's law, exhibiting a predominance of highly frequent function words and a greater proportion of lower frequency content words?
2. What is the productive Hebrew vocabulary size of Arabic speakers entering higher education programs, and how does it compare to established written and spoken Hebrew frequency lists?

### Corpus and Population Description

For data collection, the study utilizes a corpus consisting of 156 argumentative essays, written in Hebrew by native Arabic-speaking students. These essays were obtained from an entrance examination for an Israeli academic college, designed to assess the Hebrew language proficiency of candidates. The essays revolved around the argumentative topic of whether parents should

compensate children for assisting with household chores, including cleaning and caring for younger siblings. The choice of utilizing written essays as opposed to recall tests for assessing learners' productive vocabulary knowledge was made because, in the context of essay writing, learners have a more extensive window of time to retrieve words from memory. Furthermore, they can review and make corrections or alterations to their word choices if necessary (Milton, 2009; Leijten et al., 2010).

The participants in this study attended Israeli schools where Arabic is the primary language of instruction. Hebrew is formally introduced in the second grade (ages 8–9) with 2–5 hours of weekly instruction throughout their schooling (Manor & Watad, 2024). This culminates in a final exam for Arabic-speaking students (L2) in Hebrew. Although Hebrew is officially the language used in high school Hebrew classes, teachers (often native Arabic speakers) often switch to Arabic during lessons (Abu-Rabiah et al., 2023). Their expected level of proficiency in Hebrew is B1–B2 on the CEFR (Common European Framework of Reference for Languages) scale.

## Analysis Approach

In the analysis of written vocabulary, selecting the appropriate counting unit is pivotal (Nation & Coxhead, 2021; Brown et al., 2022; Maw et al., 2022). Various definitions, such as tokens, types, lemmas, flemmas, and word families, rely on how learners perceive relationships between word forms (Milton, 2009; Schmitt & Schmitt, 2020; Nation, 2021). Each word in a sample is a token, regardless of repetition, making the sample's length the count of its tokens. Types denote different (or unique) words in a sample (Jarvis & Hashimoto, 2021). Repetitions of the same tokens are counted as a single type. Tokens were not used due to susceptibility to text length and time constraints on learners, and types were avoided to prevent inflating vocabulary size counts, as noted by some researchers (Nation & Coxhead, 2021).

In this research the lemma was chosen as the counting unit. Lemma comprises diverse inflections of the same root within the same part of speech. For instance, 'מדריך' (a guide.masculine) and 'מדריכה' (a guide.feminine) are regarded as a single lemma. This choice is based on the assumption that learners are more likely to know and use the inflections within a lemma than the derivations required to form a word family (Schmitt & Schmitt, 2020). Using lemmas as a counting unit is common in studies involving second language learners (Kremmel, 2016; Treffers-Daller et al., 2018; Jarvis & Hashimoto, 2021). Compared to larger units like flemma or word family, lemma provides greater precision and involves fewer assumptions about a learner's vocabulary knowledge (Gablasova & Brezina, 2021). Word families, which encompass both

inflections and derivations of a root, may overestimate learners' knowledge. Therefore, the lemma was selected to avoid such overestimation and ensure more accurate measurement.

Unlike English, Hebrew allows multiple word forms (namely, a dictionary entry) to exist within a single written word (orthographic word) (Abu-Rabiah, 2020, 2022, 2023). For example, the word "בעתיד" 'ba'atid' includes the following word forms: 'be', ['ha'], and ''atid' (in [the] future). To ensure accurate analysis and comparison with other languages, we used Dicta's (Analytical tools for Hebrew texts) morphology tool to add vocalization (nikud) to each essay's words based on context. This allowed us to distinguish homographs (words with the same spelling but different meanings) and count all individual word forms within each orthographic word. This is especially relevant in Hebrew due to its consonantal writing system (Cook, 2016), where around 23% of words are homographic when considered in isolation from their surrounding context (Shimron & Sivan, 1994). Following this automated process, the corpus was manually annotated to accurately identify and categorize each word by its corresponding lemma.

The essays in the corpus display notable length variation, ranging from 34 to 267 orthographic words, with an average of 116. The total orthographic words in the corpus amount to 18,054, representing 27,407 word forms. Length control is unnecessary in this study, as the focus is not on comparing vocabulary size among individuals; instead, all essays are treated as one population. Opting for lemmas as the counting unit.

To delineate the productive vocabulary characteristics of written Hebrew among Arabic-speaking individuals, this study employed a comparative analysis with established Hebrew frequency lists. The primary reference was the heTenTen21 corpus (Hebrew Web Corpus 2021), which offers a frequency lemma list for written Hebrew. However, the open-access version of this list is limited to the top 1,000 most frequent lemmas. While the primary focus of the list is on words, it is worth noting that punctuation marks are also included. To assess the vocabulary size of the research population, it was necessary to supplement this with a second resource containing a broader range of lemmas.

As such, the Frequency Dictionary of Spoken Hebrew (FDOSH) (Pinto, 2018) was utilized for additional comparison. FDOSH, also based on lemmas, provides a compilation of the most frequently occurring items in conversational Modern Hebrew, derived from a corpus of film subtitles. Although FDOSH focuses on spoken rather than written Hebrew, it serves as a foundational resource for learners aiming to achieve conversational proficiency across diverse contexts. Its subtitle-based approach is well-justified, as prior studies have demonstrated that subtitle corpora closely mirror conventional spoken language (Pinto, 2018). Together, these resources enabled a more comprehensive examination of the learners' productive vocabulary.

As an initial exploration of vocabulary size in Hebrew across both L1 and L2 contexts, this study analyzes the entire corpus of 156 learner essays as a unified dataset, rather than examining individual learners. This corpus-driven approach, aligned with prior research (e.g., Milton, 2006), aims to characterize the overall productive vocabulary knowledge of the learner sample. However, future studies could build on these findings by incorporating individual-level analyses.

## Results and Discussion

1. Does the rank-frequency distribution of words in a corpus of L2 Hebrew writing produced by Arabic-speaking learners conform to Zipf's law, exhibiting a predominance of highly frequent function words and a greater proportion of lower frequency content words?

The frequency of word usage can provide valuable insights into the characteristics of a language. Table 1 displays the frequency and percentage of occurrence of lemmas in the corpus, sorted from most frequent to least frequent. Notably, the 20 most frequent lemmas on the left side of the table collectively account for nearly 59% of all word form occurrences in the corpus, while the 20 least frequent lemmas on the right side constitute less than one percent. This discrepancy signifies that the 20 most frequent lemmas are approximately 841 times more common than the 20 least frequent lemmas. Such a distribution pattern is consistent with findings in the English language, where high-frequency lemmas exhibit substantial coverage gains, while lower-frequency words exhibit diminishing gains (Schmitt & Schmitt, 2020).

For instance, the definite article 'ה' ("ha") occurs 2,977 times, covering nearly 11% of all word forms in the corpus, whereas each of the 20 least frequent lemmas appears only once, collectively covering less than half a percent of all word forms. Interestingly, there are 421 lemmas that occur only once in the corpus, and together, they account for no more than 1.54% of the total word forms. This underscores the notion that frequency serves as a reliable indicator of vocabulary utility at higher-frequency levels, but its predictiveness diminishes as one ventures into lower-frequency vocabulary in any language (Schmitt & Schmitt, 2020).

**Table 1**

*Frequency and Percentage of the Top 20 Most and Least Frequent Lemmas in the Corpus*

| Order | Lemma | Translation | Occurrences | Percentage of the Full Corpus | Order | Lemma | Translation | Occurrences | Percentage of the Full Corpus |
|---|---|---|---|---|---|---|---|---|---|
| 1 | ה | the | 2977 | 10.86% | 1004 | אסף | gathered | 1 | 0.0007% |
| 2 | הם | they.plural.masculine | 1408 | 5.14% | 1005 | בזבז | squandered | 1 | 0.0007% |
| 3 | ו | and | 1353 | 4.94% | 1006 | אתן | you.plural.feminine | 1 | 0.0007% |
| 4 | של | of | 1274 | 4.65% | 1007 | גורם | cause | 1 | 0.0007% |
| 5 | ל | to | 1255 | 4.58% | 1008 | הגן | defended | 1 | 0.0007% |
| 6 | ב | in\at | 981 | 3.58% | 1009 | הסתדר | arranged oneself | 1 | 0.0007% |
| 7 | הורה | parent | 800 | 2.92% | 1010 | הגשים | fulfilled | 1 | 0.0007% |
| 8 | ילד | boy | 779 | 2.84% | 1011 | הסתכלות | observing | 1 | 0.0007% |
| 9 | בית | home | 633 | 2.31% | 1012 | הדגמה | demonstration | 1 | 0.0007% |
| 10 | ש | that | 548 | 2.00% | 1013 | העיף | threw | 1 | 0.0007% |
| 11 | על | on | 540 | 1.97% | 1014 | אור | light | 1 | 0.0007% |
| 12 | אני | I | 515 | 1.88% | 1015 | הפחית | subtracted | 1 | 0.0007% |
| 13 | עזר | helped | 502 | 1.83% | 1016 | הובן | understood | 1 | 0.0007% |
| 14 | הוא | he.singular.masculine | 403 | 1.47% | 1017 | אקראי | random | 1 | 0.0007% |
| 15 | את | you.singular.feminine | 399 | 1.46% | 1018 | הוגדר | defined | 1 | 0.0007% |
| 16 | זה | this.feminine | 387 | 1.41% | 1019 | בטוח | safe | 1 | 0.0007% |
| 17 | לא | no | 378 | 1.38% | 1020 | בסיסי | basic | 1 | 0.0007% |
| 18 | אנחנו | we | 334 | 1.22% | 1021 | אבד | lost | 1 | 0.0007% |
| 19 | עבודה | work | 332 | 1.21% | 1022 | בעקבות | as a result of | 1 | 0.0007% |
| 20 | כל | all | 331 | 1.21% | 1023 | בנאדם | human being | 1 | 0.0007% |
| Grand Total | 20 | | 16129 | 58.85% | Grand Total | 20 | | 20 | 0.015% |

The distinction between the most frequent words and less frequent words plays a pivotal role in understanding the dynamics of vocabulary and its role in language. Vocabulary serves as the link between the tangible world and linguistic expression, connecting meanings with linguistic forms (Schmitt & Schmitt, 2020). It is primarily composed of content or lexical words, which encompass nouns, verbs, adjectives, and adverbs (e.g., 'car' and 'fresh') and are easily expandable (Cook, 2016). In contrast, function or structure words, including articles, prepositions, pronouns, and conjunctions (e.g., 'in' and 'but'), appear to carry minimal inherent meaning but are indispensable for constructing grammatically and semantically coherent language (Schmitt & Schmitt, 2020). Function words are limited in number and rarely expand (Milton, 2009; Cook, 2016). Content words provide the primary semantic content of sentences and are known as lexical words or open-class words, while function words serve as connectors between content words, labeled as grammatical words or closed class words (Abdalla et al., 2010; Vejdemo & Hörber, 2016).

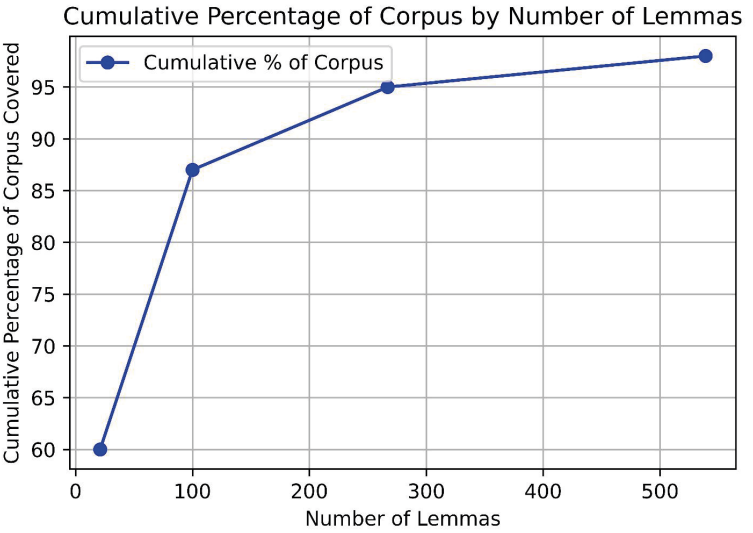
In actual language usage, function words tend to be the most frequent words across various corpora (Milton, 2009; Schmitt & Schmitt, 2020). In the current corpus, approximately 75% of the 20 most frequent lemmas are function words, while only 10% of the 20 least frequent lemmas belong to this category. The ubiquity of function words is observed consistently in corpora, irrespective of the general or specific nature of the language in the corpus and whether it is in written or spoken form. Content words, on the other hand, are context-specific and influenced by the thematic focus of the corpus (Schmitt & Schmitt, 2020). In the present corpus, which revolves around the topic of whether parents should provide compensation to children for assisting with household chores such as cleaning and caring for younger siblings, the frequent content words are inherently linked to this theme. Examples of these content words include 'הורה' ("hore") (a parent), 'ילד' ("yeled") (a son), 'בית' ("bayit") (a home), and 'עבודה' ("'avoda") (a work).

The examination of word frequency distribution, and its adherence to established linguistic patterns such as Zipf's law, is a crucial part of our analysis. A comprehensive survey of the entire corpus reveals that the distribution of lemma frequencies follows a pattern reminiscent of standard language distributions, wherein certain words are employed significantly more frequently than others (see Figure 1). This observed frequency distribution adheres to Zipf's law, a well-documented linguistic phenomenon. Zipf's law in corpus linguistics states that word frequency is inversely proportional to its rank in large language corpora (Ha et al., 2002). Under Zipf's law, a limited number of high-frequency words are exceedingly common, with their frequency declining exponentially as one progresses through the vocabulary, resulting in the majority of words being relatively rare (Schmitt & Schmitt, 2020). While initially observed for single words, research shows it applies more accurately when combining words

and n-gram phrases (Ha et al., 2002). In addition, the law extends beyond frequency, with Zipf's "law-of-meaning" suggesting that more frequent words tend to have more meanings (Ilgen & Karaoglan, 2007). In the current corpus, the top 21 most frequent lemmas collectively constitute 60% of the corpus, the top 100 encompass 87%, the top 267 cover 95%, and the top 539 account for 98% of the corpus.

**Figure 1**

*Word Frequency Distribution*




Zipf's law postulates that in a natural language corpus, the frequency of the first most common word is approximately twice that of the second most common word, which, in turn, is approximately twice that of the fourth most common word (Milton, 2009). In our current corpus, this relationship holds true, where the most frequent word, the definite article 'ה,' ("ha") occurs approximately twice as often as the second word 'הם' ("hem") (they). It is important to note that while Zipf's law is not without its imperfections (Milton, 2009), it underscores the idea that the most vital and practical vocabulary is concentrated among the most frequent words (Schmitt & Schmitt, 2020). This distribution pattern, encapsulated by Zipf's law, is not exclusive to English but holds true in numerous other languages, including Greek (Hatzigeorgiu et al., 2001), French (Cobb & Horst, 2004), and many more (Jayaram & Vidya, 2008).

2. What is the productive Hebrew vocabulary size of Arabic speakers entering higher education programs, and how does it compare to established written and spoken Hebrew frequency lists?

To quantify the productive Hebrew vocabulary size of the 156 Arabic-speaking learners in the study corpus, a process of lemmatization and de-duplication was implemented. Specifically, all inflected word form tokens were reduced to their corresponding lemmas. Subsequently, all instances of the same lemma type were collapsed, such that each distinct lemma was represented only once in the final vocabulary list, regardless of its frequency of occurrence in the corpus. This procedure yielded a total of 1,023 unique lemma types, constituting the productive vocabulary size estimate for this population of learners on the cusp of entering Hebrew-medium higher education programs.

The rationale for adopting this method of lemma-based vocabulary counting was to provide a conservative yet psycholinguistically motivated estimate of the learners' vocabulary knowledge. By focusing on distinct lexeme types rather than tokens, the analysis circumvented inflectional variation and generated a vocabulary size metric that more closely approximates the number of lexical stems productively commanded by these learners. This lemmatized vocabulary list therefore represents the core productive lexicon deployed by Arabic-speaking students as they prepare to engage with Hebrew as the language of academic instruction and scholarly content.

The average learner in our study had approximately 11 years of Hebrew instruction (1,188–1,980 hours). This translates to acquiring less than one productive lemma per lesson. This falls short of receptive vocabulary acquisition estimates (around 6.5–8 lemmas per lesson) reported by Vassiliu (2001). Interestingly, this size is comparable to the receptive vocabulary size reported in other studies: 1,654 word families for English majors in China (Miao, 2019) and 948–1,186 words for Spanish EFL learners (Montero-SaizAja, 2022). This suggests that our Arabic learners' productive vocabulary may be larger than the receptive vocabulary of learners in a foreign language context, potentially due to exposure to Hebrew outside classrooms. The 1,023 lemma size is also similar to the expected vocabulary size for English learners of French in the UK, where core vocabulary lists for CEFR B1 exams range from 1,000–1,500 items (Milton & Meara, 1998). However, it falls short of contexts like Hungary, where learners are expected to have 400 active lemmas by grade 8 (Krizsan, 2003, reported in Milton, 2009). In line with Laufer's (2017) suggestion that L2 learners typically grasp 2,000–4,000 word families despite extensive instruction (over 1,000 hours), a productive size of around 1,000 lemmas is not unexpected.

Research suggests receptive vocabulary size can be a good predictor of productive vocabulary, with a ratio ranging from 50% to 80% (Milton, 2009; Nizonkiza, 2016). Applying this range to our findings, the estimated receptive vocabulary size for our learners could be between 1,535 and 1,841 lemmas. However, this requires further investigation.

This study seeks to ascertain the extent of word overlap between the lemmas in the current L2 learner corpus and established written frequency lists

in both spoken and written Hebrew. It is well-acknowledged that word frequencies in oral and written language diverge (Lau et al., 2019), with conversational speech characterized by reduced formality and a less academic tone in contrast to written language (Milton, 2009).

**Table 2**

*Comparison of the 50 Most Frequent Lemmas in the Present Corpus with FDOSH Lemmas, Indicating Overlap*

| Order | Corpus Lemma | Translation | FDOSH Lemma | Overlap | Order | Corpus Lemma | FDOSH Lemma | Overlap |
|---|---|---|---|---|---|---|---|---|
| 1 | ה | the | הוא | Yes | 26 | צריך | אז | Yes |
| 2 | הם | they.plural. masculine | ה | No | 27 | יש | סדר | Yes |
| 3 | ו | and | את | Yes | 28 | היא | צריך | No |
| 4 | של | of | ל | Yes | 29 | קטן | רק | No |
| 5 | ל | to | לא | Yes | 30 | מ | חשב | Yes |
| 6 | ב | in\at | זה | Yes | 31 | טוב | כאן | Yes |
| 7 | הורה | parent | ב | No | 32 | אח | הלך | No |
| 8 | ילד | boy | של | No | 33 | היה | דבר | Yes |
| 9 | בית | home | ש | No | 34 | או | איש | Yes |
| 10 | ש | that | היה | Yes | 35 | שמר | אל | No |
| 11 | על | on | מה | Yes | 36 | שילם | כך | No |
| 12 | אני | I | ו | No | 37 | אמא | יותר | No |
| 13 | עזר | helped | על | No | 38 | רצה | שם | Yes |
| 14 | הוא | he.singular. masculine | כול | Yes | 39 | הרבה | יכול | No |
| 15 | את | you.singular. feminine | ידע | Yes | 40 | אם | ראה | Yes |
| 16 | זה | this.feminine | כן | Yes | 41 | זמן | עכשיו | Yes |
| 17 | לא | no | מ | Yes | 42 | משפחה | אחד | No |
| 18 | אנחנו | we | יש | No | 43 | כסף | משהו | No |
| 19 | עבודה | work | עשה | No | 44 | יכול | למה | Yes |
| 20 | כל | all | אבל | No | 45 | ניקה | בא | No |
| 21 | עזרה | help | טוב | No | 46 | נתן | זאת | No |
| 22 | כי | because | רצה | No | 47 | דעת | או | No |

***Table 2 continued***

| Order | Corpus Lemma | Translation | FDOSH Lemma | Overlap | Order | Corpus Lemma | FDOSH Lemma | Overlap |
|---|---|---|---|---|---|---|---|---|
| 23 | דבר | thing | אם | Yes | 48 | בן | זמן | No |
| 24 | עשה | did | עם | Yes | 49 | מה | נכון | Yes |
| 25 | גם | and | אמר | No | 50 | כמו | כמו | Yes |
| Overlap | | | | | | | | 50% |

When comparing the 50 most frequent lemmas in the present corpus with the 50 most frequent lemmas in the FDOSH (Frequency Dictionary of Spoken Hebrew), 50 percent overlap is observed (see Table 2). This implies that half of the 50 most frequent lemmas in the written Hebrew utilized by Arab learners coincide with the 50 most frequent lemmas in spoken Hebrew. A similar pattern of overlap emerges when considering all 1,023 lemmas in the current corpus against the 5,000 lemmas in FDOSH (see Table 3). Specifically, half of the lemmas in the corpus are among the top 1,000 most frequent lemmas in spoken Hebrew, while 30 percent of the lemmas fall within the remaining 4,000 lemmas, distributed as 15 percent in the 2,000 lemmas, 7 percent in the 3,000 lemmas, and 4 percent in both the 4,000 and 5,000 lemmas. In essence, 80% of the 1,023 lemmas in our current corpus are included in the 5,000 lemmas of FDOSH. The additional 20% of the lemmas do not appear in the FDOSH's 5,000 most frequent lemmas. This set of words encompasses not only content words like 'ניקיון' ("nikayon") (cleaning), 'הרהר' ("hirher") (ponders), and 'חירות' ("ẖerut") (freedom) but also function words like 'אנחנו' ("anaẖnu") (we).

**Table 3**

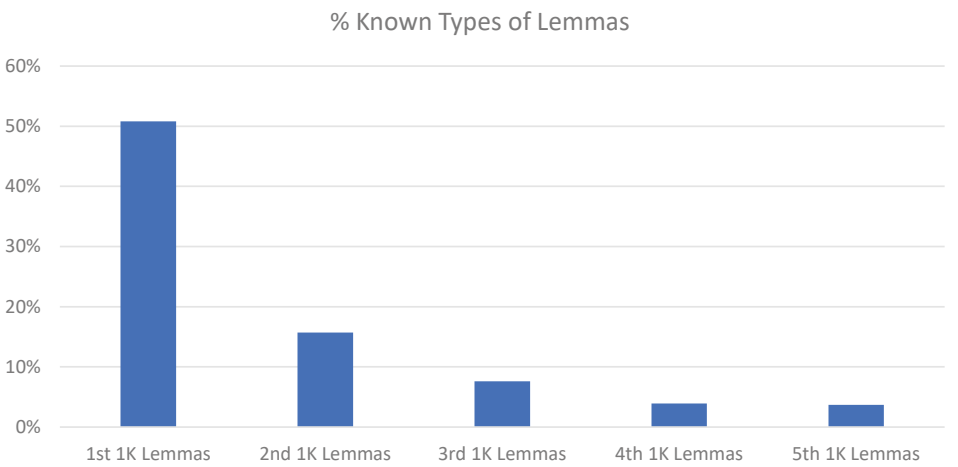
*Comparison of Lemmas in the Current Corpus and FDOSH*

| Total Lemmas in the Current Corpus | FDOSH Lemma Frequency | Overlapped Lemmas | Percentage of Overlap: (Overlapped Lemmas \ Total Lemmas in the Current Corpus) | Additional Overlap Percentage from Previous Frequency Group |
|---|---|---|---|---|
| 1023 | 1000 | 508 | 50% | – |
| | 2000 | 665 | 65% | 15% |
| | 3000 | 741 | 72% | 7% |
| | 4000 | 780 | 76% | 4% |
| | 5000 | 817 | 80% | 4% |
| | Out of the list | 206 | – | – |
| | Total | 1023 | 80% | – |

When we segment the FDOSH dataset into five distinct frequency bands, each consisting of 1,000 (1k) lemmas, and examine the occurrence of the 1,023 lemmas from the current corpus within these bands, a discernible pattern emerges. This pattern aligns with the well-documented concept of a typical learner's vocabulary knowledge, in which individuals tend to possess a more extensive command of high-frequency words compared to their familiarity with lower-frequency ones (Meara, 1992; Kweon & Kim, 2008 but cf. Crossley et al., 2019) and less sophisticated words (Abu-Rabiah, 2024). This distribution of lemmas within the frequency bands exhibits a characteristic profile, as depicted in the following graphical representation (see Figure 2). Notably, the learner's knowledge is notably robust in the higher-frequency columns on the left side, while it diminishes as one progresses towards the less frequent columns on the right side. This distinctive downward slope in vocabulary knowledge, from left to right, underscores the common observation in language acquisition research (Milton, 2009). This phenomenon highlights that learners tend to excel in acquiring and employing words that are encountered more frequently in spoken and written language, whereas their proficiency gradually declines as they confront less common and less frequently encountered lemmas.

**Figure 2**

*Distribution of Lemmas in the Current Corpus across 5 Frequency Bands, Each Consisting of 1000 (1k) Lemmas, According to FDOSH*




A comparison between the 1,023 lemmas in the current corpus and a shorter list of the 1,000 most frequent lemmas in written Hebrew, as documented in the heTenTen21 corpus, has yielded significant findings (see Table 4). The analysis identified 511 lemmas common to both datasets, representing a 50% overlap.

**Table 4**

*Comparison of Lemmas in the Current Corpus and heTenTen21, Indicating Overlap and Additional Overlap by Frequency Band*

| Total Lemmas in the Current Corpus | heTenTen21 Lemma Frequency | Overlapped Lemmas | Percentage of Overlap (Overlapped Lemmas / Total Lemmas in the Current Corpus) | Percentage of Additional Overlap from the Previous Frequency Band |
|---|---|---|---|---|
| 1023 | 50 | 45 | 4% | – |
| | 100 | 93 | 9% | 5% |
| | 500 | 343 | 34% | 25% |
| | 1000 | 511 | 50% | 16% |
| | Out of the list | 31 | – | – |
| | Total | 1023 | 50% | – |

A more in-depth examination focusing on the 50 most frequent words indicated that 45 lemmas were common to both datasets, constituting a notable 90% concordance. This level of agreement surpasses the level observed when contrasting these lemmas with the FDOSH dataset, which exhibited a 50% overlap. Similarly, the scrutiny of the 100 most frequent words unveiled 93 shared lemmas, signifying a significant 93% concurrence, compared to the 89% alignment with the FDOSH dataset.

When juxtaposing the initial 1,000 most frequent lemmas of heTenTen21 with the FDOSH dataset, a 58% overlap is observed. Furthermore, when evaluating the 1,000 most frequent lemmas of heTenTen21 against the 5,000 most frequent lemmas of FDOSH, a significant 91% level of overlap is recorded. Importantly, it is worth noting that this degree of overlap may potentially be even higher in practice, given the possible imperfections in the cleanliness of the heTenTen21 corpus, as previously acknowledged.

Collectively, these findings emphasize the significant overlap among the 1,000 most common lemmas in both written and spoken Hebrew, extending to the written Hebrew of Arabic speakers. It is evident that a considerable portion of the written vocabulary utilized by Arabic-speaking learners of Hebrew relies on these high-frequency words. However, their command of less common vocabulary appears to be more restricted. Half of their productive vocabulary is derived from the 1,000 most frequent written and spoken words in Hebrew. The majority of these words do not include academic vocabulary, suggesting that they may not be adequately prepared for higher education. This limitation could impede their ability to articulate nuanced concepts and navigate the intricacies of academic writing.

## Conclusions

This study investigated the vocabulary size and characteristics of L2 learners entering higher education, with a specific focus on Arabic speakers entering Hebrew-medium programs. By analyzing 156 argumentative essays written in Hebrew, the research revealed a productive vocabulary size of approximately 1,023 lemmas, with a distribution following Zipf's law. Notably, despite exceeding 1,000 hours of instruction, learners in contexts where the target language is used outside classrooms may still acquire a limited productive vocabulary (around 1,000 lemmas).

The analysis also revealed a significant overlap between the learners' vocabulary and established Hebrew frequency lists, such as heTenTen21 and FDOSH, which encompass both written and spoken language. In addition, the learners exhibited a typical vocabulary profile, demonstrating greater use of lemmas from the 1k frequency band compared to the 2k band, with usage progressively declining in subsequent frequency bands. The findings further indicated that learners demonstrated a strong command of high-frequency words, which were crucial for everyday communication in Hebrew. However, their mastery of less frequent vocabulary, essential for expressing nuanced ideas in academic contexts, was notably limited.

Based on these findings, it is recommended to deliberately teach academic vocabulary at the beginning of the academic journey to support learners in producing discipline-specific language and engaging effectively with academic content. Additionally, frequency lists can be utilized to assess and guide vocabulary instruction, helping to set realistic expectations for language learners. Future research should explore the gap between L2 learners' receptive and productive vocabulary sizes, particularly as they transition to higher education. Furthermore, examining how productive vocabulary size varies across different proficiency levels may yield valuable insights into the developmental trajectory of L2 lexical knowledge.

## References

Abdalla, F., Robb, M. P., & Al-Shatti, T. (2010). Stuttering and lexical category in adult Arabic speakers. *Clinical Linguistics & Phonetics, 24*(1), 70–81.

Abu-Rabiah, E. (2020). Lexical measures for testing progress in Hebrew as Arab students' L2. *Journal of Language and Linguistic Studies, 16*(3), 1096–1114. https://doi.org/10.17263/jlls.803551

Abu-Rabiah, E. (2023). Evaluating L2 vocabulary development features using lexical density and lexical diversity measures. *LLT Journal: A Journal on Language and Language Teaching, 26*(1), 168–182. https://doi.org/10.24071/llt.v26i1.5841

Abu-Rabiah, E. (2024). Exploring lexical sophistication in second language: An analysis of vocabulary using a word-rating method. *Theory and Practice of Second Language Acquisition, 10*(2), 1–24. https://doi.org/10.31261/TAPSLA.15623

Abu-Rabiah, E., Gafter, R., & Henkin, R. (2023). Where syntactic interference persists: The case of Hebrew written by native Arabic speakers. *L1-Educational Studies in Language and Literature, 23*, 1–19. https://doi.org/10.21248/l1esll.2023.23.1.399

Ambridge, B., Kidd, E., Rowland, C. F., & Theakston, A. L. (2015). The ubiquity of frequency effects in first language acquisition. *Journal of Child Language, 42*(2), 239–273.

Brown, D., Stoeckel, T., McLean, S., & Stewart, J. (2022). The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence. *Applied Linguistics, 43*(3), 596–602.

Brysbaert, M., Stevens, M., Mandera, P., & Keuleers, E. (2016). How many words do we know? Practical estimates of vocabulary size dependent on word definition, the degree of language input and the participant's age. *Frontiers in Psychology, 7*, 1116, 1–11.

Cobb, T., & Horst, M. (2004). Is there room for an AWL in French? In T. Cobb & M. Horst (Eds.), *Vocabulary in a second language: Selection, acquisition, and testing* (pp. 15–38). John Benjamins B.V.

Cook, V. (2016). *Second language learning and language teaching*. Routledge.

Craig, H. (2011). Shakespeare's vocabulary: Myth and reality. *Shakespeare Quarterly, 62*(1), 53–74.

Crossley, S. A., Skalicky, S., Kyle, K., & Monteiro, K. (2019). Absolute frequency effects in second language lexical acquisition. *Studies in Second Language Acquisition, 41*(4), 721–744.

DICTA. Analytical tools for Hebrew texts: https://morph-analysis.dicta.org.il/

Fan, M. (2000). How big is the gap and how to narrow it? An investigation into the active and passive vocabulary knowledge of L2 learners. *RELC Journal, 31*(2), 105–119.

Gablasova, D., & Brezina, V. (2021). Words that matter in L2 research and pedagogy: A corpus-linguistics perspective. *Studies in Second Language Acquisition, 43*(5), 958–961.

Ha, L., Sicilia-Garcia, E., Ming, J., Smith, F. (2002). Extension of Zipf's law to words and phrases. In *COLING 2002: Proceedings of the 19th International Conference on Computational Linguistics*.

Hatzigeorgiu, N., Mikros, G., & Carayannis, G. (2001). Word length, word frequencies and Zipf's law in the Greek language. *Journal of Quantitative Linguistics, 8*(3), 175–185.

heTenTen21 corpus: Hebrew Web corpus 2021: https://www.sketchengine.eu/hetenten-hebrew-corpus/

Ilgen, B., & Karaoglan, B. (2007). Investigation of Zipf's "law-of-meaning" on Turkish corpora. In *22nd International Symposium on Computer and Information Sciences* (ISCIS 2007) (pp. 1–6).

Jayaram, B. D., & Vidya, M. N. (2008). Zipf's law for Indian languages. *Journal of Quantitative Linguistics, 15*(4), 293–317.

Jarvis, S., & Hashimoto, B. J. (2021). How operationalizations of word types affect measures of lexical diversity. *International Journal of Learner Corpus Research, 7*(1), 163–194.

Kremmel, B. (2016), Word families and frequency bands in vocabulary tests: Challenging conventions. *TESOL Quarterly, 50*(40), 976–987.

Kweon, S. O., & Kim, H.R. (2008). Beyond raw frequency: Incidental vocabulary acquisition in extensive reading. *Reading in a Foreign Language, 20*(2), 191–215.

Lau, S. H., Huang, Y., Ferreira, V. S., & Vul, E. (2019). Perceptual features predict word frequency asymmetry across modalities. *Attention, Perception, & Psychophysics, 81*, 1076–1087.

Laufer, B. (2017). From word parts to full texts: Searching for effective methods of vocabulary learning. *Language Teaching Research, 21*(1), 5–11.

Laufer, B., & Yano, Y. (2001). Understanding unfamiliar words in a text: Do L2 learners understand how much they don't understand? *Reading in a Foreign Language, 13*(2), 549–566.

Leijten, M., Van Waes, L., & Ransdell, S. (2010). Correcting text production errors: Isolating the effects of writing mode from error span, input mode, and lexicality. *Written Communication, 27*(2), 189–227.

Lenko-Szymanska, A. (2002). How to trace the growth in learners' active vocabulary? A corpus-based study. In B. Kettermann & M. Georg (Eds.), *Teaching and learning by doing corpus analysis* (pp. 217–230). Rodopi.

Llach, M. P. A., & Gallego, M. T. (2009). Examining the relationship between receptive vocabulary size and written skills of primary school learners. *Atlantis, 31*(1), 129–147.

Mairano, P., & Santiago, F. (2020). What vocabulary size tells us about pronunciation skills: Issues in assessing L2 learners. *Journal of French Language Studies, 30*(2), 141–160.

Manor, R., & Watad, A. (2024). The complexity of teaching Hebrew in Israel's Arab school system. *Language Teaching*, *58*(1), 14–26.

Maw, T. M. M., Clenton, J., & Higginbotham, G. (2022). Investigating whether a flemma count is a more distinctive measurement of LD. *Assessing Writing, 53*, 100640.

Meara, P. M. (1992). *EFL vocabulary tests*. University College Swansea: Center for Applied Language Studies.

Miao, J. (2019). A study on English majors' vocabulary size and its correlation with scores on test for English majors grade-four. In *2nd International Conference on Contemporary Education, Social Sciences and Ecological Studies* (pp. 202–208). Atlantis Press.

Milton, J. (2006). Language lite? Learning French vocabulary in school. *Journal of French Language Studies, 16*(2), 187–205.

Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters.

Milton, J., & Meara, P. (1998). Are the British really bad at learning foreign languages?. *Language Learning Journal, 18*(1), 68–76.

Milton, J., & Treffers-Daller, J. (2013). Vocabulary size revisited: The link between vocabulary size and academic achievement. *Applied Linguistics Review, 4*(1), 151–172.

Milton, J., Wade, J. & Hopkins, N. (2010). Aural word recognition and oral competence in English as a Foreign Language. In R. Chacon-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (83–98). Multilingual Matters.

Montero-SaizAja, A. (2022). The productive vocabulary of multimodal and unimodal English as a Foreign Language learners. *Porta Linguarum: Revista Internacional de Didáctica de Las Lenguas Extranjeras. Monográfico, 4*, 139–153.

Nation, I. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review, 63*(1), 59–82.

Nation, I. S. P., & Coxhead, A. (2021). *Measuring native-speaker vocabulary size*. John Benjamins Publishing Company.

Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge.

Nation, P. (2021). Thoughts on word families. *Studies in Second Language Acquisition, 43*(5), 969–972.

Nizonkiza, D. (2016). First-year university students' receptive and productive use of academic vocabulary. *Stellenbosch Papers in Linguistics, 45*, 169–187.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76*(4), 763–782.

Pinto, J. D. (2018). *Creating a frequency dictionary of spoken Hebrew: A reproducible use of technology to overcome scarcity of data*. Doctoral dissertation, University of Texas at Austin.

Richards, B. J., & Malvern, D. D. (1997). *Quantifying lexical diversity in the study of language development*. University of Reading, Faculty of Education and Community Studies.

Rowe, M. L., Raudenbush, S. W., & Goldin-Meadow, S. (2012). The pace of vocabulary growth helps predict later vocabulary skill. *Child Development, 83*(2), 508–525.

Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning, 64*(4), 913–951.

Schmitt, N., & Schmitt, D. (2020). *Vocabulary in language teaching*. Cambridge University Press.

Shimron, J., & Sivan, T. (1994). Reading proficiency and orthography evidence from Hebrew and English. *Language Learning, 44*(1), 5–27.

Treffers-Daller, J., Parslow, P., & Williams, S. (2018). Back to basics: How measures of lexical diversity can help discriminate between CEFR levels. *Applied Linguistics, 39*(3), 302–327.

Van Zeeland, H., & Schmitt, N. (2013). Lexical coverage in L1 and L2 listening comprehension: The same or different from reading comprehension? *Applied Linguistics, 34*(4), 457–479.

Vassiliu, P. (2001). *Lexical input and uptake in the low level EFL classroom*. Unpublished PhD dissertation, University of Wales Swansea.

Vejdemo, S., & Hörberg, T. (2016). Semantic factors predict the rate of lexical replacement of content words. *PloS ONE, 11*(1), e0147924.

Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition, 30*(1), 79–95.

Webb, S., & Rodgers, M. P. (2009). The lexical coverage of movies. *Applied Linguistics, 30*(3), 407–427.

Webb, S., & Rodgers, M. P. (2009). Vocabulary demands of television programs. *Language Learning, 59*(2), 335–366.

Wesche, M., & Paribakht, T. S. (1996). Enhancing vocabulary acquisition through reading: A hierarchy of text-related exercise types. *The Canadian Modern Language Review, 52*, 154–178.

Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/23
https://doi.org/10.31261/TAPSLA.17111

CC BY SA

**Meihua Liu**
Tsinghua University, China
https://orcid.org/0000-0002-5951-4167

# Learning Style Preferences, Strategy Use and Chinese Undergraduate EFL Students' English Achievements

Abstract

Research shows that both learning style preference and strategy use are important factors affecting second/foreign language (SL/FL) learning, and that both may vary, as learners are different. Yet, studies on learning style preferences and their interaction with strategy in relation to SL/FL learning outcomes are limited. The present study thus examined Chinese undergraduate EFL (English as a FL) students' preferences for learning styles, use of learning strategies and their predictive effects on their English achievements. Analyses of 439 questionnaires revealed the following major findings: (1) the whole sample, regardless of gender, displayed no major preference for any learning style and generally preferred the kinesthetic style the most, followed by tactile, auditory, visual, group and individual styles respectively; (2) the whole sample, regardless of gender, demonstrated a medium to high level of use of the six types of learning strategies, and used metacognitive strategies the most frequently, followed by cognitive, social, compensation, memory and affective strategies, respectively; (3) cognitive strategy use significantly positively predicted all samples' English achievements, additionally, memory strategy use and the tactile style significantly predicted female students' English achievements, and (4) no significant difference was observed between genders in learning styles or strategy use except for group learning. These findings further pinpoint the importance of learning styles and strategy use in second/foreign language learning, which thus deserve continuous research.

*Keywords*: learning style, strategy use, achievement in English, gender difference, predictive effect

It is widely acknowledged that various individual factors such as motivation, emotion, learning style and strategy use affect second/foreign language (SL/FL) learning (Oxford, 1990; Sun et al., 2023). Consequently, individual factors have become an important research topic in SL/FL learning and acquisition. Of these

factors, learning style remains relatively less researched, although an increasingly greater number of studies have examined SL/FL learners' learning style preferences and their relations to SL/FL learning since the 1970s (Al Khatib & Ghosheh, 2013; Liu & Chen, 2024).

Research on learning styles in SL/FL learning primarily focuses on ESL/EFL (English as SL/FL) learners, which shows that ESL/EFL learners generally prefer kinesthetic and tactile styles the most, and group learning the least (e.g., Gao, 2016; Mozayan et al., 2013), displaying varying preferences for different styles based on individual factors such as gender, culture, and education (e.g., Alkahtani, 2016; Gao, 2016; Ha, 2019). As for the relationship between learning styles and English learning, the limited available literature reveals mixed findings (Akbarian et al., 2019; Malsawmkimi & Fanai, 2019). All these factors clearly justify more research on learning styles in relation to ESL/EFL learning outcomes.

Compared to learning styles, language learning strategies have been extensively researched in various SL/FL contexts. The results show that good learners tend to utilize more effective strategies and be more flexible in choosing more suitable strategies than poor learners (e.g., Lin et al., 2021; McMullen, 2009; Pongsukvajchakul, 2021). Nevertheless, when individual factors were considered, mixed findings have been found about the use of strategies and its relation to SL/FL learning outcomes (e.g., Gregersen & MacIntyre, 2014; Pongsukvajchakul, 2021). Hence, continuous research on language learning strategies is needed as well.

Moreover, insufficient research has been conducted to explore the interaction of learning styles with other individual factors (Lin et al., 2021; Zokaee et al., 2012). This also motivates the present research, which aims to examine Chinese university EFL students' preferences for learning styles, use of language learning strategies and their predictive effects on their English achievements.

## Literature Review

### Learning Styles

For decades, learning styles, or differences in how people learn, have caught the attention of educators and researchers as a way of meeting students' individual needs (Whitman, 2023). Keefe (1979, p. 4) defined learning styles as "characteristic cognitive, affective, and physiological behaviors that serve as relatively stable indicators of how learners perceive, interact with, and respond to the learning environment." Reid (1995) regarded learning styles as a student's natural, habitual, and preferred way of absorbing, processing, and retaining new knowledge.

Many theories have been proposed to explain learning styles, of which the most well-known include the Felder-Silverman learning/teaching style model, the Kolb experiential learning theory, Reid's perceptual learning style theory, and Fleming and Mills' sensory-based learning styles (Hawk & Shah, 2007; Whitman, 2023). Though each theory varies in how they describe learning styles and categorize learners, all believe that learners learn in different ways (Hawk & Shah, 2007). For example, Dunn and Price (1975) categorized learning styles as visual, tactile, and kinesthetic. Kolb (1984) grouped learning styles as diverging (perceiving input concretely and processing it reflectively), converging (perceiving input abstractly and processing it actively), assimilating (perceiving input abstractly and processing it reflectively), and accommodating (perceiving input concretely and processing it actively). Reid (1987) classified learning styles into six types: Auditory (learning through the oral-aural channel), visual (learning through the seeing/visual channel), kinesthetic (learning through experiential learning), tactile (learning through hands-on activities), individual learning (learning through working alone), and group learning (learning through working with others). To measure learning styles of non-native speakers of English, Reid (1987) developed the 30-item five-point-Likert Perceptual Learning Style Preference Questionnaire (PLSPQ) which covers these six categories, with five items for each category. Reid (1987) also provided three cut-off scores for major (38–50), minor (27–37), and negligible (24 or less) learning style preferences to analyze the data received from the PLSPQ.

Despite lacking consistent levels of validity and reliability (Pelegrín, 2020), PLSPQ is the most recent and widely used instrument for ESL/EFL learners and its categorization has been widely acknowledged (e.g., Akbarian et al., 2019; Alkahtani, 2016; Gao, 2016; Ha, 2019; Koglin, Arald, & Felicetti, 2021; Lethaby & Russell, 2020; Liu, 2017, 2023; Malsawmkimi & Fanai, 2019; Naserieh & Mohammad, 2013; Reid, 1987, 1995; Zokaee et al., 2012). These studies have shown that ESL/EFL students generally prefer kinesthetic and tactile learning styles the most, and group learning the least. For example, Peacock's (2001) study of 206 EFL students showed that the participants generally preferred kinesthetic and auditory learning styles the most, and individual or group learning the least. This was largely supported by Akbarian et al.'s (2019) study of 235 tertiary EFL learners from two Iranian universities, which showed that kinesthetic, auditory, visual, and tactile styles were the participants' major learning styles while individual and group ones were their minor styles, with group learning being the least preferred style. Nevertheless, different findings have also been revealed (Al Khatib & Ghosheh, 2013). Al Khatib and Ghosheh (2013) examined the preferred learning styles of 210 students of Al Ain University of Science and Technology in relation to gender, academic performance and field of study. The results showed that the students' major learning style preferences were auditory, visual and group styles, while kinesthetic,

tactile and individual styles were their minor styles. In addition, learning style preferences except for group learning preference did not vary by academic performance. Education students were commonly more tactile learners than those in other fields of study, while law students were largely more group learners and pharmacy students were ordinarily more individual learners than those in other fields. Gao (2016) found that almost all Chinese university students were multi-style learners, preferring tactile, visual, and kinesthetic styles the most, and group learning the least. On the contrary, Alkahtani (2016) found that Yanbu EFL students typically preferred auditory and group learning. Ha's (2019) research of 162 Vietnam university students also revealed that group learning was the most preferred style.

Additionally, studies show that male and female students have different preferences for learning styles. Melton (1990) discovered that female Chinese students preferred auditory and kinesthetic learning styles, while their male peers preferred tactile and individual learning styles. Hyland's (1993) study of 440 Japanese university students indicated that female students demonstrated a stronger preference for every learning style than males. This was supported by Mozayan et al.'s (2013) study of Iranian medical sciences students, which found that female students preferred five learning styles more strongly than males. Al Khatib and Ghosheh (2013) found significant gender differences in learning styles as well: Males were generally more auditory and tactile learners, whereas female students were on average more group learners. Akbarian et al. (2019) also found that male Iranian university students preferred the auditory style than females. In contrast, Zokaee et al.'s (2012) research of 54 Iranian EFL learners revealed no significant gender difference in learning style preferences.

All these findings generally reveal that learning style preferences differ with age, education, achievement, gender, culture, major field, and so on. Coupled with the relatively limited number of studies on learning styles in ESL/EFL learning and the large population of ESL/EFL learners, learning styles in relation to ESL/EFL learning outcomes deserves more research. This motivated the present research, which adopted Reid's (1987) classification and measure of learning styles because of the clarity and ease in implementing the PLSPQ and interpreting its results.

## Language Learning Strategy Use

The role of strategies in language learning has long been noticed and researched (Lee & Heinz, 2016; Oxford, 1990, 2017; Rubin, 1975; Thomas et al., 2022). Though no consensus about the definition of language learning strategy (LLS) has been reached, LLS refers to the general involvement (all behaviors, techniques, habits, methods, etc.) in language learning (Griffiths, 2020; Rubin, 1975) or the selective and intentional process in language learning (Oxford

& Cohen, 1992). Accordingly, Oxford (1990) grouped LLSs into six types: memory, cognitive, compensation, metacognitive, affective, and social, and developed the Strategy Inventory for Language Learning (SILL) to measure the use of these strategies. Oxford (1990) also provided three cut-off scores for a high (3.5–4.4), medium (2.5–3.4) and low (below 2.5) level of strategy use. As more definitions were proposed, Oxford (2017) reviewed them and summarized that language learning strategy can have many forms: it can be a technique, device, tool, or method; it can be a behavior; it can be a general tendency, or approach. Among the various forms of strategies identified in the definitions, the central feature is mental action or process (Oxford, 2017).

Oxford's influential works have inspired numerous studies on language learning strategies, which have used or adapted SILL in various SL/FL contexts (e.g., Lin et al., 2021; McMullen, 2009; Pongsukvajchakul, 2021; Rahman, 2020; Vimalakshan & Aziz, 2021). Salahshour, Sharifi, and Salahshour (2013) studied LLS use in 65 high school students via SILL and found that the students used metacognitive strategies the most frequently, and cognitive strategies the least frequently. Stracke (2016) administered SILL to 522 sixth graders and found that the participants had high use of social, affective and metacognitive strategies, and moderate use of cognitive strategies. The study also revealed significant differences in strategy use between students who perceived themselves capable of performing English tasks and self-regulating their learning, and those who did not. Rahman (2020) administered the SILL to 30 Saudi-Arabian university students and found that the students used metacognitive strategies the most frequently, followed by compensation, social, memory, cognitive, and affective strategies respectively. Pongsukvajchakul (2021) administered the SILL to 100 Thai undergraduate students, and found that the participants used social strategies the most often, followed by memory, compensatory, metacognitive, cognitive and affective strategies in English writing, respectively. The study also revealed significant differences in LLS use depending on field of study, year of study and writing frequency.

In another case, McMullen's (2009) study of 71 male and 94 female Saudi students via the SILL showed that female students used slightly more LLSs than male students. Salahshour et al. (2013) also found that females used learning strategies more frequently than males.

Concurrently, LLS use also proves to interact with many other variables in SL/FL learning, such as foreign language anxiety (e.g., Noormohamadi, 2009), self-efficacy (e.g., Stracke, 2016), and motivation (e.g., Lin et al., 2021).

All these findings indicate that individual learners' LLS use differs in terms of gender, age, beliefs, nationality/ethnicity, personality and so forth (Gregersen & MacIntyre 2014; Pongsukvajchakul, 2021). This justifies continuous research on LLS in SL/FL learning.

## Relationships between Learning Style Preferences, Strategy Use and Second/Foreign Language Achievements

Debates exist regarding the relationship between learning style and learners' academic performance or (perceived) intellectual aptitude. For example, Sun et al. (2023) did three experiments on American children and their teachers and parents, and found that learning style descriptions impacted parents' and teachers' thinking about children's intellectual aptitudes. Yet, Malsawmkimi and Fanai's (2019) research of 192 secondary school students found no correlation between students' academic achievements and learning styles, which was confirmed by Whitman (2023).

The research on learning styles and SL/FL outcomes is limited, and also reveals mixed findings. For example, Huang et al.'s (2018) study of 329 Chinese university students revealed no significant relationship between learning styles and the participants' English proficiency, while Ha (2019) found a significant relationship between Vietnam university students' learning styles and their English language proficiency. Similarly, Akbarian et al. (2019) found that Iranian university students' tactile style scores significantly correlated with vocabulary knowledge. Liu and Chen (2024) collected PLSPQ and strategy use data from 439 Chinese university students and found that learning styles were significantly positively related to and predicted the participants' English achievement.

Regarding strategy use and SL/FL learning, empirical studies reveal that good learners tend to use more effective strategies and choose different strategies to tackle different learning tasks while poor learners tend to use less effective strategies and are not so flexible in choosing different strategies (e.g., Oxford, 1990, 2017; Rubin, 1975; Salahshour et al., 2013; Vimalakshan & Aziz, 2021). Learners may also use different strategies simultaneously for different functions and in different orders (Gao & Hu, 2020). For example, Salahshour et al.'s (2013) study of 65 high school students via SILL found that proficient learners demonstrated significantly more strategy use, particularly of metacognitive and social strategies. This finding was generally consistent with that of other studies in similar contexts (e.g., Lin et al., 2021). Lin et al.'s (2021) study of 547 international university students indicated that language learning strategies significantly affected the students' literal comprehension in Chinese. Liu and Chen (2024) found that learning strategies were significantly positively related to and predicted students' English achievement.

Meanwhile, the available limited research on learning styles and strategy use generally show that the two are positively correlated with each other (Alkahtani, 2016; Gao, 2016; Lin et al., 2021; Zokaee et al., 2012). For example, Zokaee et al. (2012) revealed that Iranian university EFL students' specific learning styles correlated with specific vocabulary learning strategies. Alkahtani's (2016) research of 667 Yanbu EFL students revealed that significant correlations existed

between perceptual language learning styles and the use of language learning strategies, which was supported by the findings in Liu and Chen (2024). Gao's (2016) examination of learning styles and strategy use on 250 Chinese university EFL learners also revealed complex relationships between the two.

## Research Questions

As reviewed, language learning strategies are important factors that affect SL/FL learning, and it is likely that learning styles also play a role in SL/FL learning. Considering the limited number of studies on the two, and the mixed findings they have uncovered, the relationship between language learning strategies and learning styles deserves more research. Moreover, preferences for learning styles and use of language learning strategies vary by gender, education, nationality, culture and other individual characteristics. Coupled with rather inadequate research on learning styles in SL/FL learning, research on these two topics in relation to SL/FL learning is always worthwhile. Consequently, this research aimed to investigate Chinese undergraduate EFL students' preferences for learning styles, use of language learning strategies and their predictive effects on their English achievements. The following research questions were of particular interest:

1) What are the students' preferences for learning styles?
2) To what degree are the language learning strategies used by the students?
3) How are the students' learning style preferences related to their language learning strategy use?
4) How do the students' learning style preferences and language learning strategy use predict their English achievements?
5) What are the gender differences in learning style preferences, language learning strategy use, and their predictive effects on English achievements?

## Research Design

### Participants

The participants were 439 (209 male and 230 female) undergraduate EFL students from different Chinese universities, who had generally studied English for more than 10 years since primary school. With a mean age of 19.71 and

an age range of 17 to 22, these students came from three major disciplines: Engineering (286/65.1%, 178 male and 108 female), science (15/3.4%, 4 male and 11 female) and liberal arts (138/31.4%, 27 male and 111 female).

## Instruments

The participants in this study answered a background information questionnaire, the 30-item Perceptual Learning Style Preference Questionnaire (PLSPQ) and the 50-item Strategy Inventory for Language Learning (SILL). Both PLSPQ and SILL were 5-point Likert scales, ranging from "Never or almost never true of me/Strongly disagree" to "Always or almost always true of me/Strongly agree," with values 1–5 assigned to the descriptors respectively.

### *The PLSPQ*

The PLSPQ developed by Reid (1987) was adopted and achieved a reliability score of .86 in the present study, covering six types of learning styles: auditory, kinesthetic, visual, tactile, group and individual. Sample items included: "When the teacher tells me the instructions, I understand better," "I prefer to learn by doing something in class," and "I learn better by reading what the teacher writes on the chalkboard."

### *The SILL*

Since the present research aimed at examining students' learning style preferences and strategy use in general, the SILL designed by Oxford (1990) was utilized and achieved a reliability score of .95 in the present study, covering six types of learning strategies: memory, cognitive, compensation, metacognitive, social and affective. Sample items included: "I use rhymes to remember new English words," "I practice the sounds of English," "I read English without looking up every new word," and "I pay attention when someone is speaking English."

### *The Background Information Questionnaire*

This questionnaire aimed to collect information about the respondents such as gender, age, field of study, and year of study.

### *English Achievements*

Since the students were from different universities, they were asked to self-rate their overall English proficiency on a 10-point scale (from 1 being "very poor" to 10 "nativelike") as an indicator of their English achievements.

### Data Collection and Analysis

The study was approved by the Ethics Committee of the writer's institute. Then, all of the questionnaire items were translated into Mandarin Chinese, back-translated, and double-checked by a researcher proficient in both Chinese and English. Due to COVID-19, the survey was administered online. Students were encouraged to participate by their course teachers, yet all participation was voluntary. Finally, 439 valid questionnaires were collected within two weeks. The data was then analyzed via SPSS 20 and Mplus 7. Means and standard deviations were computed to determine preferences for learning styles and levels of strategy use; independent sample t-tests were run to examine differences between male and female students; correlation analyses were conducted to explore correlations between learning styles and strategy use; structural equation modelling (SEM) and multiple stepwise regression analyses were run to examine predictive effects of learning styles and strategy use on students' English achievements.

## Results

### Students' Preferences for Learning Styles and Use of Language Learning Strategies

As reported in Tables 1 and 2, skewness values for all the styles and strategies were between –1 and 1, indicating that all the styles and strategies had an approximately normal distribution. As shown in Table 1, the respondents scored on average from 34.27 to 37.15 on the six learning styles, meaning that they used all the styles as their minors. They scored the highest on the kinesthetic style (mean = 37.15), followed by tactile (mean = 36.76), auditory (mean = 36.57), visual (mean = 35.35), group (mean = 34.42) and individual learning (mean = 34.27) styles, respectively.

A similar pattern was observed for both male and female students who scored on average 33.51 (individual) to 36.98 (kinesthetic) and 33.60 (group) to 37.31 (kinesthetic) respectively on the six learning styles. However, male

students scored higher on auditory, visual and group styles and lower on the other three styles than their female peers. It should be noted that a significant difference occurred only in group learning ($t = 2.088$, $p = .037$), suggesting that male students preferred group learning significantly more than female respondents.

As shown in Table 2, the respondents on average scored 3.41 on affective, 3.42 on memory, 3.43 on compensation, 3.51 on social, 3.54 on cognitive and 3.76 on metacognitive strategies. This finding means that they used metacognitive strategies the most frequently, followed by cognitive, social, compensation, memory and affective strategies, respectively. It also means that the respondents had a high level, though on the low end, of metacognitive, cognitive and social strategy use, and a medium level, but on the high end, of compensation, memory and affective strategy use (Oxford, 1990).

A similar pattern was observed for both male and female students who scored on average 3.38 (affective) to 3.71 (metacognitive) and 3.44 (memory, compensation and affective) to 3.81 (metacognitive) respectively on the six types of learning strategies. Meanwhile, male students scored higher on affective but lower on all the other five types of strategies than their female peers. Nevertheless, no significant difference occurred in any type of strategies.

**Table 1**

*Means, Standard Deviations and t-test Results of Learning Styles*

| | | Auditory | Kinesthetic | Visual | Tactile | Individual | Group |
|---|---|---|---|---|---|---|---|
| Whole sample ($N$ = 439) | Mean | 36.57 | 37.15 | 35.35 | 36.76 | 34.27 | 34.42 |
| | SD | 6.55 | 6.55 | 6.81 | 6.44 | 8.74 | 8.66 |
| | skewness | –.416 | –.513 | –.061 | –.395 | –.274 | –.442 |
| Male students ($N$ = 209) | Mean | 36.63 | 36.98 | 35.67 | 36.59 | 33.51 | 35.32 |
| | SD | 6.60 | 6.71 | 7.27 | 6.68 | 9.13 | 8.63 |
| | skewness | –.233 | –.427 | –.095 | –.601 | –.247 | –.315 |
| Female students ($N$ = 230) | Mean | 36.50 | 37.31 | 35.06 | 36.90 | 34.97 | 33.60 |
| | SD | 6.51 | 6.41 | 6.36 | 6.23 | 8.33 | 8.61 |
| | skewness | –.251 | –.365 | –.389 | –.769 | –.167 | –.501 |
| t-test results | $t$ | .203 | –.538 | .943 | –.505 | –1.744 | 2.088* |
| | P | .839 | .591 | .346 | .614 | .082 | .037 |
| | Cohen's d | / | / | / | / | / | 0.20 |

*Note.* ** = $p < .01$; * = $p < .05$

Effect size of Cohen's d: small = d ≤ 0.2; medium = d = 0.5; large = d ≥ 0.8 (Cohen, 1988).

**Table 2**

*Means, Standard Deviations and t-test Results of Language Learning Strategies*

| | | Memory | Cognitive | Compensation | Metacognitive | Affective | Social |
|---|---|---|---|---|---|---|---|
| Whole sample (*N* = 439) | Mean | 3.42 | 3.54 | 3.43 | 3.76 | 3.41 | 3.51 |
| | SD | 0.70 | 0.67 | 0.56 | 0.69 | 0.73 | 0.83 |
| | skewness | –.246 | –.414 | –.246 | –.706 | –.213 | –.424 |
| Male students (*N* = 209) | Mean | 3.40 | 3.51 | 3.41 | 3.71 | 3.38 | 3.52 |
| | SD | .72 | .72 | .56 | .76 | .75 | .80 |
| | skewness | –.335 | –.501 | –.146 | –.471 | –.228 | –.613 |
| Female students (*N* = 230) | Mean | 3.44 | 3.56 | 3.44 | 3.81 | 3.44 | 3.50 |
| | SD | .68 | .63 | .56 | .62 | .72 | .86 |
| | skewness | –.496 | –.522 | .010 | –.300 | –.285 | –.302 |
| t-test results | t | –.592 | –.723 | –.729 | –1.431 | –.828 | .247 |
| | *p* | .554 | .470 | .466 | .153 | .408 | .805 |

## Relations between Learning Styles and Strategy Use

Table 3 presents the relations between learning styles and strategy use, which shows that the use of all six types of strategies was significantly positively related to different types of learning styles except for individual learning for the whole sample ($r = .285 \sim .597$, $p < .01$), male ($r = .343 \sim .644$, $p < .01$) and female ($r = .241 \sim .557$, $p < .01$) students. This means that a respondent, whether male or female, who preferred auditory, kinesthetic, visual, tactile or group styles tended to use more frequently memory, cognitive, compensation, metacognitive, affective or social strategies, or vice versa. Meanwhile, individual style was only significantly positively related to compensation strategy use for the whole sample ($r = .129$, $p < .01$) and male ($r = .240$, $p < .01$) students, though the coefficient was not strong. This means that for the whole sample and male students, a respondent who preferred individual learning tended to more frequently utilize compensation strategies, or vice versa.

**Table 3**

*Correlations between Learning Styles and Strategies*

| | Auditory | Kinesthetic | Visual | Tactile | Individual | Group |
|---|---|---|---|---|---|---|
| MemoryS | .506**/.525**/.489** | .584**/.574**/.593** | .464**/.542**/.385** | .526**/.570**/.480** | .015/.072/-.049 | .489**/.502**/.487** |
| CognitiveS | .559**/.606**/.513** | .575**/.644**/.500** | .515**/.555**/.472** | .500**/.558**/.434** | .078/.109/.038 | .445**/.459**/.444** |
| CompensationS | .395**/.518**/.281** | .401**/.428**/.374** | .400**/.511**/.289** | .397**/.523**/.271** | .129**/.240**/.012 | .285**/.343**/.241** |
| MetacognitiveS | .550**/.573**/.533** | .597**/.635**/.557** | .476**/.529**/.421** | .555**/.601**/.502** | .083/.128/.017 | .451**/.464**/.462** |
| AffectiveS | .525**/.520**/.532** | .486**/.554**/.418** | .480**/.512**/.453** | .435**/.524**/.344** | .038/.081/-.012 | .441**/.480**/.417** |
| SocialS | .548**/.551**/.547** | .551**/.605**/.506** | .480**/.513**/.453** | .462**/.493**/.436** | .023/.048/.002 | .474**/.473**/.478** |

*Note.* The three numbers in each column refer to the coefficients for the whole sample, male and female students respectively; MemoryS = memory strategies; CognitiveS = cognitive strategies; CompensationS = compensation strategies; MetacognitiveS = metacognitive strategies; AffectiveS = affective strategies; SocialS = social strategies ** = $p < .01$; * = $p < .05$.

## Predicting Effects of Learning Styles and Strategy Use on English Achievements

To examine the predictive effects of learning style preferences and strategy use on students' English achievements, structural equation modelling (SEM) and multiple stepwise regression analyses were conducted for the whole sample, for both male and female students respectively. In all situations, self-rated English proficiency was used as the dependent variable, learning style and strategy use scores were used as latent independent variables in SEM, while specific learning style and strategy use scores were used as independent variables in regression analyses. The results are displayed in Figure 1 and Table 4 respectively.

SEM analyses showed that the indicator values of the whole sample's model were: $\chi^2 = 343.063$, $p < .01$, RMSEA = .091, CFI = .898, SRMR = 0.058, meaning that the model was good (Hu & Bentler, 1999). As shown in Figure 1(a), learning strategy use was significantly positively related to learning style ($r = .833$, $p < .001$), and strategy use ($\beta = .509$, $p < .001$) significantly positively predicted the whole sample's self-rated English proficiency while learning style ($\beta = –.011$, $p > .05$) did not.

Likewise, the indicator values of male students' models were: $\chi^2 = 205.692$, $p < .01$, RMSEA = .089, CFI = .908, SRMR = .060, meaning that the model was good (Hu & Bentler, 1999). As shown in Figure 1(b), learning strategy use was significantly positively related to learning style ($r = .829$, $p < .001$), and strategy use ($\beta = .448$, $p < .001$) significantly positively predicted male students' self-rated English proficiency, while learning style ($\beta = .113$, $p > .05$) did not.

For female students, the indicator values of the model were: $\chi^2 = 203.638$, $p < .01$, RMSEA = .085, CFI = .885, SRMR = 0.063, meaning that the model was good (Hu & Bentler, 1999). As shown in Figure 1(c), learning strategy use was significantly positively related to learning style ($r = .854$, $p < .001$), and strategy use ($\beta = .669$, $p < .001$) significantly positively predicted female students' self-rated English proficiency while learning style ($\beta = –.243$, $p > .05$) did not.

**Figure 1(a)**

*The Whole Sample*

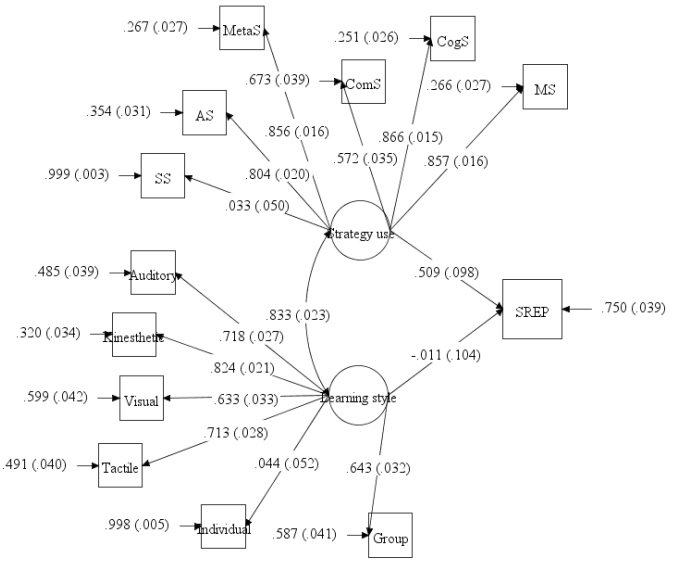



**Figure 1(b)**

*Male Students*

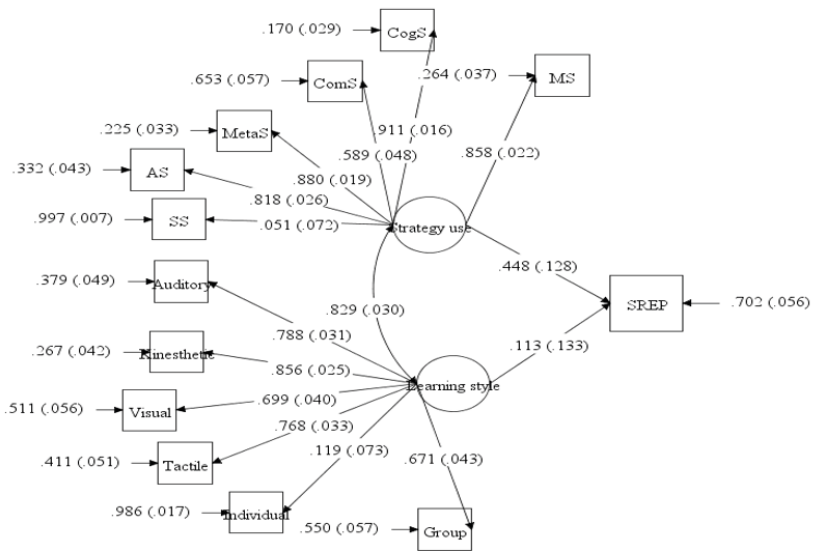


**Figure 1(c)**

*Female Students*

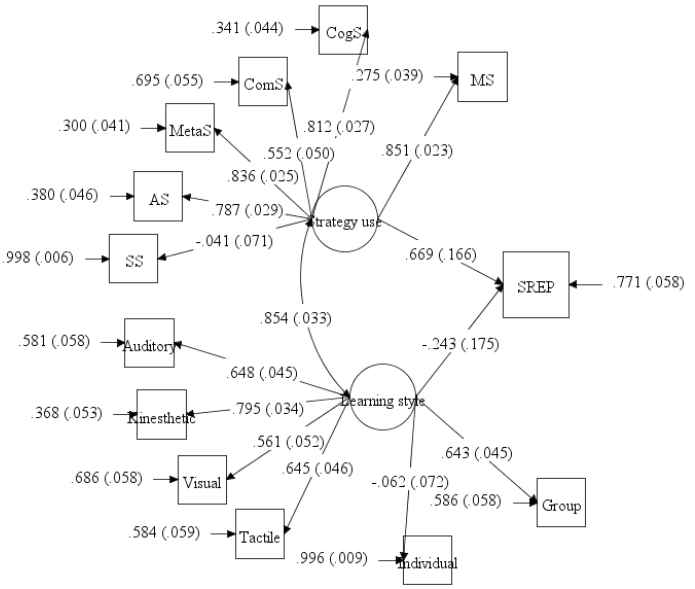



Figure 1: SEM results of learning styles and strategy use on SREP

*Note.* SREP = self-rated English proficiency; MS = memory strategies; Congs = cognitive strategies; ComS = compensation strategies; MetaS = metacognitive strategies; AS = affective strategies; SS = social strategies.

As shown in Table 4, regression analyses yielded one model with the change in $R^2$ being .255 for the whole sample (CognitiveS) and .292 for male students (CognitiveS), respectively. Namely, CognitiveS significantly positively predicted the whole sample's ($\beta$ = .505, $t$ = 12.244, $p$ = .000, $f^2$ = .34) and male students' ($\beta$ = .543, $t$ = 9.309, $p$ = .000, $f^2$ = .41) self-rated overall English proficiency, both with a large effect size. Meanwhile, regression analyses yielded three models for female students, with the change in $R^2$ being .208 for model 1 (CognitiveS), .014 for model 2 (CognitiveS and Tactile), and .016 for model 3 (CognitiveS, tactile and MemoryS). And model 3 seemed to be the best. Thus, CognitiveS ($\beta$ = .392, $t$ = 4.55, $p$ = .000, $f^2$ = .26), Tactile ($\beta$ = –.171, $t$ = –2.564, $p$ = .011, $f^2$ = .06), and MemoryS ($\beta$ = .193, $t$ = 2.183, $p$ = .03, $f^2$ = .069) were good predictors for female students' self-rated overall English proficiency, with CognitiveS and MemoryS being positive predictors, and Tactile being a negative predictor, all with a small to medium effect size.

**Table 4**

*Regression Coefficients and Significance of Predictors for English Achievements*

| | β | *t* | *p* | VIF | Cohen's $f^2$ |
|---|---|---|---|---|---|
| | The whole sample's self-rated English proficiency | | | | |
| CognitiveS | .505 | 12.244** | .000 | 1.000 | .34 |
| | Male students' self-rated English proficiency | | | | |
| CognitiveS | .543 | 9.309** | .000 | 1.000 | .41 |
| | Female students' self-rated English proficiency | | | | |
| CognitiveS | .392 | 4.55** | .000 | 2.209 | .26 |
| Tactile | –.171 | –2.564* | .011 | 1.325 | .06 |
| MemoryS | .193 | 2.183* | .03 | 2.331 | .069 |

*Note*. ** = $p \leq .01$; * = $p \leq .05$.
Effect size of Cohen's $f^2$: small = $f^2 \leq .02$; medium = $f^2 = .15$; large = $f^2 \geq .35$ (Cohen, 1988).

## Discussion and Conclusion

The present study examined perceptual learning styles, strategy use and their predictive effects on Chinese undergraduate EFL Learners' achievement in English. Statistical analyses show that both the PLSPQ and the SILL were fairly reliable and generally significantly correlated with each other.

### Preferences for Learning Styles

The present study revealed that the whole sample, as well as male and female students, displayed no major preference for any style, but used all styles as their minors on the higher end, indicating that they were multi-style learners, similar to the finding in Gao (2016) but different from that in Melton (1990) in similar contexts. This might be attributed to the different characteristics possessed by the participants of the individual studies, though they were generally Chinese learners of English. Due to the complexity of learner populations, learning styles warrant continuous research in various ESL/EFL contexts.

Meanwhile, the participants in the present research preferred the kinesthetic style the most, followed by tactile, auditory, visual, group and individual styles respectively, partially consistent with the findings in the current litera-

ture (e.g., Akbarian et al., 2019; Liu, 2023; Mozayan et al., 2013; Naserieh & Mohammad, 2013; Peacock, 2001; Reid, 1987). This could be largely attributed to the fact that more than 65.1% of participants were engineering and science students who tended to engage in hands-on projects by themselves or with others in learning. By contrast, liberal arts students might tend to read and listen more frequently in their learning, which often does not involve others as much. Their preferences for auditory and visual styles might result from the fact that, in traditional Chinese educational classrooms, teachers often dominate classes through lectures and presentations in various visual forms such as pictures and powerpoint slides. This was particularly the case during COVID-19, when teaching and learning shifted online and interaction was restricted by online meeting platforms and the Internet. As little was known about the participants' backgrounds except for their ages, genders and disciplines, it was hard to pinpoint what had caused their preferences for different styles and whether their style preferences were consistent across time. This merits attention in future studies on learning styles. Moreover, as learners are categorized differently by different learning style theories, it will be interesting to identify the features of their style preferences when gauged by different instruments. These findings will help better understand the learning style myth.

## Language Learning Strategy Use

This study shows that the whole sample as well as male and female students demonstrated a medium to high level of use of the six types of learning strategies, as defined in Oxford (1990). This finding indicates that the participants in the present research tended to have a greater use of the strategies than their peers who normally have a medium level of strategy use, as indicated by numerous current studies (e.g., Magogwe & Oliver, 2007; Tang & Tian, 2015). This was probably caused by the fact that students had to resort to various strategies when learning became more self-dependent due to COVID-19, indicating that they were able to plan, self-regulate and self-assess their own learning and become autonomous learners during the critical time of the pandemic (Lee & Heinz, 2016). This supports the claim in Gao and Hu (2020) that use of language learning strategies is influenced by different resources: community (e.g., a group of people working on the same/similar tasks for similar goals), rules (e.g., time and task requirements), and division of labor (e.g., roles and power relationships).

The study also reveals that the participants used metacognitive strategies the most frequently, followed by cognitive, social, compensation, memory and affective strategies, respectively. This finding was partially consistent with the findings reported in the current literature (e.g., Lee & Heinz, 2016;

Pongsukvajchakul, 2021; Rahman, 2020; Salahshour et al., 2013; Stracke, 2016; Tang & Tian, 2015). This might have resulted from the fact that all the participants were college students, most being engineering and science students, who were generally required to plan and monitor their hands-on projects, especially during COVID-19. When difficulties occurred, they resorted to social, compensation and memory strategies. Affective strategies were the least frequently used, probably because much attention had been given to students' affect and well-being by their universities and parents during the pandemic.

Coupled with the applications of technology in education, students now have a wide repertoire of various forms of strategies (Oxford, 2017). To tackle a task, students may employ a cluster of strategies simultaneously in different orders (Gao & Hu, 2020; Thomas et al., 2022). In the future, more case studies can be conducted to better understand how strategies are used in relation to specific tasks.

## Predicting Effects of Learning Styles and Strategy Use on English Achievements

SEM analyses show that strategy use significantly positively predicted the whole sample's, as well as male and female students' English achievements while learning style did not. Regression analyses further show that cognitive strategy use (CognitiveS) significantly positively predicted the whole sample's as well as male and female students' English achievements. In addition, memory strategy use (MemoryS) significantly positively predicted female students' English achievements, while the tactile style (Tactile) negatively predicted them. These results further support the finding that strategy use often facilitates SL/FL learning (e.g., Lin et al., 2021; Salahshour et al., 2013), while learning styles might have complex relations with ESL/EFL learning outcomes (e.g., Ha, 2019; Malsawmkimi & Fanai, 2019). This might be because learning styles, though variable as individual characteristics like gender, education and culture change, are generally more stable than strategy use. Students may be able to choose more suitable strategies to tackle different learning tasks accordingly, but they may not be able to change their learning styles so flexibly. Yet, because of the paucity of research on learning styles and ESL/EFL learning, more similar research is called for.

### Gender Differences in Learning Styles, Strategy Use and Their Predictive Effects on English Achievements

Table 1 shows that male students preferred individual learning the least, and that female students preferred group learning the least, despite their preferences for other styles remaining similar. This might be because most female students were liberal arts students who tended to study alone. Male students scored higher on auditory, visual and group styles, but lower on the other three than their female peers, partially consistent with the finding in Melton (1990) and Hyland (1993), but different from those in Mozayan et al. (2013) and Akbarian et al. (2019). In addition, as found in Zokaee et al. (2012), generally, no significant difference occurred between genders in all learning styles, except for group learning. It is worth noting that male students scored higher on affective but lower on all the other five types of strategies than their female peers, though no significant difference was observed in any type of strategies, largely consistent with the finding in McMullen (2009) and Salahshour et al. (2013). And a similar pattern of significant positive correlations between learning styles and strategies existed for male and female students, except that individual learning was significantly positively correlated with compensation strategies for only male students. Nevertheless, since few studies have explored gender differences in learning styles and strategy use, these findings need to be further confirmed in more research in various contexts.

Interestingly, this study shows that of all learning styles and strategies, only cognitive strategy use (CognitiveS) significantly positively predicted male students' English achievements, while for female students, cognitive strategy use, memory strategy use (MemoryS), and the tactile style all significantly predicted their English achievements. Considering that male and female students had similar learning preferences and usage of different strategies, this finding proves especially intriguing and deserves further research.

## Implications and Limitations

The findings of this study further attest to the importance of learning styles and strategies in ESL/EFL learning. Hence, it is important for instructors to know their students' learning styles so that they can appropriately match teaching with students' learning styles or scaffold teaching and tasks to stimulate different learning styles (Liu, 2023). Likewise, it is useful for instructors to encourage and train students to use various strategies to handle different tasks. It may also be useful to provide students with adequate resources

(e.g., various use of technology, visual and audio aids, etc.) and instruct them how to use those resources (Thomas et al., 2022) so that they use different forms of strategies more readily and more effectively. As reviewed above, multi-style preference and implementation of diverse appropriate strategies can help students make the most of the learning circumstance.

The present large-scale quantitative study is one of few that uses both SEM and regression analyses to explore predictive effects of learning styles and strategy use on students' English achievements, thus enriching the current literature. Even so, certain limitations existed. First, the findings entirely relied on quantitative data, which could be more generalizable if complemented by qualitative data. Future research can use mixed methods to provide a more comprehensive understanding of learning styles and strategy use in relation to ESL/EFL learning outcomes. Second, due to various constraints, students' English achievements were measured by their self-ratings, which might not truly reflect students' competence in English. A standardized proficiency test could be more reliable and is thus recommended for future research. Moreover, similar research is needed to confirm the findings of the present study, especially those involving SEM and regression analyses. Furthermore, as both learning styles and strategy use vary with age, education, achievement, gender, culture, and other individual characteristics, more internal factors should be integrated into the investigation of the issues in future research.

## References

Akbarian, I., Afzali-Shahri, M., Ghasemi-Rezveh, J., & Salimi, M. (2019). The relationship between perceptual learning style preferences and depth of vocabulary knowledge. *Journal of Language Horizons, 3*(2), 79–103. https://doi.org/10.22051/lghor.2020.28705.1207

Alkahtani, S. S. (2016). *Language learning strategies among Saudi EFL college students and their relationship to students' perceptual learning style, gender, academic major and proficiency level*. [Doctoral dissertation, University of Tennessee]. University of Tennessee TRACE. https://trace.tennessee.edu/utk_graddiss/4122

Al Khatib, S. A., & Ghosheh, S. K. (2013). Perceptual learning style preferences in relation to gender, academic achievement and field of study among a sample of UAE college Students. *Scholars Journal of Arts, Humanities and Social Sciences, 1*(2), 69–80.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Dunn, R., Dunn, K., & Price, G. E. (1975). *The learning style inventory*. Price Systems.

Gao, P. (2016). An analysis of the impact of language learning styles on learning strategy preferences. *Sino-US English Teaching, 13*(7), 542–551. https://doi.org/10.17265/1539-8072/2016.07.006

Gao, X., & Hu, J. (2020). From language learning strategy research to a sociocultural understanding of self-regulated learning. In M. J. Raya & F. Vieira (Eds.), *Autonomy in language education: Theory, research and practice* (pp. 31–45). Routledge.

Gregersen, T., & MacIntyre, P. D. (2014). *Capitalizing on language learners' individuality*. Multilingual Matters.

Griffiths, C. (2020). Language learning strategies: Is the baby still in the bathwater? *Applied Linguistics, 41*(4), 607–611. https://doi.org/10.1093/applin/amy024

Ha, N. T. (2019). Research learning style preferences and English language proficiency of first year economics students at University of Economics-Technology for Industries, Vietnam. *International Journal of English Language Teaching, 7*(3), 67–82.

Hawk, T. F., & Shah, A. J. (2007). Using learning style instruments to enhance student learning. *Decision Sciences Journal of Innovative Education, 5*(1), 1–19. https://doi.org/10.1111/j.1540-4609.2007.00125.x.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*, 1–55. https://doi.org/10.1080/10705519909540118

Huang, F., Hoi, C. K. W., & Teo, T. (2018). The influence of learning style on English learning achievement among undergraduates in mainland China. *Journal of Psycholinguistic Research, 47*(5), 1069–1084. https://doi.org/10.1007/s10936-018-9578-3

Hyland, K. (1993). Culture and learning: A study of the learning style preferences of Japanese students. *RELC Journal,* 24(2), 69–87. https://doi.org/10.1177/003368829302400204

Keefe, J. W. (1979). Learning style: An overview. NASSP's *Student learning styles: Diagnosing and prescribing programs* (pp. 1–17). [ERIC Document No. ED182859]. National Association of Secondary School Principles.

Koglin, G., Arald, M. O., & Felicetti, V. L. (2021). College students and perceptions of their learning styles. *International Journal of Higher Education, 7*, 1–15. https://doi.org/10.20396/riesup.v7i0.8660214

Kolb, D. A. (1984). Experiential learning: *Experience as the source of learning and development*. Prentice-Hall.

Lee, J., & Heinz, M. (2016). English language learning strategies reported by advanced language learners. *Journal of International Education Research, 12*(2), 67–76. https://doi.org/10.19030/jier.v12i2.9629

Lethaby, C., & Russell, M. (2020). A critical examination of perceptual learning styles in English language teaching. *International Review of Applied Linguistics in Language Teaching*, *58*(2), 221–237. https://doi.org/10.1515/iral-2017-0067

Lin, L., Lam, W.-I., & Tse, S. K. (2021). Motivational strategies, language learning strategies, and literal and inferential comprehension in second language Chinese reading: A structural equation modeling study. *Frontiers in Psychology*. 12:707538. https://doi.org/10.3389/fpsyg.2021.707538

Liu, M. (2023). English classroom anxiety, learning style and English achievement in Chinese university EFL students. *Sustainability,* 15, 13697. https://doi.org/10.3390/su151813697

Liu, M., & Chen, Z. (2024). Predictive and mediating effects of learning strategies and styles on Chinese undergraduate students' English achievement. *The Asia-Pacific Education Researcher,* 0123456789. https://doi.org/10.1007/s40299-023-00775-5

Lui, C. J. (2017). The perceptual learning style preferences of Hispanic students in higher education. *All Theses and Dissertations*. 6712. https://scholarsarchive.byu.edu/etd/6712

Magogwe, J. M., & Oliver, R. (2007). The relationship between language learning strategies, proficiency, age and self-efficacy beliefs: A study of language learners in Botswana. *System, 35*(3), 338–352. https://doi.org/10.1016/j.system.2007.01.003

Malsawmkimi, M., & Fanai, L. (2019). A Study of perceptual learning styles with academic achievement in private higher secondary schools in Aizawl city. *Educational Quest: An International Journal of Education and Applied Social Science, 10*(1), 21–25. https://doi.org/10.30954/2230-7311.1.2019.4

McMullen, M. G. (2009). Using language learning strategies to improve the writing skills of Saudi EFL students: Will it really work? *System, 37*(3), 418–433. https://www.learntechlib.org/p/105412

Melton, C. D. (1990). Bridging the cultural gap: A study of Chinese students' learning style preferences. *RELC Journal, 21*(1), 29–54. https://doi.org/10.1177/003368829002100103

Mozayan, M. R., Ebrahimpourtaher, A., Hoominian, Z., Khosravi, A., & Shamsi, F. (2013). Perceptual learning styles of medical sciences students. *Global Journal of Human Social Science, 13*(10), 57–62.

Naserieh, F., & Mohammad, R. A. S. (2013). Perceptual learning style preferences among Iranian graduate students. *System, 41*(1), 122–133. https://doi.org/10.1016/j.system.2013.01.018

Noormohamadi, R. (2009). On the relationship between language learning strategies and foreign language anxiety. *Journal of Pan-Pacific Association of Applied Linguistics, 13*(1), 39–52.

Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Heinle & Heinle Publishers.

Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context*. Routledge. https://doi.org/10.4324/9781315719146

Oxford, R. L., & Cohen, A. D. (1992). Language learning strategies: Crucial issues of concept and classification. *Applied Language Learning, 3*, 1–35. https://doi.org/10.4236/ce.2022.133049

Peacock, M. (2001). Match or mismatch? Learning styles and teaching styles in EFL. *International Journal of Applied Linguistics, 11*(1), 1–20. https://doi.org/10.1111/1473-4192.00001

Pelegrín, J. D. (2020). *Perceptual Learning Style Preference Questionnaire*: A review of a language learning styles survey. *Rhesis*. *International Journal of Linguistics, Philology, and Literature*, *11*(1), 319–335.

Pongsukvajchakul, P. (2021). Language learning strategies used in English writing by Thai undergraduate students. *Shanlax International Journal of Education, 9*(2), 54–59. https://doi.org/10.34293/education.v9i2.3626

Rahman, M. M. U. (2020). EFL learners' language learning strategies: A case study at Qassim University. *Advances in Language and Literary Studies, 11*(5), 6–11. http://doi.org/10.7575/aiac.alls.v.11n.5-p.6

Reid, J. M. (1987). The learning style preferences of ESL students. *TESOL Quarterly*, *21*(1), 87–111. https://doi.org/10.2307/3586356

Reid, J. M. (1995). *Learning styles in the ESL/EFL classroom*. Heinle & Heinle.

Rubin, J. (1975). *What the "good language learner" can teach us. TESOL Quarterly, 9*(1), 41–51. https://doi.org/10.2307/3586011

Salahshour, F., Sharifi, M., & Salahshour, N. (2013). The relationship between language learning strategy use, language proficiency level and learner gender. *Procedia-Social and Behavioral Sciences, 70*, 634–643. https://doi.org/10.1016/j.sbspro.2013.01.103

Stracke, E. (2016). Language learning strategies of Indonesian primary school students: In relation to self-efficacy beliefs. *System, 60*, 1–10. https://doi.org/10.1016/j.system.2016.05.001

Sun, X., Norton, O., & Nancekivell, S. E. (2023). Beware the myth: learning styles affect parents', children's, and teachers' thinking about children's academic potential. *NPJ Science of Learning,* 8:46. https://doi.org/10.1038/s41539-023-00190-x

Tang, M., & Tian, J. (2015). Associations between Chinese EFL graduate students' beliefs and language learning strategies. *International Journal of Bilingual Education and Bilingualism, 18*(2), 131–152. https://doi.org/10.1080/13670050.2014.882882

Thomas, N., Rose, H., Cohen, A. D., Gao, X., Sasaki, A., & Hernandez-Gonzalez, T. (2022). The third wind of language learning strategies research. *Language Teaching,* 55, 417–421. https://doi.org/10.1017/S0261444822000015

Vimalakshan, S., & Aziz, A. A. (2021). Investigating language learning strategies used by ESL lower secondary learners from national and vernacular school background. *International Journal of New Technology and Research, 7*(1), 41–47. https://doi.org/10.31871/IJNTR.7.1.13

Whitman, G. M. (2023). Learning styles: Lack of research-based evidence. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 1–7. https://doi.org/10.1080/00098655.2023.2203891

Zokaee, S., Zaferanieh, E., & Naseri, M. (2012). On the impacts of perceptual learning style and gender on Iranian undergraduate EFL learners' choice of vocabulary learning strategies. *English Language Teaching, 5*(9), 138–143. http://dx.doi.org/10.5539/elt.v5n9p138

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/29
https://doi.org/10.31261/TAPSLA.16445

CC BY SA

**Mohammad Hamad Al-khresheh**
Northern Border University, Arar, Saudi Arabia
https://orcid.org/0000-0001-8647-8854

# The Role of Digital Storytelling in Jordanian School EFL Classrooms: A Qualitative Exploration of Teachers' Perceptions

Abstract

The topic of digital storytelling in the field of EFL instruction has attracted increasing attention. However, its significance, effectiveness, and incorporation within educational settings in Jordan have not been thoroughly studied. In order to address this research gap, the present paper explores the teaching effectiveness, challenges, and practical applications of digital storytelling in Jordanian EFL classrooms, specifically through a qualitative analysis of teacher perspectives, incorporating semi-structured interviews with 26 EFL teachers recruited by purposive sampling. The thematic analysis revealed that EFL teachers in Jordan predominantly view digital storytelling as a valuable enhancement to English language teaching, notably for its ability to engage students and facilitate language learning. This positive perspective, however, is tempered by concerns over resource availability and curriculum integration challenges. The study also observed that digital storytelling positively impacts student engagement and language skills, with teachers reporting notable improvements in student involvement and language competencies. However, this impact varied among students, influenced by individual learning styles and access to technological resources. Lastly, regarding integrating digital storytelling into the curriculum, teachers identified several benefits, including the versatility of digital storytelling in achieving educational objectives and its effectiveness in facilitating interactive learning. Yet, they also faced challenges, particularly regarding inadequate technological infrastructure and difficulties aligning digital storytelling with rigid curricular structures. These findings provide crucial insights into the application of digital storytelling in EFL education in Jordan, reflecting its potential to transform teaching practices amidst practical implementation challenges. The paper concludes with limitations and future research suggestions, intending to encourage deeper scholarly and practical involvement with this novel teaching style.

*Keywords*: EFL, digital storytelling, Jordanian EFL teachers, Jordanian EFL classrooms, teaching methods, technology use, teachers' perceptions.

Over the past few decades, there have been significant changes in English as a Foreign Language (EFL) teaching, driven mainly by integrating technology into educational frameworks. This phenomenon holds particular significance within digital storytelling, a novel instructional approach that uses digital technologies to construct and communicate narratives. Using a combination of text, graphics, and audio in digital storytelling has been recognized as an effective tool for promoting language learning and enhancing educational involvement (Abdolmanafi-Rokni, 2014; Robin, 2008; Sadik, 2008).

However, the utilization of digital storytelling in specific educational contexts, such as EFL classes in Jordan, remains underexplored. This is particularly noteworthy in a nation where English instruction is compulsory but often delivered through traditional, teacher-centred methods, resulting in persistently below-average proficiency levels (Aljaberi, 2021; Aljaraideh, 2020; Al-Sharah et al., 2021; Mansi, 2019). The situation underscores the need for educational approaches that prioritize learner engagement, interactivity, and communication. In light of this, it is surprising that while previous scholarly work has extensively investigated the role of technology in EFL settings, there has been a relative lack of focus on the specific utility of digital storytelling (Haga, 2023; Luan et al., 2023). Moreover, empirical research examining the implementation of this pedagogical approach in Jordanian EFL classrooms is notably absent. To fill this research gap, the present study explores the educational effectiveness, challenges, and practical applications of digital storytelling in Jordan, specifically through a qualitative analysis of teacher perspectives. This study's findings can significantly influence curriculum design, pedagogical practices, and educational policies related to technology use in EFL contexts. Aligned with the primary research objective, the study posits the following research questions:

1. How do EFL teachers in Jordan perceive the role of digital storytelling in enhancing English language teaching and learning in their classrooms?
2. Based on teachers' observations, what impact does digital storytelling have on student engagement and language skills in EFL classrooms in Jordan?
3. What specific challenges and benefits do EFL teachers in Jordan identify when integrating digital storytelling into their curriculum?

## Literature Review

### Digital Storytelling: Definition and Components

Digital storytelling represents a modern narrative practice that artfully integrates the traditional craft of storytelling with diverse digital multimedia

elements, including images, audio, and video. This innovative approach enables individuals to articulate and disseminate personal, cultural, or historical narratives in an engaging and interactive format (Meletiadou, 2022). Digital storytelling converges conventional storytelling methods and modern multimedia components, resulting in a wide-ranging narrative structure (Setiyorini, 2020). Emerging as a dual-purpose medium encompassing artistic expression and pedagogical utility, multimedia incorporates text, music, images, and frequently video and animation to enhance the narrative encounter (James et al., 2019). The multimedia elements in question function not just as supplemental features but also as intrinsic components that augment the emotional and cognitive depth of the narrative. For example, the deliberate incorporation of audio can introduce additional dimensions of emotion and context. At the same time, visual components can effectively depict characters and environments, thus enhancing audience engagement to a greater extent (Robin, 2008).

Within educational settings, the application of digital storytelling encompasses more than just fostering student involvement (Yuliani & Hartanto, 2022). Producing digital storytelling transforms into an educational endeavour incorporating several linguistic proficiencies, encompassing written expression, spoken communication, and occasionally textual analysis and investigation (Ramalingam et al., 2022). Digital storytelling embodies contemporary pedagogical approaches, including communicative and task-based language teaching. It offers a comprehensive and holistic learning experience that aligns with modern educational principles (Robin, 2008).

Furthermore, this methodology also addresses the cognitive and emotional dimensions of learning. Creating a digital story necessitates integrating critical thinking, problem-solving, and decision-making abilities, as students are tasked with effectively structuring their thoughts and selecting the most appropriate multimedia components (Salem, 2022). The emotional involvement of students is heightened by the frequently intimate character of storytelling, resulting in increased recall and a more personally significant learning experience (Yang & Wu, 2012).

Practically, digital storytelling emerges as a pivotal pedagogical approach, harnessing the power of multimedia technology to revolutionize the educational narrative, moving beyond traditional storytelling methods to a more integrated multimedia experience. This innovative strategy enables learners to weave together text, images, audio, and video to craft engaging narratives, merging the art of storytelling with digital prowess (Maureen et al., 2020). Projects like the Personal Narrative Project allow students to explore and articulate personal experiences through detailed scriptwriting and the selection of multimedia elements, enhancing their narrative skills and language proficiency (Dreher, 2012). Similarly, the Cultural Exchange Project encourages students to engage with and present different cultural perspectives, promoting research skills, critical

thinking, and empathy (Oakley et al., 2018). These activities foster creative and digital literacy and enrich students' understanding of language and culture, illustrating digital storytelling's role as a dynamic educational tool. By providing immersive, context-driven language and cultural learning opportunities, digital storytelling is an essential asset in modern education, equipping students with the skills to navigate a digitally interconnected world.

However, successfully incorporating digital storytelling in educational environments requires specific technological requirements. The successful deployment of this novel teaching tool is influenced by various aspects, such as the availability of hardware and software and the technical literacy levels of teachers and students (Sadik, 2008; Meletiadou, 2022).

## Technological Tools in Digital Storytelling

Integrating diverse hardware and software technologies is critical in digital storytelling, particularly within EFL instruction. Essential hardware components for EFL digital storytelling include an array of devices such as computers, tablets, smartphones, and digital cameras, each playing a distinct role in the creation, editing, and presentation of digital narratives tailored to language learning (Robin, 2008; Christiansen & Koelzer, 2016). Computers and tablets, recognized for their robust processing capabilities and full multimedia support, are predominantly utilized for EFL story creation and interactive engagements (Churchill, 2020). Smartphones, with their widespread accessibility and advanced features, have become indispensable tools for capturing audio-visual content in EFL contexts. Moreover, digital cameras provide superior options for high-quality visual storytelling, enhancing the visual appeal of language learning materials.

On the software spectrum, many applications and platforms are employed to streamline the creation and dissemination of EFL digital stories (Urbieta & Peñalver, 2021). Advanced video editing software like Adobe Premiere Pro and Final Cut Pro enables intricate narrative constructions and productions, suitable for EFL content creation. More accessible software such as iMovie and Windows Movie Maker offer essential editing functionalities, facilitating the creation of EFL learning materials for beginners (Castillo-Rodriguez et al., 2022). Specialized digital storytelling software like Storybird and Book Creator, which provide tailored templates and resources, are especially beneficial for constructing educational stories in EFL settings (Kazazoglu & Bilir, 2021).

Incorporating Learning Management Systems like Moodle, Blackboard, and Canvas has been pivotal in embedding digital storytelling within EFL curricula, simplifying content distribution and accessibility for language learners (Masters, 2020; Yu & Lu, 2021). These systems furnish educators with assessment tools

and interactive feedback mechanisms, enriching the EFL learning experience within a storytelling context.

The emergence of augmented reality and virtual reality technologies marks a new frontier in immersive storytelling experiences, particularly beneficial in EFL education (Karacan & Akoğlu, 2021). Augmented reality applications blend digital content with the physical world, creating interactive and enhanced narratives conducive to language learning (Yastibaş et al., 2022). Virtual reality further immerses users in entirely virtual environments, offering unparalleled opportunities for engaging in historical recreations or scientific simulations, thus enriching the EFL learning experience (Arvanitis & Krystalli, 2021; Merchant et al., 2014).

The technological landscape in digital storytelling for EFL is both diverse and progressive, encompassing a broad spectrum of hardware and software engineered to create, edit, and disseminate stories dynamically and interactively. This technological progression has not only refined the storytelling experience but has also broadened the scope of educational applications, especially within the realm of EFL classrooms.

## Efficacy of Digital Storytelling in Language Learning

Digital storytelling has received considerable interest due to its potential to enhance multiple aspects of language learning. Numerous studies support the beneficial effects of digital storytelling on language learning: expanding vocabulary, enhancing listening comprehension, and fostering cultural awareness (Sadik, 2008; Tran, 2021; Yang & Wu, 2012). Creating a digital story frequently requires a profound involvement with language, as students are motivated to investigate novel terminology and idiomatic idioms to communicate their narratives proficiently. This involvement expands their lexical repertoire and offers a significant context for utilizing newly acquired words, consequently augmenting retention and comprehension (Huang, 2022; Lim et al., 2022).

Furthermore, using multimedia elements in digital storytelling provides a distinct opportunity for enhancing one's listening skills (Zhong & Shen, 2002). Students are presented with diverse auditory components, including narrative, background music, and sound effects, all of which collectively enhance the auditory encounter. The intricacy of this phenomenon can enhance students' listening comprehension skills, allowing them to interpret significance from a diverse range of auditory stimuli (Yang & Wu, 2012).

In addition to its language aspects, digital storytelling also functions as a means to facilitate cultural discovery (Shen & Peng, 2022). Including cultural themes or settings in narratives can serve as a means for students to acquire knowledge and understanding of diverse cultural norms, beliefs, and practices.

Including cultural enrichment in language learning increases the experience by emphasizing the socio-cultural dimensions of language usage (Sadik, 2008).

In addition to the advantages above related to language learning, digital storytelling has garnered praise for its capacity to cultivate a diverse set of non-technical skills that are increasingly recognized as essential in the modern educational environment of the 21st century (Abdolmanafi-Rokni, 2014; Orak & Al-khresheh, 2021). Creating a digital story necessitates a significant level of critical thinking, as students must carefully strategize their narratives' structure, select suitable multimedia components, and seamlessly integrate them into a cohesive storyline. According to Aljaberi (2021), engaging in planning and executing tasks fosters creativity and necessitates teamwork, mainly when undertaken as a collective endeavour. Consequently, this practice cultivates crucial skills that transcend the boundaries of the language classroom.

## The Role of Digital Storytelling in EFL Contexts

The genesis of digital storytelling in EFL can be traced back to around 2008, marking a period of significant experimentation and early adoption in language education (Wu & Chen, 2020). During this nascent stage, teachers began exploring the potential of digital narratives to enhance language learning, recognizing the unique blend of visual, auditory, and textual elements that these tools offered (Pandhiani, 2022). Pioneering efforts during this period laid the groundwork for integrating digital storytelling in EFL curricula, demonstrating its effectiveness in engaging learners and facilitating language learning. Notable early studies and implementations, such as those by McLellan (2007) and Robin (2016), highlighted how digital storytelling could be harnessed to develop essential language skills, particularly vocabulary enrichment, grammatical understanding, and oral proficiency. These initial explorations played a crucial role in shaping the methodologies and pedagogical strategies that are now widely recognized in the field of EFL.

Digital storytelling has established a well-regarded track record of improving linguistic abilities following these early developments. The inherent focus of digital storytelling on language makes it an ideal tool for teaching EFL (Abdolmanafi-Rokni, 2014; Aljaberi, 2021; Indriani & Suteja, 2023; Tatli et al., 2022). Yuliani and Hartanto (2022) underscored its practicality within online English courses, specifically focusing on its relevance for pre-service teachers. Studies conducted by Alemi et al. (2022), Dewi and Sari (2022), and James et al. (2019) have provided evidence that the utilization of digital storytelling can substantially enhance English language learners' speaking and writing skills. Octaviani and Hesmatantya (2018) expanded the scope of the study

to encompass the assessment of story-retelling abilities in kindergarten pupils, emphasizing the tool's adaptability across several educational tiers.

In EFL teaching, storytelling emerges as an essential pedagogical instrument. Its application transcends conventional didactic methodologies, fostering the holistic enhancement of linguistic abilities (Yuliani & Hartanto, 2022). Fundamentally, storytelling assists in the augmentation of vocabulary, the comprehension of grammatical intricacies, and the development of narrative competencies. These elements are pivotal in thoroughly mastering spoken and written English (Abdolmanafi-Rokni, 2014). Storytelling narratives' engaging and contextual nature provides a conducive environment for language learning, propelling learners to organically assimilate and employ new vocabulary and structural patterns (Indriani & Suteja, 2023).

The integration of digital storytelling marks a significant enhancement in language learning, representing a step in applying contemporary teaching strategies. By integrating multimedia facets like audio, visual imagery, and interactive elements, digital storytelling significantly enhances the conventional narrative paradigm (James et al., 2019). This contemporary adaptation effectively captivates diverse learning preferences and broadens language education's scope and inclusivity. Digital storytelling in EFL contexts renders the learning process more immersive and expands the horizons of cultural and contextual comprehension, thus providing learners with a more expansive viewpoint (Robin, 2008).

Furthermore, extending beyond linguistic skill development, digital storytelling has also demonstrated significant psychological and socio-cultural benefits. Zarifsanaiey et al. (2022) found that digital storytelling can enhance social and emotional intelligence, suggesting its potential for holistic educational development. Complementing these findings, studies by Fan (2022) and Pandhiani (2022) revealed that this approach fosters critical engagement among learners and strengthens their sense of identity. Belda-Medina (2021) highlighted its role in positively influencing teacher candidates' attitudes towards diversity and inclusivity, while Rusda et al. (2023) used digital storytelling to reinforce environmental literacy. Such research accentuates that digital storytelling is a tool for academic growth and a catalyst for social consciousness and proactive engagement in societal issues.

Studies by Kim et al. (2023) and Luan et al. (2023) reveal that incorporating digital storytelling markedly elevates the quality of education. It notably bolsters student engagement, stimulates creative thought processes, and aids in the comprehensive grasp of intricate ideas. The interactive and multimedia-enriched digital storytelling approach is critical to creating an engaging and immersive educational environment. According to Kim et al. (2023), the tool facilitates students' active participation in self-expression. Luan et al. (2023)

also demonstrated its potential to enhance students' inclination to communicate in a flipped classroom environment. Zuhriyah (2023) and Fitri et al. (2022) have made valuable contributions by providing empirical evidence that highlights the favourable influence of the tool on students' views and motivations. Despite the positive results observed, there is a notable emphasis on the necessity for enhanced pedagogical and technological readiness. Belda-Medina (2021) and Fitri et al. (2022) have emphasized the need for forthcoming research and practical applications to consider this particular element, thereby paving the way for further advancements in this rapidly developing sector. This call for exploration and advancement in digital storytelling aligns closely with the context of English education in Jordan, where innovative approaches are essential to address the existing challenges in language learning.

In the context of Jordanian EFL classrooms, the integration of digital storytelling is garnering special attention, mainly due to its alignment with the rising trend of digital literacy among students and its potential to modernize traditional language teaching methods (Aljaraideh, 2020). This approach, known for its interactive essence, not only stimulates active engagement and creativity in learners but also effectively enhances language competencies within a dynamic and culturally diverse educational setting. The positive impact of storytelling, especially in its digital form, on language learning has been well-documented (Meletiadou, 2022; Salem, 2022; Yang & Wu, 2012), highlighting the benefits of narrative-based techniques and the amplifying effect of technology in language learning.

Despite compulsory English education in Jordan, proficiency levels often fall short of regional and global benchmarks (Aljaraideh, 2020; Mansi, 2019). The prevalent use of conventional, teacher-oriented teaching methods, which focus heavily on memorization and grammatical rules, has been critiqued for not adequately fostering practical communication skills (Canals & Al-Rawashdeh, 2018; Al-Sharah et al., 2021). In this light, digital storytelling emerges as a promising pedagogical alternative. It emphasizes thorough mastery of linguistic skills and fosters critical reasoning and cultural comprehension, addressing the gaps in traditional methods. For Jordanian students, who often face limited opportunities for authentic language use and cultural immersion, digital storytelling could be particularly advantageous. However, successfully implementing this innovative approach requires overcoming logistical challenges such as curriculum modification, teacher training, and developing a robust digital infrastructure.

Despite the considerable body of research on the function of technology in educational contexts, including EFL classrooms, a notable deficiency exists in understanding the precise usefulness and influence of digital storytelling. The disparity is especially notable in non-Western contexts such as Jordan, where conventional instructional approaches prevail, and English language

ability is not optimal. Previous research predominantly concentrates on generic technological tools or the possible advantages of digital storytelling, such as learning vocabulary and developing critical thinking skills. However, there is a notable oversight in these studies about the impact of digital storytelling on communicative abilities and cultural comprehension. The primary objective of this study is to fill the existing knowledge gaps by investigating the utilization and effects of digital storytelling in EFL classrooms in Jordan.

# Research Method

## Research Design

The research design utilized in this study is based on a qualitative paradigm, specifically adopting a phenomenological approach. The study's primary aim is to understand the role of digital storytelling in EFL classrooms in Jordan. The rationale for selecting this design is supported by the phenomenological emphasis on investigating subjective experiences, thereby providing a detailed understanding of EFL teachers' beliefs, perspectives, and difficulties (Neuman, 2014; Yin, 2018). The utilization of the phenomenological approach facilitates a deeper examination of the experiences of educators, surpassing superficial descriptions and delving into the intricacies of incorporating digital storytelling into teaching methods. The framework offered facilitates the systematic examination of the multifaceted nature of these experiences. It demonstrates a particular proficiency in uncovering fundamental themes and recurring patterns that shed light on the motivations and methods behind digital storytelling. Also, it explores the perceived advantages and obstacles encountered throughout its implementation. Including this methodology enhances the study's methodological rigour, assuring that the findings make a valuable contribution to the current body of research.

## Participants

The research utilizes a purposive sampling method to select 26 EFL teachers from diverse educational settings in Jordan, deliberately chosen for their direct relevance to the study's focus on digital storytelling in EFL instruction. This technique ensures the inclusion of participants likely to provide rich, varied insights due to their unique experiences and perspectives in implementing digital storytelling techniques in different educational contexts (Emmel, 2013).

In Jordan, as mentioned earlier, English language education is a compulsory component of public and private school curricula, beginning from the first year of primary education. On average, students engage in five hours of English instruction per week, a commitment reflecting the importance placed on English proficiency in the national educational agenda. The curriculum, standardized and overseen by the Jordanian Ministry of Education, is designed to develop linguistic skills and foster a deeper cultural understanding, integrating global perspectives with language learning (Alkhaldi & Kayapinar, 2022). However, it is essential to note that the practical application of this curriculum is subject to variation across different schools. Factors such as the availability of resources, the institution's technological infrastructure, and the teaching staff's specific expertise can influence how English is taught, resulting in diverse educational experiences in English language learning across the country (Alzubi, 2023).

The participants in this study represent a cross-section of this system, working under varied conditions. While some teach in well-resourced urban schools with extensive access to technological tools such as interactive whiteboards and digital storytelling software, others come from less advantaged rural or suburban areas, where such resources are limited. This disparity in access to technology is a significant factor in the adoption and effectiveness of digital storytelling techniques in their classrooms.

The selected teachers' ages range from 26 to 54 years, providing a spectrum of perspectives influenced by generational differences in pedagogical approaches and familiarity with digital technologies. Their educational backgrounds are equally diverse, spanning from bachelor's degrees to PhDs in fields related to English language education. This range ensures a variety of theoretical perspectives and practical experiences in EFL teaching. The sample also encompasses a broad range of teaching experiences, from newcomers to experts, contributing to diverse views on the challenges and potential of digital storytelling in EFL instruction.

The primary aim of selecting such a varied group is to capture a deep understanding of the role, challenges, and advantages of digital storytelling in the context of Jordan's EFL teaching environment. Table 1 details the participants' demographics, educational backgrounds, and teaching contexts.

**Table 1**

*The Participants' Characteristics*

| **Variables** | | **Frequency** | **Percentage** [%] |
|---|---|---|---|
| Gender | Male | 20 | 77 |
| | Female | 6 | 23 |
| Qualifications | Bachelor's | 15 | 58 |
| | Master's | 10 | 38 |
| | Doctorate | 1 | 4 |
| Teaching Experience | 1–5 | 5 | 19 |
| | 6–10 | 10 | 38 |
| | 11–15 | 4 | 15 |
| | More than 15 | 7 | 28 |
| **Total** | | 26 | 100 |

This table provides a detailed breakdown of the characteristics of the 26 EFL teachers participating in the study. The gender distribution indicates a higher representation of male participants (77%) than female participants (23%). Regarding academic qualifications, the majority hold a Bachelor's degree (58%), followed by Master's degree holders (38%), and a minority with a doctorate (4%). The teaching experience among the participants varies, with the largest group having 6–10 years of experience (38%), followed by those with more than 15 years (28%). Participants with 1–5 years and 11–15 years of experience constitute 19% and 15% of the sample, respectively.

This participant profile, particularly the gender distribution, shows a deviation from the global trend in the teaching profession, where females often outnumber males. It is crucial to understand that the higher proportion of male participants in this study is partially influenced by the structure of the Jordanian education system, where male and female teachers often work in separate schools. This separation made it more feasible to access male teachers during the research, contributing to the higher representation of males in this sample. Moreover, the purposive sampling technique was employed to ensure the inclusion of participants who could provide informed and diverse insights into integrating digital storytelling in EFL instruction rather than strictly aiming for a gender-balanced representation. As a result, while the gender distribution in this study diverges from broader trends, it uniquely reflects the specific dynamics of the EFL teaching environment in Jordan, thereby providing valuable perspectives on using digital storytelling in this distinct educational context.

## Instrument

The primary data for this study was collected through semi-structured interviews, an approach chosen for its balance of directed focus and participant-driven expression (Yin, 2018). The interview is structured into two main parts. The first part collects demographic information from the participants, including gender, qualifications, and teaching experience. This helps in understanding the diverse backgrounds of the EFL teachers involved. The second part of the interview consists of five open-ended questions designed to elicit detailed responses on specific aspects of digital storytelling in EFL education. These questions explore the teachers' perspectives on the role of digital storytelling in EFL pedagogy, their observations of student engagement during digital storytelling activities, challenges faced in its implementation, perceived benefits, and suggestions for better integration in educational settings. For a detailed view of the interview structure, refer to Appendix 1, which contains the interview format.

## Content Validity and Reliability

Multiple efforts were taken to validate and ensure the reliability of the data analysis methodologies to ensure the scientific rigour of the study. A panel of three subject matter experts specializing in EFL instruction thoroughly evaluated the content validity of the interview questions. The feedback they provided was used to improve the questions and ensure they accurately captured the elements of interest (Neuman, 2014). To ensure internal reliability, a pilot study was done with five EFL teachers who were not among the main study participants. The pilot study results were analysed and cross-checked to determine the effectiveness and clarity of the questions. This preliminary practice gave valuable insights into the reliability of the interview questions and allowed for any necessary tweaks before the main study (Cohen et al., 2017).

## Data Collection

The data collection for this study was executed with thorough attention during the first semester of the academic year 2023–2024. Central to this phase were the semi-structured interviews conducted with 26 carefully selected EFL teachers in Jordan, chosen for their diverse experiences and expertise. These interviews were scheduled with consideration for the participants' convenience to facilitate maximum engagement and responsiveness. The settings chosen for these interviews, primarily within the teachers' educational institutions, were

intended to provide a comfortable and familiar atmosphere conducive to open and honest communication.

Before the interviews, each participant was given a detailed briefing about the study's objectives, the procedures to be followed, and the ethical standards upheld, including strict adherence to confidentiality and the voluntary nature of their participation. Informed consent was duly obtained from all participants, ensuring compliance with ethical research guidelines.

The interviews were audio-recorded with the participant's explicit permission to ensure accurate capture of their responses for thorough analysis. These recordings were then transcribed in detail and systematically coded. The coding process was integral to the analysis, identifying significant themes and insights and ensuring a structured and wide-ranging examination of the data. This methodical approach was crucial in guaranteeing the reliability and depth of the research findings, thus providing a nuanced understanding of the role of digital storytelling in EFL instruction.

## Data Analysis

Thematic analysis was used to interpret the qualitative data collected through semi-structured interviews in this study, a method lauded for effectiveness in qualitative research (Cohen et al., 2017). This approach proved ideal, enabling a deep and structured exploration of participants' experiences and perspectives organized into distinct thematic categories. The intricate process involved the meticulous coding of interview transcripts, a critical step in identifying and analyzing emerging themes. These themes, representing the core elements of participants' responses, were systematically organized to reflect the depth and complexity of the data. The flexibility and comprehensive scope of thematic analysis made it the method of choice, ensuring a nuanced interpretation of the rich, detailed narratives provided by the participants.

## Ethical Consideration

The key objective was thoroughly explained to participants before the main study. They were also informed that their participation was entirely voluntary, that they had the right to confidentiality, and that they could opt out of the study at any time. Informed permission forms were signed and collected, and rigorous steps were implemented to anonymize all identifiable information, ensuring participant confidentiality.

# Findings

This section explores the insights from interviews with 26 Jordanian EFL teachers, focusing on their perspectives on using digital storytelling in educational settings. The data are organized into five major thematic categories: teachers' perceptions of digital storytelling, integration with instructional practices, student engagement levels, encountered challenges, and observed benefits. These themes cooperatively provide an in-depth understanding of the multifaceted role of digital storytelling in the EFL classroom, revealing its advantages and limitations.

## Teachers' Perceptions of Digital Storytelling in EFL Teaching

This theme elicited various opinions from the interviewed teachers. The preponderance of responses was generally favourable. Exclamations such as "excellent," "very high," and "advances the educational reality" encapsulated the enthusiasm felt by the majority of participants regarding the function of digital storytelling in the classroom. They elaborated that it serves multiple purposes, including breaking up "stagnation and monotony" in the classroom and enhancing students' "knowledge" and "abilities and talents." Several participants remarked that it "enables students to comprehend the most challenging subjects and topics" and facilitates "unconventional summaries of academic topics." One participant summed up the consensus: "I believe digital storytelling plays an important role in enhancing students' English language learning." However, not all of the responses were positive. Two participants described digital storytelling as "moderate," "kind of good," or "fairly good." One participant stated that "difficult availability" prevented them from using it frequently. Despite these divergent opinions, the consensus was that digital storytelling has a positive impact on enhancing English language instruction, albeit with some reservations and challenges.

## Integrating Digital Storytelling into EFL Instructional Practices

Teachers provided different insights into their varied strategies and outcomes regarding digital storytelling. While several teachers strongly supported digital storytelling, citing its positive impact on student engagement and learning outcomes, there were differing levels of implementation and perspectives. For instance, one participant shared a successful application, "I used a cartoon story to help my students learn pronouns, and the method generated a high level

of engagement." This highlights the use of multimedia as a tool for teaching specific language elements.

Teachers detailed diverse pedagogical objectives achieved through digital storytelling. One noted its utility in vocabulary learning: "Digital storytelling helped my students learn new vocabulary through listening." Another highlighted its effectiveness in enhancing listening and speaking skills: "I have used it to improve my students' listening and speaking skills with positive results." However, the extent of digital storytelling usage varied among participants. While some had integrated it frequently into their lessons, others were yet to employ it, as indicated by a teacher's remark, "No, I have not had the opportunity to use digital storytelling in my classroom."

The thematic analysis of the interviews also revealed a multifaceted landscape in the utilization of digital storytelling among participants, characterized by a spectrum of frequency, duration, and diverse student demographics. A teacher elucidated the selective integration of digital storytelling, noting, "Digital storytelling is a strategic choice in my classroom, especially for lessons where student engagement tends to wane. It is not a staple but a special addition for enhancing understanding of complex topics." This reflects a judicious use of digital storytelling for maximizing impact in specific instructional contexts. Another teacher provided insight into the application of digital storytelling with younger learners: "In my experience teaching lower grades, Digital Storytelling has been instrumental in capturing the fleeting attention of young minds. It is a monthly feature in my lesson plans, seamlessly blending with traditional teaching methods to enrich the learning experience." This highlights the role of digital storytelling in augmenting regular curricular activities tailored to the needs of younger students. In contrast, a teacher working with older adolescents shared a different perspective: "My approach with Digital Storytelling for students aged 15 to 16 is quite measured. I employ it for project-based learning or when unpacking more abstract concepts. It is not frequent, but its occasional use throughout the semester brings a dynamic edge to the learning process." These insights underscore digital storytelling's nuanced and varied application in EFL teaching. The teachers' narratives reveal how the flexibility of digital storytelling makes it a valuable teaching tool across various age groups and learning objectives, adaptable to the evolving requirements of diverse classroom environments.

A few teachers expressed reservations about comparing digital storytelling with traditional methods. One teacher's comment, "The first skill anyone learns is listening, not speaking," suggests a preference for conventional listening-based approaches. Despite these varied experiences and views, the prevailing sentiment among the majority favoured the effectiveness and versatility of digital storytelling in enhancing EFL instruction.

## Student Engagement and Language Skills Development in EFL through Digital Storytelling

Teachers provided illuminating insights into the impact of digital storytelling on student engagement in EFL classrooms. A substantial increase in engagement levels was a common thread in their responses. One teacher vividly described the change: "The level of student engagement was high; they were visibly more attentive and participative." Another echoed this sentiment, adding, "Students were not just listening; they were actively involved. Their participation soared, especially during storytelling sessions." This heightened engagement was particularly pronounced in smaller to medium-sized classes. A participant teaching a mid-sized class shared, "In my class of 20 students, digital storytelling transformed the learning dynamics. Students who were usually reticent participated enthusiastically." The same teacher elaborated, "It was not just about listening to stories; students became eager to share their own, which is a huge leap in engagement." These observations suggest that digital storytelling captivated students and actively involved them in learning, making the classroom environment more dynamic and interactive. Teachers attributed this to the compelling nature of storytelling and the multimedia elements that resonated well with the students, fostering a more immersive and participative educational experience.

In addition to engagement, participants also shed light on the influence of digital storytelling on language skills development. They observed improvements in critical areas such as vocabulary expansion, enhanced comprehension, and increased fluency in spoken English. For instance, one teacher observed: "Digital storytelling has significantly contributed to vocabulary building in my younger learners," while another reflected: "Students demonstrate greater speaking confidence and fluency during storytelling activities." However, the impact on language skills was not uniformly experienced across all educational contexts. Participants noted variations based on factors like the age of students and class sizes. Younger learners in smaller groups were reported to show more substantial gains in language skills, whereas in classes with older students, particularly in larger groups, the effectiveness in enhancing language skills varied more.

Teachers stressed the significance of having the proper technological and educational framework for successfully implementing digital storytelling. One participant pointed out, "Effective digital storytelling relies heavily on the right tech setup. Without interactive whiteboards or reliable internet, it is challenging to engage students fully." Another highlighted the role of the classroom environment, saying, "It is not just about the technology. The classroom atmosphere needs to be dynamic and conducive to storytelling." Challenges in catering to diverse student groups were also noted. A teacher shared, "In larger classes,

it is hard to find stories that appeal to everyone. When the content does not resonate, student engagement drops sharply." This sentiment was echoed by another teacher who observed, "Diverse interests in my class mean I need a variety of storytelling approaches to maintain effectiveness, which can be quite demanding." These reflections underline the complexities of implementing digital storytelling in varied classroom settings, particularly emphasizing the necessity for appropriate technological resources and a versatile approach to accommodate diverse student needs and interests for optimal engagement and language skill enhancement.

Teachers generally recognized digital storytelling as an effective tool for boosting student engagement in EFL classrooms. While positive, its impact on language skill development depended on various contextual factors. This underscores the need for tailored digital storytelling approaches to optimize engagement and language learning outcomes effectively.

## Challenges and Obstacles in Digital Storytelling Implementation

This theme illuminated some difficulties teachers face in successfully integrating digital storytelling into EFL classrooms. Infrastructure constraints emerged as a significant barrier, with teachers frequently noting the lack of essential technology and resources. One teacher highlighted, "Interactive whiteboards are rarely available for every lesson," and another lamented the general "lack of computing and software resources in educational settings."

Curricular limitations also posed challenges, particularly in integrating digital narratives into existing, content-heavy syllabi. "Fitting digital storytelling into our dense curriculum is challenging," one teacher observed, pointing to the intricacies of aligning narratives with specific linguistic objectives. Time constraints further compounded these difficulties, with teachers feeling pressured to deliver results efficiently, notwithstanding the substantial preparation time digital storytelling modules demand. One participant noted, "Digital storytelling takes time to prepare, which adds pressure to our tight schedules."

Class size and a lack of adequate instructional tools were additional hurdles. Statements like "The student-to-teacher ratio is quite high in some classes" and "We only have one room equipped for interactive learning" were common, reflecting the logistical challenges in implementing digital storytelling. Furthermore, some teachers expressed their limitations in creating integrated educational materials and addressing the diverse needs of their student populations.

Adding to these challenges, three teachers mentioned not utilizing digital storytelling due to resource scarcity or a lack of interest, both on their part and among students. "I have not used digital storytelling, so I cannot comment on

its benefits," admitted one teacher, while another voiced frustration about the "lack of interest or response" from students. This spectrum of experiences and perceptions underscores that while digital storytelling is a promising pedagogical tool, its effectiveness is influenced by various factors, including resource availability, teacher and student enthusiasm, and broader educational culture. Despite these challenges, it is noteworthy that some teachers encountered minimal or no significant obstacles in implementing digital storytelling, indicating a diverse range of experiences across different educational settings.

## Benefits and Advantages of Digital Storytelling in EFL Teaching

Teachers praised its versatility and adaptability in various learning environments. One teacher highlighted, "Digital storytelling facilitates the easy storage and reuse of educational materials," which is particularly beneficial in settings with limited resources. This aspect of digital storytelling allows for efficient content recycling, making it a cost-effective teaching tool. Another teacher lauded digital storytelling as "an advanced and efficient pedagogical instrument," pointing out its capability to streamline and enhance learning. This efficiency is seen in how digital stories can condense complex information into engaging, digestible formats, making learning more accessible for students.

A significant benefit noted was the role of multimedia elements, like audio and video, in enriching learners' sensory and cognitive experiences. Such multimedia integration caters to different learning styles and preferences, potentially leading to higher retention rates and a deeper understanding of the subject matter. One teacher observed, "The use of multimedia in digital storytelling has visibly increased student engagement and interaction," suggesting that this approach can effectively capture and maintain students' attention, a crucial element in effective language teaching. Moreover, teachers recognized the potential of digital storytelling to foster a more interactive and participative classroom environment. One teacher said: "Students become active learners rather than passive recipients of information, engaging with content in a way that encourages deeper involvement and understanding."

The consensus among the teachers was that digital storytelling, with its blend of multimedia elements and engaging narratives, offers a rich, dynamic, and efficient method for language teaching. It not only aids in appealingly presenting information but also significantly contributes to an interactive and immersive learning experience, enhancing the overall quality of EFL education.

# Discussion

## Discussion of Research Question 1

The findings were predominantly skewed positive when examining the first research question about EFL teachers' perceptions of digital storytelling in Jordan. This favourable view is likely influenced by several key factors inherent to the educational context in Jordan. Digital storytelling signifies a transformative shift from conventional teaching methodologies, aligning with the contemporary emphasis on dynamic, student-centred learning approaches. This evolution is well-documented in educational research (Robin, 2008; Yang & Wu, 2012), supporting its growing adoption and effectiveness in diverse learning environments. Teachers explicitly appreciated how digital storytelling revitalizes conventional teaching, with one stating, "It brings a fresh perspective to our vocabulary lessons and enhances language skills like listening and speaking interactively."

The alignment of digital storytelling with specific educational outcomes in EFL was another critical factor contributing to its positive reception. Teachers found that digital storytelling effectively dovetails with essential language learning objectives. This is evident in comments like, "Digital storytelling directly supports our curriculum goals, particularly in engaging students more deeply with the language." Teacher preparedness and proficiency in utilizing digital tools also play a significant role. Those with training or a natural inclination towards technology in education might perceive digital storytelling more positively, seeing it as a valuable tool in their instructional arsenal. Furthermore, the cultural relevance of digital storytelling content and its role in enhancing student engagement cannot be overlooked. Customized stories that reflect local contexts and students' interests likely increase engagement and make learning more meaningful and enjoyable. A teacher highlighted this aspect: "Students connect better with culturally relevant and engaging stories, making the learning experience more impactful." Generally, this broad optimism is consistent with prior research highlighting the benefits of digital storytelling in language teaching (Fan, 2022; Fitri et al., 2022).

However, despite these positive aspects, a minority of teachers expressed reservations, primarily due to practical implementation challenges like limited resources and integration issues with the existing curriculum. Such concerns are echoed in broader educational technology literature (Canals & Al-Rawashdeh, 2018; Haga, 2023) and reflect the need for adequate support and resources for successful implementation. To put it briefly, while the general percep-

tion of digital storytelling in EFL instruction is overwhelmingly positive, as evidenced by the teachers' testimonies, it is nuanced by practical considerations. The findings suggest that addressing these challenges and leveraging the method's alignment with educational goals and cultural relevance are crucial to maximizing its potential in the EFL classroom.

## Discussion of Research Question 2

In exploring the impact of digital storytelling on student engagement and language proficiency, the study reveals a multifaceted array of outcomes influenced by various factors. Teachers reported a noticeable increase in student engagement, which aligns with research endorsing interactive learning environments (Mansi, 2019). For example, one teacher noted, "Digital storytelling has significantly increased student participation and attention in my class," suggesting the method's ability to engage students more profoundly. The variation in engagement levels among different groups can be attributed to individual learning styles and technological familiarity. Students accustomed to digital environments likely find digital storytelling more engaging, while those with limited tech exposure may experience a steeper learning curve. Moreover, socio-economic background also plays a role, as students with greater access to technology outside of school are more likely to engage actively with digital storytelling methods in the classroom. This disparity highlights the need for equitable access to technological resources to ensure all students benefit from digital learning tools.

The study's findings also underline significant improvements in language proficiency, especially in vocabulary and conversational skills, as a direct result of integrating digital storytelling in EFL classrooms. Teachers have observed notable advancements, with one teacher remarking, "Since integrating digital storytelling, there has been a noticeable improvement in students' vocabulary and conversational abilities." This observation aligns with the effectiveness of multi-modal language-learning approaches as highlighted by Tran (2021). The immersive quality of digital storytelling, which synthesizes visual, auditory, and narrative elements, offers a contextual and engaging language learning experience. It not only aids in the retention and application of language skills but also fosters active learner participation. This participatory aspect is crucial, leading to deeper cognitive engagement and enhancing language retention, thereby underscoring the method's critical role in developing language skills.

Furthermore, the success of digital storytelling in fostering language proficiency is significantly influenced by external factors, as highlighted by teachers. The availability of necessary resources, such as interactive whiteboards, and a conducive learning environment are pinpointed as essential for maximizing

the pedagogical benefits of digital storytelling. This reliance on contextual elements for effective digital storytelling implementation resonates with previous findings in educational technology research (Rusda & Puspitasari, 2023; Sadik, 2008), which emphasize that educational technology's impact often hinges on supportive external conditions. Therefore, the efficacy of digital storytelling in enhancing language skills is not only a function of its inherent educational features but also depends on the broader educational ecosystem in which it is implemented. In a nutshell, while digital storytelling has positively impacted student engagement and language proficiency, its effectiveness is mediated by a constellation of factors. These include not only the individual characteristics of students but also the broader infrastructural and institutional context. Teachers' optimism towards the advantages of digital storytelling underscores its potential as an instructional tool. However, this potential is fully realized only when these diverse elements are adequately addressed, a sentiment that resonates with broader discussions in educational technology.

## Discussion of Research Question 3

A multifaceted picture emerges in addressing the third research question, which probes the specific challenges and advantages Jordanian EFL teachers face in integrating digital storytelling. A key challenge identified by teachers is the lack of adequate infrastructure, notably the absence of interactive whiteboards and sufficient processing resources in some schools. This issue, as one teacher vividly described, "limits our ability to engage students with digital stories fully," reflects systemic flaws within the educational sector, a concern echoed in prior research (Aljaberi, 2021).

Furthermore, the rigidity of the existing curriculum was frequently cited as a barrier to integrating digital storytelling effectively. Teachers noted that "fitting digital stories into our tightly-packed curriculum is a constant struggle," indicating institutional limitations favouring traditional instructional approaches and a potential gap in preparing teachers for innovative methodologies. Time constraints also surfaced as a significant concern, aligning with broader educational discourse emphasizing the need for careful planning and resource allocation in adopting new teaching methods (Lim et al., 2022).

On the flip side, teachers identified several advantages of digital storytelling, which offer a counterbalance to these challenges. The ability to save and reuse digital storytelling materials was highlighted as a precious feature, especially in resource-limited contexts. As one teacher said, "The reusability of digital stories helps us overcome some resource challenges." This practical aspect of digital storytelling contributes to its appeal as an effective instructional tool. Teachers also underscored the versatility of digital storytelling in achieving

diverse educational objectives, from skill enhancement to cognitive engagement. “Digital storytelling allows us to target multiple learning goals engagingly,” shared a teacher, reflecting the method’s capacity to cater to different pedagogical needs. This advantage aligns with the growing body of research supporting multimodal, interactive methods in education for enriching learning experiences (Belda-Medina, 2021; Indriani & Suteja, 2023).

In summary, exploring the third research question reveals a landscape where challenges coexist with significant advantages, shaping teachers’ perceptions and implementing digital storytelling in EFL settings. Despite facing infrastructural, curricular, and time-related hurdles, the method’s adaptability, practical benefits, and pedagogical effectiveness present compelling reasons for its integration into teaching practices. This intricate back-and-forth of difficulties and advantages, embedded within the broader context of educational technology, underlines the evolving nature of teaching methodologies and the need to navigate these dynamics effectively.

## Implications

This study sheds light on significant implications for the theoretical framework and practical execution across educational domains, notably in pedagogical methodologies, curriculum innovation, and policy formulation.

### Theoretical Implications

The favourable reception of digital storytelling among teachers in Jordan highlights its transformative potential in language education methodologies. This enthusiasm underscores the ability of digital storytelling to augment the EFL curriculum, laying a theoretical foundation for incorporating technology-driven educational tools. It challenges conventional teaching models, promoting a transition towards interactive, learner-centric educational environments. This study contributes substantially to the ongoing dialogue about technology’s role in education, advocating for a critical reassessment of current pedagogical approaches by showcasing the capacity of digital storytelling to engage learners actively and bolster language acquisition.

### Practical Implications

Identified challenges in curriculum enhancement, notably the absence of vital technological infrastructure and difficulties in aligning digital storytelling with established curriculum objectives, are pivotal. Educational bodies and policymakers are urged to prioritize allocating resources to bridge these gaps, whether by securing investments for the requisite technology or initiating professional development programs to equip teachers with the skills necessary for implementing digital storytelling effectively. A tangible example of the application of digital storytelling in EFL classrooms involves students crafting digital narratives that reflect personal experiences or align with curriculum themes. An assignment might require students to create a digital story about a cultural festival, integrating research, narrative development, and multimedia production. This approach influences theoretical language skills in tangible scenarios and cultivates cultural sensitivity and digital literacy. The efficacy of digital storytelling is contingent upon the availability of resources and institutional backing. Implementing policies that support integrating digital storytelling's curriculum—such as financial investment in technology and teacher training—is critical. Schools might, for example, introduce policies that ensure access to tablets or digital storytelling software, guaranteeing that all learners can engage in these enriching educational activities. Achieving the full educational potential of digital storytelling in EFL teaching necessitates a unified approach among educators, administrators, and policymakers. This collaborative effort should address logistical and infrastructural challenges, creating an environment conducive to the success of digital storytelling as an educational tool.

## Limitations and Recommendations

The study has various limitations that should be considered and pave the way for further research. First, the study is limited to EFL teachers in Jordan, which may limit its relevance to broader situations. Further studies may benefit from a more diverse participant pool, possibly extending the research to different school systems or nations. Second, the study relies heavily on qualitative interview data, which may introduce subjective biases. Future studies could improve the validity of findings by combining qualitative and quantitative data sources, such as standardised tests or surveys. Third, the study focuses entirely on teachers' perspectives, ignoring students' experiences and reactions to digital storytelling. Future efforts should aim for a more thorough understanding

by including the students' perspectives. Considering these limitations, future studies could better understand how digital storytelling affects EFL education.

## Conclusion

In conclusion, the research sheds significant light on the complexities and opportunities associated with using digital storytelling in Jordanian EFL classrooms. The study affirms a generally positive perception among teachers regarding the utility of digital storytelling as a tool for engagement and improving specific language skills such as vocabulary acquisition, speaking, and listening. These findings support the developing consensus regarding the advantages of interactive, multimedia-based pedagogies by corroborating the positive evaluations found in prior educational research. However, the study also highlights the numerous obstacles teachers experience when incorporating digital storytelling into their curricula. These range from infrastructural limitations and lack of resources to curricular constraints and time pressures. These obstacles align with broader issues identified in the literature on educational technology, indicating that teachers confront universal challenges outside of Jordan. The research provides a well-rounded comprehension of the function and influence of digital storytelling in EFL classrooms. It emphasises its potential to considerably enhance language learning experiences and the practical obstacles that impede its widespread adoption. Thus, the study is essential to the discourse on educational technologies, providing educators, policymakers, and future researchers with actionable insights.

**Funding**

'The authors extend their appreciation to the Deanship of Scientific Research at Northern Border University, Arar, KSA for funding this research work through the project number "NBU-FFR-2024-2905-03."'

# References


Abdolmanafi-Rokni, S. J. (2014). Digital storytelling in EFL classrooms: The effect on the oral performance. *International Journal of Language and Linguistics, 2*(4), 252. https://doi.org/10.11648/j.ijll.20140204.12

Alemi, M., Salmani Givi, S., & Rezanejad, A. (2022). The role of digital storytelling in EFL students' writing skill and motivation. *Language Teaching Research Quarterly*, *32*, 16–35. https://doi.org/10.32038/ltrq.2022.32.02

Aljaberi, N. (2021). Perceptions and beliefs of the teachers of kindergarten and the first primary stage for employing digital technologies in the education process in Jordan. *International Journal of Progressive Education, 17*(5), 87–101. https://doi.org/10.29329/ijpe.2021.375.7

Aljaraideh, Y. A. (2020). The impact of digital storytelling on academic achievement of sixth grade students in English language and their motivation towards it in Jordan. *Turkish Online Journal of Distance Education, 21*(1), 73–82. https://doi.org/10.17718/tojde.690345

Alkhaldi, A. A., & Kayapinar, U. (2022). English textbook challenge in Jordan: An in-depth analysis. *International Journal of English Language and Literature Studies, 11*(4), 155–166. https://doi.org/10.55493/5019.v11i4.4616

Al-Sharah, N., Awwad, S., Abd Hamid, J., Shukri, S. M., & Ariffin, I. A. (2021). EFL teachers' perceptions towards English language teaching during the pandemic: Jordanian perceptive. *Eurasian Journal of Applied Linguistics, 7*(2), 171–185.

Alzubi, A. (2023). The role of multimedia tools in Hashemite Kingdom of Jordan education classroom teaching in the digital era. *European Journal of Interactive Multimedia and Education, 4*(2), e02303. https://doi.org/10.30935/ejimed/13378

Arvanitis, P., & Krystalli, P. (2021). Digital storytelling and augmented reality-based scenarios for foreign language teaching. In M. Montebello (Ed.), *Digital Pedagogies and the Transformation of Language Education* (pp. 1–22). IGI Global.

Belda-Medina, J. (2021). Promoting inclusiveness, creativity and critical thinking through digital storytelling among EFL teacher candidates. *International Journal of Inclusive Education, 26*(2), 109–123. https://doi.org/10.1080/13603116.2021.2011440

Canals, L., & Al-Rawashdeh, A. (2018). Teacher training and teachers' attitudes towards educational technology in the deployment of online English language courses in Jordan. *Computer Assisted Language Learning, 32*(7), 639–664. https://doi.org/10.1080/09588221.2018.1531033

Castillo-Rodriguez, C., Díaz, I. C. S., & Lage, J. M. D. (2022). Digital storytelling for EFL preschool classroom: Tools, activities and themes proposed by preservice teachers. *Espiral. Cuadernos del Profesorado, 15*(31), 82–93. https://doi.org/10.25115/ecp.v15i31.8087

Christiansen, M. S., & Koelzer, M. L. (2016). Digital storytelling: Using different technologies for EFL. *MexTESOL Journal, 40*(1), 1–14.

Churchill, N. (2020). Development of students' digital literacy skills through digital storytelling with mobile devices. *Educational Media International, 57*(3), 271–284. https://doi.org/10.1080/09523987.2020.1833680

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education*. Routledge.

Dewi, P., & Sari, D. E. (2022). Perception of digital storytelling in overcoming fear for speaking English through interdisciplinary project of gender issues. *IDEAS: Journal on English Language Teaching and Learning, Linguistics and Literature, 10*(2). https://doi.org/10.24256/ideas.v10i2.2748

Dreher, T. (2012). A partial promise of voice: Digital storytelling and the limits of listening. *Media International Australia*, *142*(1), 157–166. https://doi.org/10.1177/1329878X1214200117

Emmel, N. (2013). *Sampling and choosing cases in qualitative research*. SAGE.

Fan, Y. S. (2022). Facilitating content knowledge, language proficiency, and academic competence through digital storytelling: Performance and perceptions of first-year medical-related majors. *Journal of Research on Technology in Education,* 1–21. https://doi.org/10.1080/15391523.2022.2110337

Fitri, H. A., Mohamad, M., Harianingsih, I., & Wadi, A. S. (2022). Digital storytelling-based tasks on speaking classroom; seen from ESL learners' lenses. *Jurnal Tatsqif, 20*(2), 180–198. https://doi.org/10.20414/jtq.v20i2.6289

Haga, S. (2023). Digital transformation in foreign language education: Examining the position of the EFL teacher in ICT literacy diffusion. *Studies in Technology Enhanced Learning, 3*(2). https://doi.org/10.21428/8c225f6e.d8ba13fb

Huang, H. D. (2022). Examining the effect of digital storytelling on English speaking proficiency, willingness to communicate, and group cohesion. *TESOL Quarterly, 57*(1), 242–269. https://doi.org/10.1002/tesq.3147

Indriani, S., & Suteja, H. (2023). Fostering reading interest through digital storytelling for young learners in the early childhood. *Journal of Education and Learning (EduLearn), 17*(2), 301–306. https://doi.org/10.11591/edulearn.v17i2.18372

James, P. R. A. P., Yong, K. L., & Yunus, M. M. (2019). Hear me out! Digital storytelling to enhance speaking skills. *International Journal of Academic Research in Business and Social Sciences, 9*(2), 190–202. https://doi.org/10.6007/ijarbss/v9-i2/5533

Karacan, C. G., & Akoğlu, K. (2021). Educational augmented reality technology for language learning and teaching: A comprehensive review. *Shanlax International Journal of Education, 9*(2), 68–79. https://doi.org/10.34293/education.v9i2.3715

Kazazoglu, S., & Bilir, S. (2021). Digital storytelling in L2 writing: The effectiveness of Storybird Web 2.0 tool. *Turkish Online Journal of Educational Technology-TOJET, 20*(2), 44–50.

Kim, D., Park, H. R., & Vorobel, O. (2023). Enriching middle school students' learning through digital storytelling: A multimodal analytical framework. *ECNU Review of Education*, *7*(2). https://doi.org/10.1177/20965311231182159

Lim, N. Z. L., Zakaria, A., & Aryadoust, V. (2022). A systematic review of digital storytelling in language learning in adolescents and adults. *Education and Information Technologies, 27*(5), 6125–6155. https://doi.org/10.1007/s10639-021-10861-0

Luan, L., Yi, Y., Hwang, G. J., Dong, Y., Jing, B., & Liu, J. (2023). Facilitating EFL learners' willingness to communicate amidst the pandemic: A digital storytelling-based online flipped learning approach. *Innovation in Language Learning and Teaching*, 1–16. https://doi.org/10.1080/17501229.2023.2241436

Mansi, G. K. A. (2019). The effectiveness of using digital stories in developing listening skills among ninth grade female students in Jordan. *The International Journal for Talent Development, 10*(1), 3–17. https://doi.org/10.20428/ijtd.v10i1.1520

Masters, J. (2020). Digital storytelling, information, and education. In A. Tatnall (Eds.), *Encyclopedia of Education and Information Technologies*. Springer, Cham. https://doi.org/10.1007/978-3-030-10576-1_196

Maureen, I. Y., van der Meij, H., & de Jong, T. (2020). Enhancing storytelling activities to support early (digital) literacy development in early childhood education. *International Journal of Early Childhood, 52*, 55–76.

McLellan, H. (2007). Digital storytelling in higher education. *Journal of Computing in Higher Education, 19*, 65–79. https://doi.org/10.1007/bf03033420

Meletiadou, E. (2022). Using educational digital storytelling to enhance multilingual students' writing skills in higher education. *IAFOR Journal of Education, 10*(2), 111–130. https://doi.org/10.22492/ije.10.2.06

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*, 29–40. https://doi.org/10.1016/j.compedu.2013.07.033

Neuman, W. (2014) *Social research methods: Qualitative and quantitative approaches*. Pearson.

Oakley, G., Pegrum, M., Xiong, X. B., Lim, C. P., & Yan, H. (2018). An online Chinese-Australian language and cultural exchange through digital storytelling. *Language, Culture and Curriculum, 31*(2), 128–149. https://doi.org/10.1080/07908318.2017.1386193

Octaviani, A., & Hesmatantya, V. (2018). Using Teaching Proficiency through Reading and Storytelling (TPRS) in teaching English for young learners. *Tell: Teaching of English Language and Literature Journal, 6*(2), 78. https://doi.org/10.30651/tell.v6i2.2134

Orak, D. S., & Al-khresheh, M. (2021). In between 21st century skills and constructivism in ELT: Designing a model derived from a narrative literature review. *World Journal of English Language, 11*(2), 166–176. https://doi.org/10.5430/wjel.v11n2p166

Pandhiani, S. M. (2022). Toward empowered identities: An exploration of female ESL learners' literacy through digital storytelling. *Journal of Asian and African Studies*, 002190962211259. https://doi.org/10.1177/00219096221125910

Ramalingam, K., Jiar, Y. K., & Mathiyazhagan, S. (2022). Speaking skills enhancement through digital storytelling among primary school students in Malaysia. *International Journal of Learning, Teaching and Educational Research, 21*(3), 22–35. https://doi.org/10.26803/ijlter.21.3.2

Robin, B. R. (2008). Digital storytelling: A powerful technology tool for the 21st century classroom. *Theory Into Practice, 47*(3), 220–228. https://doi.org/10.1080/00405840802153916

Robin, B. R. (2016). The power of digital storytelling to support teaching and learning. *Digital Education Review,* 30, 17–29.

Rusda, F. S., Ali, F., & Puspitasari, D. (2023). Indonesian young EFL learners' experiences in learning environmental literacy through digital storytelling: a case study. *Pedagogy: Journal of English Language Teaching, 11*(1), 59–74. https://doi.org/10.32332/joelt.v11i1.5252

Sadik, A. (2008). Digital storytelling: A meaningful technology-integrated approach for engaged student learning. *Educational Technology Research and Development*, *56*(4), 487–506. https://doi.org/10.1007/s11423-008-9091-8

Salem, A. A. M. S. (2022). Multimedia presentations through digital storytelling for sustainable development of EFL learners' argumentative writing skills, self-directed learning skills and learner autonomy. *Frontiers in Education, 7*. https://doi.org/10.3389/feduc.2022.884709

Setiyorini, T. J. (2020). The effect of using digital storytelling on students' intrinsic motivation for learning vocabulary. *Didaktis: Jurnal Pendidikan Dan Ilmu Pengetahuan, 20*(1), 54–65. https://doi.org/10.30651/didaktis.v20i1.4336

Shen, X., Hao, C., & Peng, J. E. (2022). Promoting EFL learners' willingness to communicate through transmediation in a digital storytelling workshop. *Journal of Multilingual and Multicultural Development*, 1–18. https://doi.org/10.1080/01434632.2022.2086257

Tatli, Z., Saylan, E., & Kokoç, M. (2022). Digital storytelling in an online EFL course: Influences on speaking, vocabulary, and cognitive load. *Participatory Educational Research*, *9*(6), 89–112. https://doi.org/10.17275/per.22.130.9.6

Tran, D. T. (2021). Storytelling to enhance speaking and listening skills for English young learners: A case study at language centers in Binh Duong province. *Vietnam Journal of Education, 5*(2), 29–36. https://doi.org/10.52296/vje.2021.64

Urbieta, A. S., & Peñalver, E. A. (2021). Multimodal discourse in digital storytelling: An assessment tool proposal. *Computer Assisted Language Learning Electronic Journal, 22*(2), 14–25. https://reunir.unir.net/handle/123456789/11519

Wu, J., & Chen, D. T. V. (2020). A systematic review of educational digital storytelling. *Computers & Education*, *147*, 103786. https://doi.org/10.1016/j.compedu.2019.103786

Yang, Y. T. C., & Wu, W. C. I. (2012). Digital storytelling for enhancing student academic achievement, critical thinking, and learning motivation: A year-long experimental study. *Computers & Education, 59*(2), 339–352. https://doi.org/10.1016/j.compedu.2011.12.012

Yastibaş, A. E., Baturay, M. H., & Çinar, C. A. (2022). Augmented reality-based digital storytelling in language teaching and learning: A suggested lesson plan. *Emerging Concepts in Technology-Enhanced Language Teaching and Learning*, 42–65. https://doi.org/10.4018/978-1-7998-8981-6.ch003

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.

Yu, S., & Lu, Y. (2021). *An introduction to artificial intelligence in education*. Springer Nature

Yuliani, S., & Hartanto, D. (2022). Digital online learning by using digital storytelling for preservice teacher students. *International Journal of Language Education, 6*(3), 221. https://doi.org/10.26858/ijole.v6i3.20408

Zarifsanaiey, N., Mehrabi, Z., Kashefian-Naeeini, S., & Mustapha, R. (2022). The effects of digital storytelling with group discussion on social and emotional intelligence among female elementary school students. *Cogent Psychology, 9*(1). https://doi.org/10.1080/23311908.2021.2004872

Zhong, Y. X., & Shen, H. Z. (2002). Where is the technology-induced pedagogy? Snapshots from two multimedia EFL classrooms. *British Journal of Educational Technology, 33*(1), 39–52. https://doi.org/10.1111/1467-8535.00237

Zuhriyah, M. (2023). Students' perceptions on a multimodal digital storytelling book use in a writing class during covid-19 pandemic.w *English Language Teaching, Applied Linguistic and Literature, 4*(1), 1–10. https://doi.org/10.21154/eltall.v4i01.5451

Appendix 1

# The Interview

**Demographic Information**

| | | | | |
|---|---|---|---|---|
| Gender | □ Male | | □ Female | |
| Qualifications | □ Bachelor's | □ Master's | □ Doctorate | |
| Teaching Experience | □ 1–5 | □ 6–10 | □ 11–15 | □ More than 15 |

**Interview Questions**

1. How do you see the role of digital storytelling in enhancing the teaching and learning of English in your classroom?
2. Could you describe your strategies for integrating digital storytelling into your EFL lessons? How do you adapt these methods for different classroom contexts or age groups, and what specific goals do you aim to achieve with this approach?
3. In your experience, how has digital storytelling influenced not only student engagement but also their language skills in your classroom? Can you share examples where you have observed notable changes in student participation, interest levels, and improvements in specific language abilities during digital storytelling activities?
4. What are the challenges you have faced in integrating digital storytelling into the English language curriculum? This could include technical issues, student engagement, or educational concerns.
5. In your opinion, what are the main advantages of using digital stories in the English language classroom? Have you noticed any specific improvements in students' language skills, engagement, or other areas?

Theory and Practice of Second Language Acquisition
vol. 11 (1), 2025, pp. 1/5
https://doi.org/10.31261/TAPSLA.17036

CC BY SA

# Rachelle S. Savitz, Leslie D. Roberts, & Jason DeHart (Eds.), *Teaching Challenged and Challenging Topics in Diverse and Inclusive Literature: Addressing the Taboo in the English Classroom* Routledge, 2023, 218 pp.

Considering the sheer amount of fresh American and European publications, it is safe to say there is a resurgence of interest in the interdisciplinary field of addressing taboo and controversial issues in (language) education. Upon closer reading of the scholarly work in either general education (e.g., Hess, 2009; Kubota, 2014; Ortega-Sánchez, 2022) or foreign language education in particular (e.g., Ludwig & Summer, 2023), it is also apparent that the majority of authors draw heavily on theory rather than illustrate how it translates into teaching practice, even though they recognize the latter as a primary concern for the general readership. To see an exception to this trend, let us turn to a refreshing volume edited by Rachelle S. Savitz, Leslie D. Roberts, and Jason DeHart, in which the authors go a long way in closing the notorious theory-practice gap. Although *Teaching Challenged and Challenging Topics in Diverse and Inclusive Literature: Addressing the Taboo in the English Classroom* is based entirely on the contributions of US authors responding to recent challenges in US literacy and language arts education, it addresses issues of global relevance and there are many reasons for me to bring this book to the attention of English language teachers (at all grade levels) and teacher educators in other contexts as well.

Firstly, it is remarkable how the authors go about addressing taboo and controversial issues in non-compartmentalizing ways and describe classroom activities in which students are led to understand the complex interplay and co-existence of such issues with other aspects of social/personal life in particular

settings. For example, Chapter 4 (Dail, Koch, Witte, & Vandever) suggests that if teachers wish to discuss the topic of non-normative gender identities in class (cf. Pakuła, Pawelczyk, & Sunderland, 2015), they should not do so by separating gender from other identity-relevant concepts. More specifically, in describing *Lily and Dunkin* (Gephart, 2016) as a young adult novel suitable for gender-focused instruction, the authors explain their choice of text as follows:

> While the novel deals with serious issues such as gender identity and bipolar disorder, it does so in a manner that focuses on friendship and adolescent struggles accompanying that at its core. The characters experience typical middle school issues such as bullying and family tensions. The novel also expresses the tensions adolescents experience between authentically balancing who they truly are on the inside with the persona they project to the world. (p. 52)

Similarly, Chapter 6 (Kruep & Popov) indexes several pieces of young adult literature in which clusters of taboo and controversial issues are used to portray how they operate conjointly in characters' lives and actions. Consider, for instance, racism, sexual assault, sexism, language barriers, financial distress, and discrimination as one cluster in *The House on Mango Street* (Cisneros, 1984). Clearly, the use of such texts creates a space for rich follow-up discussions, in which the study of multiple perspectives and aspects allows students to develop complex understandings of the issues at hand. It is also typical of most chapters that the authors present extensive lists—rather than singular examples—of texts for students at all grade levels, along with ideas and materials for related classroom activities.

Secondly, it is appealing how, instead of sugarcoating, the editors and authors advise teachers to haul not only teaching practices from the particular chapters, but also safeguarding practices which can help them respond appropriately to the resistance or challenges they might meet in the micro- or macro-contexts of schools. Thus, in Chapter 7, as part of their engagement with taboo and controversial issues, Smith and Warren suggest that teachers should become familiar with "policy and standards documents in which social justice standards are explicitly signaled and can be used for safeguarding practices" (p. 95; see description of the social justice standards taken from Learning for Justice, 2021). Similarly, in Chapter 11, Schucker draws attention to the framework of global competency standards (OECD, 2018), urging teachers to acknowledge "global and intercultural competence as a desirable outcome of a 21st century education" and inviting students to "investigate the world, recognize perspectives, communicate ideas, and take action" (p. 143). Further, the authors in Chapter 4 finish off with a step-by-step guide for teachers going

through a complaints procedure regarding classroom materials and activities, and offer a checklist for preventive action.

Thirdly, and in contrast with the publications cited earlier (Hess, 2009; Kubota, 2014; Ludwig & Summer, 2023; Ortega-Sánchez, 2022), it is laudable how the twelve chapters in this book bring forth one prime argument—literature is a fundamental gateway to learning about taboo and controversial issues—by way of amassing the teaching practices and perspectives of multiple stakeholders, including classroom practitioners, school librarians, teacher educators, and students at all grade levels. Regardless of which grade level or issue category is in focus, each chapter grows out of Sims Bishop's (1990, p. 9) seminal idea of depicting stories as mirrors, windows, and sliding glass doors through which readers view, enter, and think about "worlds that may be real or imagined, familiar or strange." Thus, the authors bring together an exciting set of contemporary literary resources, ranging from interlingual and biographical picture books (Chapters 1, 2, and 8) through young adult literature (especially Chapters 3, 4, 6, and 7) to graphic novels, in which words and images are combined to engage struggling and unmotivated readers as well (Chapter 12). An interesting addition to these is the discussion, in Chapter 3 (Waymouth, Newvine, Fleming, Margolis, Mellon, & Middaugh), of multivoiced narratives which "de-center the importance of a singular character's identities, perceptions, and realities" (p. 34) and examine a given subject from the viewpoints of multiple individuals.

The construction of the authors' central argument progresses along descriptions of teaching from the lowest to the highest grades, with regular additions of theoretical and practical insights. Starting off with a call for early intervention, Buchholz and Garcia Reyes in Chapter 1 present scenes from a young learners' class to demonstrate how "books shared in classrooms send messages to children about which parts of themselves are valid and welcomed resources for school literacy engagements and what parts of themselves should be left at the classroom door" (p. 6). In Chapter 8, López-Robertson and del Rocio Herron argue that through the use of picture books and teacher scaffolding, even pre-kindergarten students can be prompted to think about what it means to be bilingual and live in a bilingual family, why families might move from one country to another, and what it means to have pride in one's linguistic and racial heritage. It is not surprising, then, that the authors of all chapters demonstrate an engagement with education for social justice, diversity, and equity, and a dedication to classroom instruction which is susceptible to all types and aspects of learner identity.

Regarding this type of instruction in the upper grades, the authors address a range of complex and interrelated concepts which they think students should learn about—consider, for instance, the discussions of social justice, social capital, socioeconomic status, and gentrification in Chapter 5 (Bianchi-

Pennington & Banack), of disabilities (e.g., dyslexia) and other related disorders (e.g., ADHD, autism) in Chapter 9 (Poynter & Savitz), and of queerness (e.g., in individuals, families, and communities) and gender norms in Chapter 10 (K. N., H. S., Carter, & Villanueva). To get a sense of how the authors lead students towards unpacking such concepts, it is worth looking at the types of questions posed—for instance, in Chapter 2 (Bentley, Broemmel, & Douglass)—to first- and third-graders: What makes you unique? Have you ever felt invisible? Why should we help people feel seen?

With the layering of such examples across grade levels, the twelve chapters seek to bring readers to two important conclusions. The first is that rather than teachers being banned from addressing taboo and controversial issues, they should be supported in selecting age- and level-appropriate texts and materials to gradually prepare students to confront those issues in informed and culturally responsible ways. The second is that education should focus not only on a value-free training of cognitive skills and intellect but also on (1) building communities inclusive of various identities, (2) engaging students morally and emotionally, and (3) fostering a reflexive stance to self and to social phenomena—all with respect to literacy and literature-based activities.

Coherent on the whole and relevant from beginning to end, the argumentation in the volume prompts three critical remarks. First, although the editors offer a rationale for sorting the chapters into three sections, the organization of the volume is not entirely clear. For instance, picture books as a teaching tool are discussed in Chapters 1–2 and then again in Chapter 8, gender-focused content is presented in Chapter 4 and picked up again in Chapter 10, and some other themes, in general, are simply revisited in the book's last section (or so it seems). In my view, it would have been better to put together chapters which focus either on a given age group, a specific literary genre, or a specific issue category. Second, despite the unquestionable merits of all the authors working on a joint set of arguments and concepts, the result of it, in the book's present form, is some repetitiveness (especially theory-wise) across the chapters. Again, this issue may have been resolved by structuring the chapters or sections differently. Third, something that many English language teachers might miss when reading this book is some guidance on how the suggested activities, especially if adopted for regular and long-term use, should be integrated into the main curricular content assigned to particular grade levels (see Gear, 2018, for comparison). If, as I suggest, the contents of the book are transferred to ELT contexts where little or no curricular time is dedicated to literacy instruction per se, then the question of which issue categories to cover and at what length (see Hess, 2009) will need to be addressed as well.

All in all, I have found it timely to come across a volume which foregrounds the voices of classroom practitioners well-versed in fostering critical literacy practices within socio-cultural, political, and educational settings where such

practices are increasingly pushed back rather than supported. Besides speaking up for teacher autonomy and expertise in assessing and dealing with issues relevant to students, the authors offer practically oriented chapters replete with engaging vignettes, transparent tables and visuals depicting learning materials and products, and links to websites for further resources. The content of the book, I believe, is a useful complement to current models of ELT as well as literacy and language arts instruction, in which student-centered and inquiry-based learning, social and emotional skills, and comprehension instruction are brought to the fore (see Gear, 2018). Finally, it is crucial to add that the authors' construction of an extended argument is significant both in (1) framing literature as a stepping stone to 21st century instruction that sensitizes youth to complex real-life issues and (2) shaping the emerging field of taboo and controversial issues, in which current grassroots ideas are slow to grow into coherent frameworks for supporting future research and instructional practice.

*Kornél Farkas*
https://orcid.org/0000-0002-1783-4656
Avicenna International College, Budapest, Hungary

## References

Cisneros, S. (1984). *The house on Mango Street*. Vintage Books.

Gear, A. (2018). *Powerful understanding: Helping students explore, question, and transform their thinking about themselves and the world around them*. Pembroke Publishers.

Gephart, D. (2016). *Lily and Dunkin*. Delacorte Press.

Hess, D. (2009). *Controversy in the classroom: The democratic power of discussion*. Routledge.

Kubota, R. (2014). "We must look at both sides"—But a denial of genocide too? Difficult moments on controversial issues in the classroom. *Critical Inquiry in Language Studies, 11*(4), 225–251. https://doi.org/10.1080/15427587.2014.968071

Learning for Justice. (2021). *Social justice standards*. Retrieved from https://www.learningforjustice.org/frameworks/social-justice-standards

Ludwig, C., & Summer, T. (Eds.) (2023). *Taboos and controversial issues in foreign language education: Critical language pedagogy in theory, research and practice*. Routledge.

OECD. (2018). *Preparing our youth for an inclusive and sustainable world: The OECD PISA global competence framework.* Retrieved from https://www.oecd.org/education/Global-competency-for-an-inclusive-world.pdf

Ortega-Sánchez, D. (Ed.) (2022). *Controversial issues and social problems for an integrated disciplinary teaching.* Springer.

Pakuła, Ł., Pawelczyk, J., & Sunderland, J. (2015). *Gender and sexuality in English language education: Focus on Poland*. British Council.

Sims Bishop, R. (1990). Mirrors, windows, and sliding glass doors. *Perspectives, 6*(3), ix–xi.

## 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 | **Centered, Bold, Title Case Heading**<br>Text begins as a new paragraph. |
| 2 | **Flush Left, Bold, Title Case Heading**<br>Text begins as a new paragraph. |
| 3 | ***Flush Left, Bold Italic, Title Case Heading***<br>Text begins as a new paragraph. |
| 4 | **Indented, Bold, Title Case Heading, Ending With a Period.** Text begins on the same line and continues as a regular paragraph. |
| 5 | ***Indented, Bold Italic, Title Case Heading, Ending With a Period.*** 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.

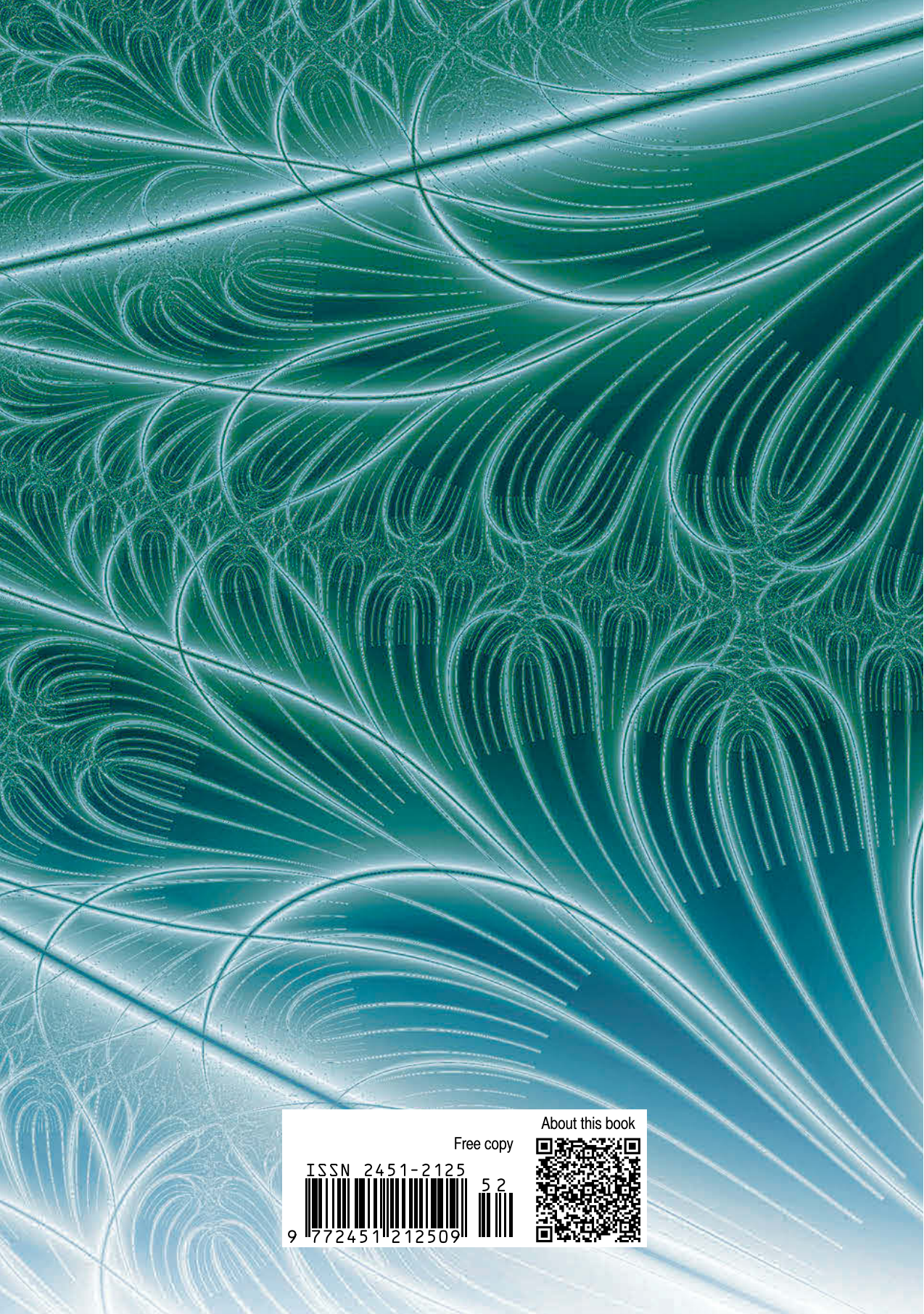
Reachel, L. H. (2001). *Native languages and toponyms: Origins, meaning, and use* (Doctoral dissertation). Available from ProQuest dissertation and theses database. (Document ID 1964749161).



Copy editing: Gabriela Marszołek-Kalaga
Proofreading: Joanna Zwierzyńska
Typesetting: Tomasz Kiełkowski
Cover preparation for printing: Paulina Dubiel

Electronic version is the original one.
The journal was previously published in printed form with the
**ISSN 2450-5455**
The journal is distributed free of charge
**ISSN 2451-2125**

Published by
**Wydawnictwo Uniwersytetu Śląskiego**
**ul. Bankowa 12B, 40-007 Katowice**
www.wydawnictwo.us.edu.pl
e-mail: wydawnictwo@us.edu.pl
First impression. Printed sheets: 17,75. Publishing sheets: 21,0.


About this book
Free copy
ISSN 2451-2125
52
9 772451 212509