

INTERNATIONAL CONFERENCE OF WITHIN AND BEYOND THE CLASSROOM



ISTANBUL TECHNICAL UNIVERSITY
SCHOOL OF FOREIGN LANGUAGES

Book of
Proceedings

INTERNATIONAL CONFERENCE OF
WITHIN AND BEYOND THE CLASSROOM
BOOK OF PROCEEDINGS

Editors: Assoc. Prof. Nilüfer Ülker, Dr. Gürkan Temiz, Inst. Sinem Bayındır Musaoğlu,
Inst. Esra Meşe

Istanbul, 2024

e-ISBN 978-975-561-714-5

İTÜ Yayınevi. No: 2024.2/26

© ITU Press

All rights of this book are reserved and all publication rights belong to "ITU Press". All or any part of this book cannot be published, printed, filmed or used indirectly without the permission of the publisher. It cannot be reproduced by dublication, photocopy or any other technique, and cannot be transferred to an environment that can be processed on computers or typesetting machines. The responsibility of all texts and visuals published in the book belongs to the author(s).

International conference of within and beyond the classroom: book of proceedings/editors, Nilüfer Ülker, Gürkan Temiz, Sinem Bayındır Musaoğlu, Esra Meşe. -- Istanbul: ITU Press, 2024.

243 pages. -- (İTÜ Yayınevi. No: 2024.2/26)

ISBN 978-975-561-664-3

1. English language--Study and teaching-- Congresses

PE1068.T8 I56 2024

CIP

İTÜ YAYINEVİ

Sertifika No: 70051

İTÜ Ayazağa Kampüsü

Mustafa İnan Kütüphanesi

34469 Maslak İSTANBUL

0212 285 75 05

www.ituyayinevi.itu.edu.tr / ituyayinevi@itu.edu.tr



Published by Istanbul Technical University, School of Foreign Languages,
Istanbul, TÜRKİYE

Reproduction of content from this abstract book is permissible with publisher consent, provided (1) it is not for sale or profit, (2) the author is notified, and (3) proper credit is given to the "*Within and Beyond the Classroom Conference: Book of Proceedings.*"

The authors bear full responsibility for the content of their abstracts and guarantee their originality, absence of prior publication, and exclusive submission to this platform. The opinions expressed in the abstracts solely represent the individual authors' perspectives and may not align with those of the editor or reviewers.

The content of this publication is arranged based on the alphabetical order of the authors.

©2024 School of Foreign Languages, Istanbul Technical University

e-ISBN 978-975-561-714-5

INTERNATIONAL CONFERENCE OF WITHIN AND BEYOND THE CLASSROOM

BOOK OF PROCEEDINGS

Chief Editor

Assoc. Prof. Nilüfer ÜLKER

Istanbul Technical University

Co-Editors

Dr. Gürkan TEMİZ

Istanbul Technical University

Sinem Bayındır MUSAOĞLU

Istanbul Technical University

Esra MEŞE

Istanbul Technical University

Istanbul Technical University

School of Foreign Languages

Istanbul, Türkiye

May 31- June 1, 2024

Table of Contents

Conference poster		iii
Acknowledgments		iv
Committees		v
Conference Programme		vii
Proceedings		
Duygu D. Yalınkaya Emrah Görgülü	Gender Disparities in Teacher-Student Interactions	2
Eftelya Şahin	Assessing the Effectiveness of AI in Educational Assessment	19
Gürkan Temiz Elif Nazlı Kafadar Onur Kubilay Pak	AI and ELT: Pre-Service Language Teachers' Conceptions on AI and Its Implementation	38
Hacer Şivil	Implications of ChatGPT in Professional Life	51
H. Kübra Er Büşra Dağdemir	Teacher Feedback Literacy: EFL Writing Teachers' Beliefs in the Context of Higher Education	71
Maria Victoria Gálvez Frutos	Podele (Podcast Created by and for Spanish as a Second Language Students)	85
Melani Konca	Teachers' Beliefs and Practices of Vocabulary Teaching in EFL Classrooms	108
Meryem Büşra Ünsal Enisa Mede	Reflections of ESP Students on Online Asynchronous Discussions for an Advanced Speaking Skills Course	123
Natia Zviadadze Sopio Kipiani Tea Liklikadze	Blended Learning – Saving Class Time	140
Nida Gürbüz Emine Yılmaz	AI-Driven Assessment Feedback in English Language Teaching	156

Pınar Kır Gürkan Temiz	Exploring the Reflections of Turkish EFL Preparatory School Students towards “English” from an ELF Standpoint	177
Sinem B. Musaoğlu Emrah Görgülü	Interlanguage Pragmatics: The Effects of Explicit Teaching of Speech Acts on EFL Learners’ Pragmatic Competence	191
Yusuf Cengiz	Comparison of EAP Exam Texts with Undergraduate Readings: A Coh-Metrix Analysis	213

WITHIN AND BEYOND THE CLASSROOM

ISTANBUL TECHNICAL UNIVERSITY SCHOOL OF FOREIGN
LANGUAGES THIRD INTERNATIONAL ELT DAYS

The Future of Language Education: Adapting to a Changing Landscape

SAVE
THE
DATE



MAY 31- JUNE 1, 2024



ITU Mustafa Kemal Auditorium
School of Foreign Languages
Maçka Campus, Istanbul/Türkiye



itueltdays.itu.edu.tr

PLENARY SPEAKERS



Prof. Martin Dewey
King's College London



Prof. Gölge Seferoğlu
California State University



Prof. Derin Atay
Bahçeşehir University



**Assoc. Prof.
Anna Lia Proietti Ergün**
Yıldız Technical University

ii



İTÜ



SCHOOL of
FOREIGN
LANGUAGES

İTÜ



250th
YEAR
1973-2023

ACKNOWLEDGMENTS

As the organizers of the first international conference of ITU School of Foreign Languages (SFL), ITU ELT Days Within and Beyond the Classroom, we are thankful for the support and collaboration of all our colleagues. We would like to extend our heartfelt gratitude to **Prof. Martin Dewey**, **Prof. Derin Atay**, **Prof. Gölge Seferoğlu** and **Assoc. Prof. Anna Lia Proietti Ergün** for graciously accepting to be a part of our event and share their insights about the current issues in the field of foreign language teaching with us and our participants.

We would like to extend our special thanks to **Rector Prof. Dr. Hasan Mandal** for his support in this publication, and to our previous Rector **Prof. Dr. İsmail Koyuncu (2020-2024)** and Vice-Rector **Prof. Dr. Şule İtir Satoğlu (2020-2024)** for their contributions and guidance throughout this conference. We are also thankful to our Secretary General **Asst. Prof. Ümit Karadoğan** and our ITU Directorate of Communication for their continuous assistance.

With deep gratitude, we also recognize our scientific committee for their contributions in curating our conference. Their commitment and insights have been instrumental in elevating the quality of this gathering.

We appreciate the presenters for their time, effort, and dedication in delivering illuminating and inspiring presentations.

We would also like to extend our gratitude to the ITU SFL instructors for their cooperation and support in this process of preparations, and we would like to welcome all the participants whose interest and enthusiasm will make this conference a success.

Assoc. Prof. Nilüfer Ülker
On behalf of the Organizing Committee

Conference Chair

Assoc. Prof. Nilüfer ÜLKER

Director of School of Foreign Languages, Istanbul Technical University

Organizing Committee

Esra MEŞE

Istanbul Technical University

Esma KOT ARTUNÇ

Istanbul Technical University

Gürkan TEMİZ

Istanbul Technical University

Pınar KIR

Istanbul Technical University

Nalan ŞAN

Istanbul Technical University

Nergis KOPARAN

Istanbul Technical University

Sevda GÖNCÜ ERGÜN

Istanbul Technical University

Sinem BAYINDIR MUSAOĞLU

Istanbul Technical University

Özlem AYAR GEMALMAZ

Istanbul Technical University

Valentina ELMETTI

Istanbul Technical University

Bekir ŞERİFOĞLU

Istanbul Technical University

Scientific Committee

Prof. Ayşe Gürel
Boğaziçi University

Prof. Cristiano Corsini
Università Roma Tre

Prof. Deniz Melanlioğlu
İstanbul University

Prof. Julie Aydın
Social Sciences University of Ankara

Prof. Kenan Dikilitaş
the University of Stavanger

Prof. Leyla Cosan
Marmara University

Prof. Mustafa Durmuş
Hacettepe University

Prof. Nevin Ozkan
Ankara University

Assoc. Prof. Abdullah Nejat Töngür
Maltepe University

Assoc. Prof. Deniz Ortaçtepe
University of Glasgow

Assoc. Prof. Doğan Yüksel
The Open University

Assoc. Prof. Emrah GÖRGÜLÜ
İstanbul Sabahattin Zaim University

Prof. Enisa Mede
Bahçeşehir University

Assoc. Prof. Francesca Calamita
University of Virginia

Assoc. Prof. Mehmet Atay
Kocaeli University

Assoc. Prof. Mehmet Yalçın Yılmaz
İstanbul University

Assoc. Prof. Nur Yiğitoğlu Aptoula
Boğaziçi University

Assoc. Prof. Pınar Başkan
Marmara University

Assoc. Prof. Tarık Demir
Gazi University

Assoc. Prof. Tuncer Can
İstanbul Cerrahpaşa University

Dr. Dilek İnal
İstanbul Üniversitesi

Dr. Donald Staub
İzmir Economy University

Dr. Ece Genç Yöntem
Yeditepe University

Dr. Giuseppe Maugeri
Università di Urbino

Dr. Gül Deniz Demirel Aydemir
Hacettepe University

Dr. Merve Selçuk
Altınbaş University

Dr. Mesout Kalmı Salı
Çanakkale Onsekiz Mart University

Dr. Sevda Çimen
Zonguldak Bülent Ecevit University

Dr. Şeyma Kökcü
Tokat Gaziosmanpaşa University

Dr. Stefania Cavagnoli
Università di Roma 2 Tor Vergata

Dr. Yavuz Kurt
Marmara University

İTÜ SCHOOL OF FOREIGN LANGUAGES
INTERNATIONAL CONFERENCE OF WITHIN AND BEYOND THE CLASSROOM

DAY 1
May 31, 2024

TIME	SESSIONS			
08:30 – 09:30	Registration			
09:30 – 09:45	Welcoming Speech Istanbul Technical University			
09:45 – 10:00	Piano Recital			
10:00 – 10:45	Keynote Speaker I: Prof. Dr. Derin ATAY “Effective Professional Development in Higher Education: Complexities of Teacher and Student Learning”			
10:45 – 11:00	Coffee Break			
	Concurrent Sessions			
	Room I A101	Room II D103	Room III D104	Room IV D105
11:00 – 11:30	Arev Babahan Improving Vocabulary Acquisition Practices: Examining the Influence of Retention Strategies on Lower-Level English Learners via Action Research	Leila Jedi Sari Biglar & Aysegül Liman Kaban Exploring the Effect of Mobile-assisted Task-based Learning on Vocabulary Achievement and Student Attitude	Melani Konca Teachers' Beliefs and Practices of Vocabulary Teaching in EFL Classrooms	Yusuf Cengiz Comparison of an EAP Exam Texts with Undergraduate Readings: A Coh-Metrix Analysis
11:30 – 12:00	Jafar Moazzez The Effect of Lesson Study on Iranian EFL Intermediate Learners' Writing Accuracy	H. Kübra Er & Bişra Dağdemir Teacher Feedback Literacy: EFL Writing Teachers' Beliefs in the Context of Higher Education	Mehmet Ulus Teachers' Perspective on Challenges against Learner Autonomy in EFL Context	Zeynep Bilgehan Can Change through Virtual Exchange: An Autoethnography of a Virtual Exchange Instructor in Higher Education in Turkey

12:00 – 12:30	Özge Kuru & Abdullah Nejat Töngür Making Peace with Literature in ELF Classes: Integrating ‘Twitterature or Instapoetry’ in ELT	Alireza Sabzehparvar An In-depth Critical Examination of the Emergence and Development of Teacher Feedback Literacy in Language Education	Gizem Mutlu Gülbak & Ayşe Kızıldağ Teacher Resilience: What Differences CPDs can make?	Devrim Günay Utilizing Translation as Critical Reflexive Practice in Language Teacher Education: A Case for Untranslatability
12:30 – 13:15	LUNCH BREAK			
13:15 – 14:00	Keynote Speaker II: Prof. Dr. Martin DEWEY “Promoting Critical Engagement with Normativity and Perception Bias in Teacher Education”			
14:00 – 14:15	Coffee Break			
14:15 - 15:00	Workshop: Assist. Prof. Ufuk Keleş Autoethnography as the New Paradigm in Qualitative Research in ELT: The Whats, Hows, and Whys			
	Concurrent Sessions			
15:00 – 15:30	Arev Babahan Utilizing English as a Means to Elevate Students' Understanding of Istanbul's Culture	Esra Meşe & Ufuk Keleş Collaborative Autoethnography as a Transformative Teacher Education Tool and a Means to Voice to 'mystory'	Özge Kuru A Critical Examination: Integrativeness in SLA within the Global Englishes Paradigm	Steve Brown & Hélène Delpont Engaging with Society around Climate and Biodiversity
15:30– 16:00	Nalan Şan & Derin Atay Language Self-efficacy Beliefs and Technology-based Self-regulated English Learning Strategies of Second Language Learners in a Rapidly Digitalizing World	Pınar Kır & Gürkan Temiz Exploring the Reflections of Turkish EFL Preparatory School Students towards “English” from an ELF Standpoint	Ildiko Furka English for Academic Purposes, Industry 4.0 and the VUCA World: A Case for the Transferability of Skills from Teaching English as a Foreign Language	Hacer Şivil Implications of ChatGPT in Professional Life
16:00 – 16:30	Duygu Doğruçan Gender Disparities in Teacher-Student Interactions	Sinem Bayındır Musaoğlu & Emrah Görgülü Interlanguage Pragmatics: The Effects of Explicit Teaching of Speech Acts on EFL Learners’ Pragmatic Competence	Bora Arga & İnci Melike Tokcan A Comparison of the Perceptions of Pre-service and In-service Teachers on Incorporating Authentic Materials in EFL Classrooms	Seçil Cengiz Benefits and Sustainability of Teacher Research as a Professional Developmental Tool

DAY 2
June 1, 2024

SESSIONS				
Registration				
Keynote Speaker III: Prof. Dr. Gölge SEFEROĞLU				
“Leveraging Unfolding Innovations and AI to Enhance Language Learning and Teaching”				
Coffee Break				
Concurrent Sessions				
	Room I A101	Room II D204	Room III D104	Room III D105
11:00 – 11:30	Eftelya Şahin Assessing the Effectiveness of AI in Educational Assessment	Nida Gürbüz & Emine Yılmaz AI-Driven Assessment and Feedback in English Language Teaching	Gökçe Gök Unlocking the Power of AI in Education: Reflections from an Online Course	Gizem Arslan Değirmenci Unraveling the Power of Emotions: How Emotional Intelligence Shapes Success in Learning English as a Foreign Language among University Students
11:30 – 12:00	Buse Uzuner Instructors Matter: Investigating The Relationship Between Grit, Well-being and Self-efficacy of EFL Instructors	Meryem Büşra Ünsal & Enisa Mede Reflections of ESP Students on Online Asynchronous Discussions for an Advanced Speaking Skills Course	Gürkan Temiz , Nazlı Kafadar & Onur Pak AI and ELT: Pre-service Language Teachers' Conceptions on AI and its' Implementation	Ekaterine Kurdadze, Nino Demetradze & Ekaterine Archvadze Developmental Approach to Peer Observation
12:00 – 12:30	Melike Ubuz & Enisa Mede The Relationship among Foreign Language Classroom Anxiety, Willingness to Communicate and Personality Traits of EFL Learners	Zeynep Oğul & Sena Elibal İçüz Assessment of Learning or Assessment For Learning	Natlia Zviadadze, Sopio Kipiani & Tea Liklikadze Blended Learning – Saving Class Time	Elif Aydın Yazıcı & Kenan Dikilitaş Beyond Language Barriers: A Metasynthesis of Translanguaging-Focused Teacher Training and its Impacts on Teachers

12:30 – 13:15	LUNCH BREAK			
13:15 – 14:00	Keynote Speaker IV: Assoc. Prof. Anna Lia Proietti ERGÜN Finding Joy in Language Learning: The Positive Psychology Perspective			
14:00 – 14:15	Coffee Break			
14:15 - 15:00	Workshop II: Assoc. Prof. Francesca Calamita Feminist Pedagogy, Women's Studies, and SLA: Interdisciplinary Encounters Towards Gender Equality			
	Concurrent Sessions			
15:00 – 15:30	Maria Victoria Galvez Frutos PODELE (PODCAST Por Y Para Estudiantes De Español Como Segunda Lengua) PODELE (PODCAST created by and for Spanish as a Second Language Students)	Tuba Arabacı What is at Stake? Corpus in English for Academic Purposes	Mert Çetinok Exploring Being an English Language Teacher in Turkey: A Model Investigating the Effects of Teacher Autonomy, Personality Traits, and Foreign Language Teaching Anxiety on Teacher Job Satisfaction	Gülçin Yonca Akdeniz & Gökhan Öztürk A Mixed-Methods Study on Foreign Language Teaching Anxiety among Turkish EFL Teachers
15:30 – 16:00	Nimet Çopur Producing a Jocular Frame in Response to Student Turns Violating Participation Framework in L2 Classrooms	Valentina Elmetti & Esra Meşe The Interplay of Enjoyment and Anxiety in Learning Foreign Languages other than English	Ghada Badawi & Amani Hassan How Could Students' Grammar Errors Inform Instruction in Our Integrated Skill Class?	
16:15 – 16:30	Closing Remarks & RAFFLE			

PROCEEDINGS

Gender Disparities in Teacher-Student Interactions

Duygu Doğrucan Yalınkaya¹ Emrah Görgülü²

¹Inst., Istanbul Technical University, Türkiye

²Assoc. Prof., Istanbul Sabahattin Zaim University, Türkiye

Abstract

Classroom interactions have a significant role to continue gender roles through teacher-student or student-student dynamics. The fundamental principle of equality dictates that every student should be given fair treatment and provided with equal access to educational opportunities. The same amount and quality of attention and care are needed to be given to the students to balance their educational development regardless of gender. However, as a result of comprehensive research, it is seen that gender neutrality is not present in teachers' interaction patterns. Additionally, the way interactions are distributed changes based on the gender of the teacher. The purpose of this study is to investigate how students of both genders are dealt with by teachers with different genders in an English preparatory school context. Classes that were led by two female and two male teachers were observed, and the relevant parts were videotaped. Subsequently, the videos were transcribed and analyzed by utilizing an adapted version of Sinclair and Coulthard's Classroom Analysis Model (1992). A statistical analysis, employing the independent t-test in Statistical Package for the Social Sciences V25.0 (SPSS V25.0), was then used to find out the statistical significance of the observed disparities. Qualitative analysis revealed an unequal distribution of teacher-student interaction, with a tendency towards providing more attention to male students. However, the t-test results did not show statistically significant differences. The findings of the study are discussed in the context of existing literature on gender and classroom interaction, and pedagogical implications are suggested to improve the learning experiences of all students, regardless of gender. Additionally, the study underlines the need for further research to elucidate the underlying reasons for observed gender disparities in teacher-student interactions, thereby contributing to a deeper understanding of gender-related dynamics in an educational context.

Keywords: Classroom interaction; gender and classroom interaction, teacher-student interaction, gender inequalities in classroom, gender bias

1. Introduction

This section provides an overview of the present study. It outlines its background, main purposes, significance, research questions, scope, and limitations. Then, it includes a brief overview of current research with the structure of the thesis, and definitions of key terms.

1.1 Statement of the Problem

Schools play a pivotal role in preparing students for adulthood and future careers. The interactions within schools have a big effect on broader societal issues such as gender inequalities. Gender bias in the classroom, particularly in course materials and teacher-student interactions, is a critical issue. According to a study conducted by UNICEF in Türkiye, textbooks still have elements promoting traditional gender roles. Men, for instance, are depicted in public roles while women are limited to domestic responsibilities (UNICEF, 2003). Besides, teachers hold an authority in the classroom and they impact learners deeply. Therefore, the interactions that take place in schools have an impact on students in terms of shaping their future interactions with others. Research indicates that boys often receive more attention and higher quality interactions from teachers compared to girls (Becker, 1981; Baker, 1986; Sadker & Sadker, 1986; Jones, 1989). However, there is only one study on this issue within the Turkish context (Bağ et al., 2014). Therefore, this study aims to address this gap.

The Turkish National Education Law promotes equal opportunities for girls and boys (Milli Eğitim Temel Kanunu; madde: 8). However, research shows that there is an inequality in classrooms and teaching materials. Sunderland (2010) exemplifies issues such as the female characters' not being visible, gender role stereotypes, and discourse in English Language Teaching (ELT). These inequalities and one-sidedness have a big influence on female learners' cognitive and communicative development. What is worse, it might cause female learners to feel alienated, and limit their classroom interactions. This study investigates whether boys and girls receive same amount of attention from teachers and delves into gender-biased interactions in ELT. It aims to raise awareness and to foster equal and not biased educational interactions.

1.2 Research Questions

The study addresses the following research questions in the EFL classroom:

1. Is there any difference between the number of the acts of female and the male teachers directed to male and/or female students?
 - a. Is there any difference between the number of the initiating acts of female and the male teachers directed to male and/or female students?
 - b. Is there any difference between the number of follow-up acts that are directed to the female and/or male students?
2. Is there any difference in the number of acts that female and male students are provided?
 - a. Is there any difference in the number of initiating acts that female and the male students are provided?
 - b. Is there any difference between the number of follow-up acts that female and male students are provided?

2. The Study

2.1 Literature Review

In many societies, values and conventions build the personal traits and behaviors of both men and women. Before starting to examine the topic deeply, it is necessary to define *sex* and *gender*. Sex refers to the biological disparity of being male or female, while gender alludes to the socially influenced traits of each sex. West and Zimmerman (1987) and Butler (1990) argue that gender is not an inherent trait but a practice, which means it is something individuals perform and act not something they biologically possess. Gender is rooted in biological sex but it overdraws these disparities and adapts them to settings in which they are not relevant. Examples include women wearing pink and men wearing blue without any biological reason, or women's crying is being highly acceptable while men's crying is not. Gender is more associated with nurture, whereas sex with nature. Nevertheless, nature and nurture are closely connected, and as Eckert and McConnell-Ginet (2016) state, there is no clear-cut separation between the two concepts. Eckert and McConnell-Ginet (2016) escalates Simone de Beauvoir's idea that gender is not inherent but constructed. Immediately from birth, children are treated according to gender norms, which influence how they behave and what their interests are. By adulthood, gendered divisions of labor become clear, with women often relegated to domestic roles and men given public, creative spaces.

Gender favoritism continues in academia even though there is the presence of many female academicians and students in it. Özyaydınlık (2014) notes that only 30% of students in technical departments and only 5.6% of rectors are female. Gender inequalities extend to access to resources and decision-making procedures, sustained through tools like language and education. The gendered process has led to increasing social inequalities and limited women's access to social resources worldwide. It has confined women to limited roles and has led to gender pay gaps for the same work. This inequality has provoked global backlash, including in Türkiye. Feminist movements have challenged patriarchal ideologies that privilege men over women. That is why it led to the rise of gender studies.

2.1.1 Gender and Education

Education has an important impact on societal norms and gender roles. Atatürk stated that education is worthy in terms of societal benefits (Uysal, 1981). Schools shape society by embedding ideological growth influencing language and power dynamics (Delamont, 1980). Parham (2013) emphasized that gender inequality comes out of prolonged educational practices. Despite aims for gender equality, education often reinforces stereotypes and biases. Teachers have a crucial role to ensure gender equality among students in the classrooms. Appleby (2009) found that teachers can impact gender equity. Besides, Uysal (1981) emphasized how textbooks impact improvement of students; however, many studies put forward that EFL (English as a Foreign Language) textbooks often perpetuate gender biases.

2.1.2 Gender and EFL

Gender issues in ELT (English Language Teaching) include pronoun usage and textbook content. Generic "he" is being replaced by "him or her" and "s/he." Efforts to desex vocabulary are ongoing but sexist terms persist. Thanks to studies on gender representation in EFL

textbooks, it has been revealed that males are often more prominent and depicted in more active and leading roles. Bahman and Rahimi (2010) stated there are significant gender gaps in Iranian EFL textbooks. Hartman and Judd (2008) noted that male references often dwarf female ones, with women typically portrayed in domestic roles. Porreca (1984) found women did not take enough place in EFL textbooks, and they were depicted with limited occupational roles, which has a negative impact on the empowerment of female learners.

2.1.3 Gender and Classroom Interactions

Interaction is crucial in understanding ourselves and the world. Social interactions let us form and challenge our identities, by emphasizing the significance of an unbiased classroom atmosphere. Research by Duffy et al. (2001), Farooq (2000), and others have examined the influence of gender on performance, classroom dynamics, and academic achievement. Cornill (2007) noted the interest in gender and classroom interactions began in the 1950s. Modern expectations are that all students, regardless of gender, should be given equal attention in schools to help them get ready for their futures.

In EFL classrooms, in which students often depend severely on teacher interaction because of language barriers, Farooq (2000) highlighted that the primary interaction for learners at a basic level is through teachers. Brown (1994) and Caudron (1993) emphasized the significance of teacher questions and feedback in constructing language competence. Nevertheless, studies like Bailey (1993) and Omvig (1989) noted that teachers interact differently with different gendered students. Brophy (1985) stated that male students generally receive more attention from teachers than female students (Hopf & Hatzichrisou, 1999; Bellamy, 1994).

Many studies suggest that male students have a dominance in classroom interactions. They receive more feedback, attention, and opportunities to speak (Butters, 1986; Alcon, 1994; Sunderland, 2000). Teachers have a tendency to ask more questions and provide more feedback to boys. They mostly perceive girls as more hard-working and self-sufficient (Farooq, 2000). Stake and Katz (1982) and Younger and Warrington (1996) noted that boys receive more warnings and criticism, probably because of more frequent disciplinary issues. Aukrust (2008) found that boys generally participate more than girls and interrupt teachers more often.

In Türkiye, Bağ et al. (2014) examined gender and teacher-student interaction in EFL context, finding uneven dissemination of teacher attention. This study is significant as it is among the first to explore this issue in Türkiye's EFL setting by utilizing video-taped classroom interactions.

Next section will detail the research methodology used in this study.

2.2 Methodology

This study is designed as quantitative research, which aims to create numerical data or data transformable into usable statistics.

2.2.1 Research Design Research Participants and Setting

The study aims to identify gender bias in classrooms by examining if boys and girls receive similar amounts and types of attention from teachers, and if gender of the teacher influences this. Conducted over two weeks, it involved observational research where participants were unaware of the study's aim. They were videotaped for 4 hours in total and 20 minutes of each were analyzed. The parts that were analyzed are the ones that had lots of teacher-student interaction.

The study was conducted at a private foundation university's preparatory school, focusing on teacher-student interactions based on gender. The university's School of Foreign Languages offers an extensive English preparatory program with levels A1 to B2. To ensure representative results, two B1+ and two B1 classes were selected, taught by two male and two female teachers. Only classes with at least B1 level students were chosen to ensure sufficient proficiency for communication. Teachers' age ranges 25 to 30 while their teaching experience ranges 3 to 6. They all graduated from the top two universities in Türkiye. A smartphone camera was the main data collection tool. The researcher used a frequency chart to document the different kinds of feedback each teacher gave, distinguishing between interactions with male and female students.

2.2.2 Procedure

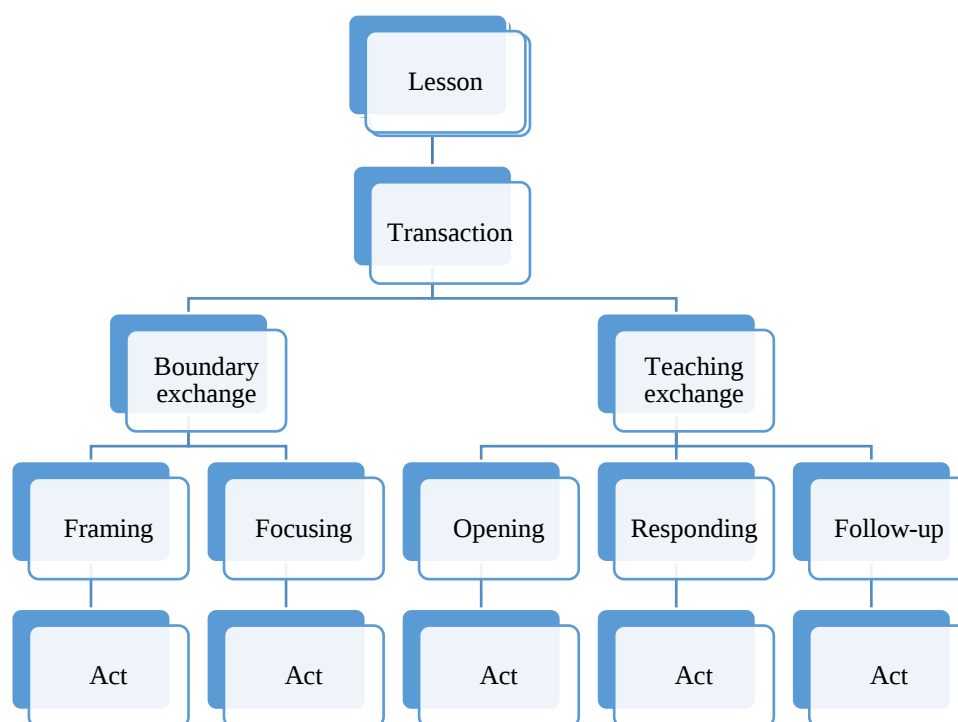
After recording the lessons, the researcher transcribed all speech using the online application, Trint. The frequency of Initiation Response Feedback (IRF) chart was used to document teacher-student interactions. Interactions with male students were noted in black, while those with female students were noted in blue. The data was then analyzed to see if the teachers gave more attention to either male or female students, as well as what kind of attention was given to female and/or male students by female and /or male teachers. An independent t-test using SPSS (V25) was used to see if the results were statistically significant.

2.2.3 Data Analysis Procedure

2.2.3.1 Sinclair and Coulthart's Classroom Analysis Model

The data analysis made use of Sinclair and Coulthart's Classroom Discourse Analysis Initiation-Response-Feedback (IRF) Model (1992), specifically developed to describe teacher-pupil interaction. The researcher chose to use categories from a model developed version by Francis and Hunston (1992) because of the presence of non-traditional and communicative style in which the class was conducted. Sinclair and Coulthart's model includes five ranks: lesson, transaction, exchange, move, and act.

Figure 1: Sinclair and Coulthard's Model



Source: Sinclair and Coulthard, 1992

Figure 1 demonstrates the basic interaction model in a teacher-based instruction. In Sinclair and Coulthard's model (referred to as S&C), the highest rank is the lesson. However, they had difficulty in demonstrating how a lesson is made up of transactions (Coulthard, 1985: 123). The other four levels, however, have well defined structures. Transactions are made up of boundary and teaching exchanges. Sinclair and Coulthard describe boundary exchanges as indicators for the beginning or end of a lesson, transaction, or a change in topic, often using words like "alright," "now," or "OK," with rising intonation and a brief pause. Exchanges are composed of moves such as focusing, framing, opening, responding, and following-up. Finally, moves are made up of "one or more acts" (ibid: 124-125).

For this study, only the exchange, move, and act levels were analyzed. Teaching exchanges consist of an opening move (Initiation), an answering move (Response), and a follow-up move (Feedback). The IRF model tracks how teachers initiate interaction (I), students respond (R), and teachers provide feedback (F).

The following is an example of a typical IRF exchange:

Initiation (I): Teacher: "What is this?"

Response (R): Student: "It's a table."

Feedback (F): Teacher: "Good."

Exchanges serve four main purposes: informing, directing, eliciting, and checking.

In the S&C model, acts are the lowest rank, forming the basis of moves such as framing, focusing, opening, responding, and following-up. Sinclair and Coulthard (1992: 19-219) identified twenty-two acts in their model, which Francis and Hunston (1992: 128-133) later adapted (see Table 1).

Table 1: Acts of Sinclair and Coulthard

Code	Act	Function	Realization
gr	Greeting	Self-explanatory	hi', 'hello' 'good morning'
re-gr	Reply-greeting	Self-explanatory	hi', 'hello' 'good morning'
inq	Inquire	Elicits information	wh-questions and ellipted forms
n.pr	Neutral proposal	Elicit information	yes no questions
m.pr	Marked Proposal	Elicits agreement	Questions seeking a 'yes' or 'no' answer, 'don't/aren't you'
Ret	Return	Seeks clarification of previous utterance	Questions, often ellipted
conf	Confirm	Gives or asserts agreement	'Yes' or 'no' items including NV, repetition and paraphrase
rej	Reject	Rejects suggestions or Presuppositions	'Yes' or 'no' items including NV or silence.
ter	Terminate	Terminates an exchange while acknowledging previous utterance	Low key 'yes' or 'no' items and their variants, or repetition
rec	Receive	Acknowledges previous utterance	'Yes' or 'no' items including NV and repetition
prot	Protest	Objects to preceding utterance, disputing its correctness	'Yes' or 'no' items and their variants
acc	Accept	Shows T has heard correct information	Yes', 'Good', 'Fine'
ack	Acknowledge	Shows initiation (or response) has been understood	'Yes', 'Ok', 'mhmm', 'Wow'
eng	engage	Providing minimal feedback without interrupting	'mm', 'yeah' and low echoes
ref	reformulate	To acknowledge preceding utterance and/or offer a revised version of it	Statement which paraphrases preceding utterance
z	Aside	T talking to himself/herself	Statement/question/command
cl	Clue	Gives extra information	Statement/question/command
com	Comment	Exemplifies/expands/justifies	Statement/tag question
con	Conclusion	Summarizes	'So, what we've been doing is...'
el	Elicitation	Requests answer	Question
e	Evaluate	Evaluates	'Good', 'Interesting', 'Yes'
l	Informative	Provides information	Statement

Code	Act	Function	Realization
I	Loop	Returns to point before S's answer	'Pardon?', 'Again?'
me	Marker	Marks boundary in discourse	Well', 'OK', 'Right'
rep	Reply	Provides appropriate reply to elicitation	Statement/question/nod
^	Silent Stress	Highlights maker	Pause
S	Starter	Provides information to facilitate response	Statement/question/command

Source: Sinclair and Coulthard, 1992

The following example and explanation further of Heinel (2017) simplify how exchanges, moves and acts interact in discourse:

- (I) Teacher: Where are you from (el)
 (R) Student: Eugene (rep)
 (F) Teacher: Eugene ok (e) Oregon State then (com)

In this exchange, the teacher makes an initiating (I) move which consists of an elicitation (el) act. This elicitation is the only act in the move and it can stand on its own, therefore it is the 'head' and indicates that the exchange is a Teacher Elicit type of exchange. In the response (R) move the student makes a reply (rep). In the third and final follow-up (F) move there are two acts, the teacher first evaluates (e) the student response, and then comments (com) on it. (Heinel, 2017)

2.2.3.2 Application of the Model in Data Analysis

To answer the research questions, the S&C model was utilized at the exchange and act levels. Not all acts were used in the analysis, only those frequently used and relevant to the study. The researcher coded the transcribed speech, focusing on initiating acts (inquire, neutral proposal) and follow-up acts (comment, reformulate, acknowledge, engage).

The transcribed data of Teacher 1 is as below:

Teacher 1: [00:00:02] Yeah. Number five. Who is number five. Cüneyt. Wow. How nice Cüneyt. Yeap. Cüneyt can you tell me what's your department.

Student A: [00:00:10] My department is Islamic economics and Finance.

Teacher 1: [00:00:13] Finance. Okay. What, what do you think about the advantages of your department.

Student A: [00:00:16] My department is new department. Maybe I can find a job easily.

Teacher 1: [00:00:21] Easily yeah.

Student A: [00:00:23] But disadvantages. Math sometimes can be hard.

The coded part of this dialog is as in Table 2:

Table 2: Sample coded chart

Initiation	Act	Response	Feedback	Act
Yeah. Number five. Who is number five? Cüneyt. Wow. How nice Cüneyt. Yeap. Cüneyt can you tell me? what's your department.	inq	My department is Islamic economics and Finance.	Finance. Okay..	ack
What, What do you think about the advantages of your department	inq	My department is new department. Maybe I can find a job easily.	Easily yeah.	Eng
		But disadvantages. Math sometimes can be hard.	Math. Hmm. For the studying part you think math would be maybe difficult. Thank you.	ack ref

Black coloring has been used as the student who is speaking is male. For the female students blue colored writing has been preferred.

2.3 Results and Discussion

2.3.1 Results and Findings

This chapter illustrates the study's findings through tables, focusing on teacher-student interactions. The study investigates the frequency of different interaction types (initiation and follow-up acts) by teachers towards male and female students. Sinclair and Coulthards' discourse analysis model was utilized.

As it can be seen in sample table (Table 3), Teacher #1 used 134 acts in total, which were directed towards 5 female and 11 male students. She used 7 inquire acts (inq) for female students and 28 for male students; 3 neutral proposal acts (n.pr) for females and 5 for males; 10 acknowledgment acts (ack) for females and 22 for males; 12 engage acts (eng) for females and 19 for males; 3 comment acts (com) for females and 18 for males; and 1 reformulate act (ref) for females and 6 for males (Table 3).

To account for the differing numbers of female and male students, the number of acts per student was calculated (Table 4). Teacher #1 used more inq, ref, and com acts towards male students, while using more n.pr and eng acts towards female students. The usage of ack was the same for both genders.

Teacher #1

Table 3: A sample table -Female Teacher #1- Total distribution of acts

	Female student (N=5)	Male students (N=11)
inq	7	28
n.pr	3	5
ack	10	22
eng	12	19
com	3	18
ref	1	6
TOTAL (N=134)		

Table 4: Sample Table -Teacher #1-Acts provided per female and male students

	Female Students	Male students
Inq	1,4	2,545
n.pr	0,6	0,454
Ack	2	2
Eng	2,4	1,727
Com	0,6	1,636
Ref	0,2	0,545

Teacher #2

Teacher #2 used 172 acts in total towards 8 female and 9 male students. The distribution is as follows: 22 inq acts for females and 34 for males; 11 n.pr acts for females and 14 for males; 20 ack acts for females and 33 for males; 5 eng acts for females and 0 for males; 4 com acts for females and 11 for males; and 8 ref acts for females and 10 for males.

The mean number of acts per student reveals that Teacher #2 used more inq, n.pr, ack, and ref acts towards male students, while using more eng and com acts towards female students.

Teacher #3

Teacher #3 used 133 acts in total towards 8 female and 9 male students. The distribution is: 12 inq acts for females and 13 for males; 7 n.pr acts for females and 4 for males; 28 ack acts for

females and 26 for males; 7 eng acts for females and 9 for males; 13 com acts for females and 6 for males; and 5 ref acts for females and 3 for males.

The mean number of acts per student indicates that Teacher #3 directed more inq and eng acts towards male students, and more n.pr, ack, com, and ref acts towards female students

Teacher #4

Teacher #4 used 151 acts in total towards 12 female and 14 male students. The breakdown includes: 15 inq acts for females and 16 for males; 6 n.pr acts for females and 5 for males; 25 ack acts for females and 31 for males; 10 eng acts for females and 9 for males; 12 com acts for females and 10 for males; and 4 ref acts for females and 8 for males.

The mean number of acts per student shows that Teacher #4 used more ack and ref acts towards male students, while using more inq, n.pr, eng, and com acts towards female students.

2.3.1.1 Gender-Related Findings

Female and male teachers' usage of acts towards female and male students has been analyzed individually. However, analyzing the data from a gender perspective is a necessity as well.

1. Male teachers directed more acts to students overall than female teachers.
2. Male teachers directed more acts to female students than female teachers did.
3. Female teachers directed more acts to male students than male teachers did.
4. Female teachers directed more acts to male students than to female students overall.
5. Female teachers used more initiation acts towards male students.
6. Female teachers used more follow-up acts towards male students.
7. Female teachers directed more inq, n.pr, ack, com, and ref acts towards male students, but more eng acts towards female students.
8. Male teachers directed more acts to female students than to male students overall.
9. Male teachers used more initiating acts and follow-up acts towards female students.
10. Male teachers directed more n.pr, ack, com, and ref acts towards female students, but more inq and eng acts towards male students.
11. Male teachers directed more follow up acts to female students than to male students.
12. Male teachers directed more n.pr, ack, com and ref to their female students than to their male students.
13. Male teachers directed more inq, eng to their male students than to their female students.

When we analyze the table (Table 5), we can conclude that:

From the students' perspective, male students received more acts overall, including more initiating and follow-up acts, and specifically more inq, ack, com, and ref acts. Female students received more n.pr and eng acts.

An independent samples t-test has been used to analyze the findings and it showed no significant differences in the number of acts used by female and male teachers towards female and male students nor in the number of acts received by female and male students.

Table 5: The means of act usage of teachers

	Female Teachers			Male Teachers		
	Per students	Per female students	Per male students	Per students	Per female students	Per male students
inq	2,7005	2,0750	3,1610	1,6375	1,6250	1,6540
n.pr	0,9850	0,9875	1,0045	0,6695	,8330	,5115
initiation act	1,84275	1,53125	2,08275	1,1535	1,229	1,08275
ack	2,5585	2,2500	2,8330	3,3265	3,3745	3,2735
eng	1,1155	1,5125	,8635	1,0315	,9995	1,0710
com	0,0970	,5500	1,4290	1,2145	1,5830	,8570
ref	0,7475	,6000	,8280	0,5635	,5830	,5355
follow-up act	1,1296	1,2281	1,4884	1,534	1,635	1,43425
Total mean	1,3673	1,3292	1,6865	1,4072	1,4997	1,3171

2.3.2 Discussion

This section discusses the research questions related to relevant literature

1. Is there any difference between the number of the female and the male teachers' acts directed to male and/or female students in the EFL classroom?

a. Is there any difference between the numbers of the female and the male teachers' initiating acts directed to male and/or female students in the EFL classroom?

b. Is there any difference between the numbers of the female and the male teachers' follow up acts directed to the male and/or female students in the EFL classroom?

2.3.2.1 Findings Related to Teachers' Gender

Female teachers directed:

- more acts to their male students than to their female student in total.
- more initiation acts to male students than to female students.
- more follow up acts to male students than to female students.
- more *inq*, *n.pr*, *ack*, *com* and *ref* to their male students than to their female students.
- *eng* to their female students than to their male students.

Male teachers directed:

- more acts to their female students than to their male students in total.
- more *n.pr*, *ack*, *com* and *ref* to their female students than to their male students.

- more *inq*, and *eng* to their male students than to their female students.
- more initiating acts to female students than to male students.
- more follow up acts to female students than to male students.

However, when the difference of number of acts that female and male teachers gave is analyzed by SPSS it has been found that there is no statistical difference.

When the findings of the study are compared to the relevant literature, they align with studies by Farooq (2000), Sunderland (1996), Bailey (1993), and Omvig (1989), which reported teachers' greater attention to boys and their interacting with different gendered students in different ways. However, Bellamy (1994) noted male teachers interacted more with male students, and Kelly (1988) and Jones and Dindia (2004) both female and male teachers interacted more with male students than they did with female ones. Brophy (1985) suggested male teachers directed their attention less to female students than did female teachers, which showed no consistency with this study. Yopez (1994) stated that teachers showed remarkable egalitarianism in their interaction with different gendered students, aligning with this study's statistical findings only.

2. Is there any difference in the number of acts directed to female and the male students in the EFL classroom?

- Is there any difference in the number of initiating acts directed to female and the male students in the EFL classroom?
- Is there any difference between the number of follow up acts provided to female and male students in the EFL classroom?
- Male students were provided more acts than female students in total.
- Male students were provided more initiating acts than female students were.
- Male students were provided more follow-up acts than female students were.
- Male students were provided more *inq*, *ack*, *com* and *ref* acts than female students.
- Female students were provided more *n.pr* and *eng* than male students.

These findings are consistent with Munro (1987), Merrett and Wherdall (1992), and Younger, M., & Warrington, M., & Williams, J. (1999), who reported male students receiving more questions and responses. They also align with Brophy (1985), Kelly (1988), Dindia (2004), and Spaulding (1963) regarding greater interaction with male students. Sadker & Sadker (1986) found male students received more praise and criticism, which this study partially supports. Finally, the study echoes Bağ et al. (2014), which is the first study conducted to find out the nature of student-teacher interaction related to gender issues in Türkiye. The results of the related study were concluded that:

In Female Teacher's class:

"The mean of the initiating moves (A&NA) directed to male students was higher than the mean of the initiating moves directed to female students. There was a fairly well balanced distribution of attention when directing NA and follow-up moves to the students, only slightly favouring females with her NA initiating moves and males with her follow-up moves" (Bağ et al.2014: 74).

In Male Teachers class:

“The mean of the initiating moves (A&NA) directed to female students was higher than the mean of the initiating moves directed to male students. The mean of the follow-up moves provided to female students was higher than the mean of the follow-up moves provided to male students” (Ibid: 74).

The findings of this study show an alignment with Bağ et al. (2014), which revealed that female teachers provided more initiating acts and follow-up moves to male students, while male teachers directed more initiating and follow-up moves to female students. Both studies are consistent in indicating similar patterns in teacher-student interactions based on gender in Turkish EFL classrooms.

3. Conclusion

This research aimed to contribute to the field of ELT by examining gender dynamics in EFL classrooms. It particularly analyzed how female and male teachers directed their initiating and follow-up acts to male and female students differently. Four teachers were videotaped for 20 minutes each, and the videos were transcribed, and analyzed using an adapted Sinclair and Coulthard model, and then with SPSS. The study found that female teachers provided more initiation and follow-up acts to male students, while male teachers directed more to female students. However, these differences were not statistically significant. Male students received more attention overall, aligning with previous studies like Bağ et al. (2014).

Despite the lack of statistical significance, it is clear that there is a tendency to give more attention to male students. There is not any enough study related to the issue in Türkiye’s ELT setting. Thus, similar studies should be conducted to explore these dynamics further. Gender of the teacher might have an impact on interaction patterns, and there is a need for investigation into the reasons behind these disparities. Teachers should reflect on their interaction patterns and avoid gender biases. To achieve this, they are suggested to use self-observation and role-playing activities to promote gender equality. In addition, in-service training, seminars, and workshops can help teachers broaden their understanding of gender issues and create a more inclusive learning environment.

References

- Alcón, E. (1994). The role of participation and gender in non-native speakers' classroom interaction. *Working Papers on Language, Gender and Sexism*, 4 (1): 51–68.
- Appleby, R. (2009). The spatial politics of gender in EAP classroom practice. *Journal of English for Academic Purposes*, 8(2):100-110. doi:10.1016/j.jeap.2008.09.004
- Aukrust, V.G. (2008). Boys' and girls' conversational participation across four grade levels in Norwegian classrooms: Taking the floor or being given the floor? *Gender and Education*, 20 (3): 237–252.
- Bağ E. , Martı L. , Bayyurt Y. (2014) Gender and Classroom Interaction: Examining A Female and A Male Teacher's Moves Directed Towards Female and Male Students in Two EFL Classrooms in Turkey, *Dergipark*, 31 (1): 59-79.
- Bahman, M., & Rahimi, A. (2010). Gender representation in EFL materials: An analysis of English textbooks of Iranian high schools. *Procedia - Social and Behavioral Sciences*, 9: 273-277. doi:10.1016/j.sbspro.2010.12.149
- Bailey, S. M. (1993). The current status of gender equity research in American schools. *Educational Psychologist*, 28: 321–339.
- Baker, D. (1986). Sex differences in classroom interactions in secondary science. *Journal of Classroom Interaction*, 22: 212-218.
- Batters, Julia. Boys Really Think Languages Are Just Girl-Talk?. *Modern Languages*, 67:2 (1986): 75-79.
- Becker, J.R. (1981). Differential treatment of females and males in mathematics classes. *Journal for Research in Mathematics Education*, 12: 40-53.
- Bellamy, N. (1994). Bias in the classroom: Are we guilty? *Science Scope*, 17: 60–63.
- Brophy, J. (1985). Teacher praise: A functional analysis. *Review of Educational Research*, 51: 5-32.
- Brown, H. 1994. *Teaching By Principles*. Englewood Cliffs: Prentice Hall.
- Butler, Judith. 1990. *Gender Trouble: Feminism and the Subversion of Identity*. New York and London: Routledge.
- Chaudron, C. 1993. *Second Language Classrooms: Research on teaching and Learning: Cambridge*: Cambridge University Press.

- Cornill, D.K. (2007). *Gender studies and foreign language teaching*. Tübingen: Narr
- Coulthard, M. (1985). *An Introduction to Discourse Analysis*. Harlow: Longman.
- Duffy, J., Warren, K., & Walsh, M. (2001). Classroom interactions: Gender of teacher, gender of student, and classroom subject. *Sex Roles*, 45 (9/10): 579 –593.
- Eckert, P, & McConnell-Ginet, S. (2016). *Language and gender*. Cambridge: Cambridge University Press.
- Farooq, M.U. (2000). Examining a male teacher's attention in a mixed-sex EFL Japanese high school classroom based on the Sinclair-Coulthard model. Unpublished MA thesis, University of Birmingham.
- Francis, G. & Hunston, S. (1992). Analysing everyday conversation. In Coulthard, M., *Advances In Spoken Discourse Analysis* (123-161). London: Routledge.
- Hartman, P. L. & Judd, E. L.(1978). *Sexism and TESOL Materials*. TESOL Quarterly, 12(4):383-393.
- Heinel, H. (2017). The Application of Sinclair and Coulthard's Model for Analyzing Classroom Discourse to a Non-Traditional Japanese Conversation Class.
- Hopf, D., & Hatzichristou, C. (1999). Teacher gender-related influences in Greek schools. *Educational Psychology*, 69: 1–18.
- Jones, M.G. (1989). Gender differences in teacher-students interactions. *Journal of Teacher Education*, 49: 33-38.
- Jones, S., & Dindia, K. (2004). A meta-analytic perspective on sex equity in the classroom. *Review of Educational Research*, 74 (4): 443–471.
- Kelly, A. (1988). Gender differences in teacher-pupil interactions: A meta-analytic review. *Research in Education*, 39: 1–23.
- Merrett, F., & Wheldall, K. (1992). Teachers' use of praise and reprimands to boys and girls. *Educational Review*, 44(1): 73–79.
- Munro, Fran. "Female and Male Participation in Small-Group Interaction in the ESOL Classroom." Unpublished term project. Graduate Diploma in TESOL. Sydney College of Advanced Education, 1987.
- Omwig, C. P. (1989). Teacher/student classroom interaction in vocational education. Kentucky: Department of Vocational Education, College of Education, University of Kentucky. (ERIC Document Reproduction Service No. ED342943)
- Özaydınlık, K. (2014). Toplumsal Cinsiyet Temelinde Türkiye’de Kadın ve Eğitim. *Sosyal Politika Çalışmaları Dergisi*, 33: 93-112.

- Parham, F. (2013). Gender representation in children's EFL textbooks. *Theory and Practice in Language Studies*, 3(9): 1674-1679. <https://doi.org/10.4304/tpls.3.9.1674-1679>
- Porreca, K. (1984). Sexism in current ESL textbooks. *TESOL Quarterly*, 18(4): 705– 724.
- Sadker, M., & Sadker, D. (1986). Sexism in the classroom: from grade school to graduate school. *The Phi Delta Kappan*, 67(7): 512-515.
- Sadker, M., & Sadker, D. (1975). *Towards an analysis of discourse: the English used by teachers and pupils*. Oxford: OUP.
- Sinclair, J. & Coulthard, M. (1992). *Toward an analysis of discourse*. In Coulthard, M., *Advances In Spoken Discourse Analysis* (1-34). London: Routledge.
- Spaulding, R.(1963). *Achievement, Creativity and SelfConcept: Correlates of TeacherPupil Transactions in Elementary School*. Washington, DC: US Department of Health Education and Welfare (Cooperative Research Project No. 1352)
- Stake, J.E., & Katz, J.F. (1982). Teacher-pupil relationships in the elementary school classroom: Teacher-gender and pupil-gender differences. *American Educational Research Journal*, 19(3): 465–471.
- Sunderland, J. (1996). *Gendered discourse in the foreign language classroom: Teacherstudent and student-teacher talk, and the social construction of children's femininities and masculinities*. Unpublished doctoral dissertation. University of Lancaster, England.
- Sunderland, J. (2000). Issues of language and gender in second and foreign language education. *Language teaching*, 33(4): 203-223.
- Sunderland, J. (2010). New understandings of gender and language classroom research: texts, teacher talk and student talk. *Language Teaching Research*, 4(2): 149– 173.
- UNESCO .(2003) *Convention for the Safeguarding of the Intangible Cultural Heritage*, Paris, 17 October 2003
<http://www.unesco.org/culture/ich/index.php?pg=00006>
- Uysal, Ş. (1981). *Atatürk ve çağdaş eğitim*. Atatürk ve eğitim. Ankara: Ted Yayınları. Yaşam memnuniyeti araştırması. (2011). Available: http://www.tuik.gov.tr/pretablo.do?alt_id=41.
- West, C., & Zimmerman, D. 1987. Doing gender. *Gender and Society*, 1:125–151.
- Yepez, M.E. (1994). An observation of gender-specific teacher behavior in the ESL classroom. *Sex Roles*, 30(1/2): 121–133.
- Younger, M., &Warrington, M., & Williams, J. (1999). The gender gap and classroom interactions: reality and rhetoric? *British Journal of Sociology of Education*, 20(3): 325–341.

Assessing the Effectiveness of AI in Educational Assessment

Eftelya Şahin*

*Inst., Istanbul University-Cerrahpaşa, Türkiye

Abstract

A crucial part of education, assessment offers insightful information about student learning outcomes and guides instructional strategies. Given the rapid growth of artificial intelligence (AI) technologies, there is growing interest in learning more about how they could be used in educational evaluation. AI has the power to automate and improve several assessment-related processes, providing prospects for increased precision, effectiveness, and individualized feedback. With a focus on the assessment tool ChatGPT, this study proposal aims to examine the effectiveness and impact of AI in educational assessment. The research questions include how AI may be used to test students more accurately, efficiently, and with individualized feedback. They also concern ChatGPT's and other AI technologies' usefulness as a tool for assessing students' knowledge and skills. The major objectives of this study are to examine the efficacy of AI-based assessment tools, look into their potential to automate grading processes, compare ChatGPT with other AI tools, and look into the potential benefits and limitations of utilizing AI in educational evaluation. A sample of prep school students will participate in the research, which will use a quantitative research approach. Studies comparing the accuracy and dependability of assessments based on AI and traditional approaches were carried out. Additionally, AI tools were used to analyze essay exams, and the results including scores, and feedback were compared to teacher evaluations.

Keywords: Artificial Intelligence (AI), assessment, ChatGPT

1. Introduction

A key component of education is assessment, which informs instructional practices and offers important insights into the learning outcomes of students. Artificial intelligence (AI) technologies are developing at a rapid pace, opening up new opportunities to improve the assessment procedure. Artificial Intelligence can enhance and mechanize many duties associated with assessments, ensuring heightened precision, efficacy, and the provision of personalized feedback.

The capabilities of the AI assessment tool ChatGPT are the specific subject of this study, which investigates the efficacy and implications of AI in educational evaluation. There is a greater need than ever for accurate and effective evaluation tools as educational contexts become more diversified.

Insights into student learning outcomes and the development of instructional strategies are provided by assessment, which is a critical component of education. It is a crucial procedure that aids instructors in evaluating the efficacy of their instructional strategies and gives pupils the chance to comprehend their development and potential areas for growth.

However, conventional evaluation techniques, including exams and essays, frequently fall short of effectively capturing students' knowledge, talents, and skills. These techniques may take a long time, be subjective, and not immediately offer feedback to help in students' learning. Due to its ability to provide intelligent feedback generation, automated grading, and flexible learning environments, AI-based assessment systems have the potential to completely change the educational landscape. These systems assess student replies and deliver personalized feedback in real-time using machine learning algorithms and techniques for natural language processing. This helps teachers save time while simultaneously giving students access to fast, personalized feedback that improves their learning. Although AI-based evaluation technologies like ChatGPT have a lot of potentials, it is vital to examine their efficacy and influence on educational assessment. The purpose of this research is to examine ChatGPT, Bard, and MyEssai with a particular focus on how well AI performs in educational evaluation. This study aims to offer insightful information that can educate educators, policymakers, and researchers on the integration of AI in educational settings by exploring the potential gains and limitations of AI-based assessment systems.

1.1 Research Questions

A. How can artificial intelligence be effectively utilized in educational assessment to improve accuracy, and efficiency, and provide personalized feedback?

The first research question aims to explore the effective utilization of artificial intelligence in educational assessment, with a focus on improving assessment accuracy, efficiency, and the provision of personalized feedback.

B. How effective is ChatGPT as an assessment tool for measuring student knowledge and skills?

The second research question investigates the effectiveness of ChatGPT, a specific AI-based assessment tool, in measuring student knowledge and skills, providing insights into its potential as a reliable assessment instrument.

C. How effective is ChatGPT in comparison with other AI tools?

The third research question compares the effectiveness of ChatGPT with other AI tools, such as BARD and MYESSAI, to assess their relative performance as assessment tools, shedding light on the comparative strengths and limitations of different AI-based assessment approaches.

2. Background

2.1 The Role of Assessment in Education

Assessment plays a crucial role in education. To see the progress of students' teachers, have to assess their students' skills and knowledge. Taras (2005) mentions that the stages or procedures necessary to carry out an evaluation are known as the assessment process. In earlier times, different types of assessments were used to assess and evaluate students. Traditionally, assessments have strongly relied on written assignments, tests, and teacher evaluations. Lambert and Lines (2000) also advocate those formal contexts and methods, such as timed exams scored under specific guidelines, are a part of assessment. However, it is now believed that these are not sufficient to assess students' progress. Furthermore, these methods are not efficient in terms of time. Nowadays educators need more flexible, diverse, and inclusive methods to assess their students and give detailed feedback.

2.2 Artificial Intelligence and Education

Artificial Intelligence (AI) has been transforming many areas including education. Many different AI tools or bots can be used in different parts of education. Both educators and students can benefit from AI in education. Some of these benefits are intelligent tutoring systems, translation, learning another language, automated grading, and feedback. AI's capability to bring together human and machine learning offers a great possibility of more efficient teaching and learning than either AI or humans could do on their own (Luan et al., 2020). That shows as humans we can work collaboratively with AI.

There are many other aspects we can mention about the opportunities of AI. For instance, educators can design curriculums, create lesson plans, and find interesting activities with the help of AI. It is also a lot easier to access information with the help of different AI tools such as ChatGPT or BARD. Educators can track their students' attendance or progress in class with these new technologies. According to Karaca and Telli (2019), AI tools help teachers in many parts of education such as course content preparation, individual performance assessment, and selecting the best methodology for the class.

AI creates a more diverse and inclusive environment for different kinds of students. Alkan (2024) states that AI has made it possible to predict differences between the students and the individualized educational content, allowing for the creation of arrangements in the instructional materials. Students with specific needs or disabilities can now access knowledge much more easily thanks to artificial intelligence. The aim is to provide special (impaired)

students full and equal participation in social life. In this way, people with disabilities are included in autonomous and free energy. Tools that convert text to sound can be used to educate visually impaired pupils and systems that convert audio to text can be used to train hearing-impaired pupils. As a result, intelligent technologies improve communication between educators and students, enable information access, make it easier to identify students who are not engaged in the course, and improve teacher-student feedback (Nalbant, 2021).

2.3 Artificial Intelligence and Assessment

Most of the traditional assessments have some drawbacks. Most of them are not authentic especially multiple-choice exams. However, it should be noted that some traditional assessments are really useful and successful. To overcome the drawbacks of traditional assessments, artificial intelligence can be used. Several AI techniques have been created that enable a more continuous picture of performance and insights into learning, in contrast to traditional assessment practices that may take discrete snapshots of performance.

Firstly, using AI as an assessment tool aligns with the constructivist model of learning which is about students and their creating of their understanding and knowledge. With personalized feedback and tailored questions, students can benefit from AI.

An additional important theory that encourages the use of AI in evaluation is cognitive learning theory. The processes of obtaining, organizing, and using knowledge are the main topics of this theory. Artificial intelligence (AI) tools such as ChatGPT and BARD including natural language processing (NLP) algorithms, are made to simulate cognitive processes by examining student responses to find specific patterns, misunderstandings, and opportunities for development. This makes it possible to analyze a learner's cognitive ability more thoroughly than just determining correct or incorrect answers. AI assesses the students' work not just as right or wrong but also give a detailed explanation and feedback for the student.

On the other hand, AI could be useful because it has no human emotion. Errors are minimized using artificial intelligence. Artificial intelligence can make more impartial and professional assessments because it is devoid of human emotions. For instance, a teacher might read the exam paper and project assignments incorrectly and give the wrong grades. However, artificial intelligence programs, reduce errors and make the fewest mistakes possible.

According to Coşkun and Alper (2024), there is a positive connection between AI and human assessments, notwithstanding some individual variations. The similarity between assessments made by AI and teachers indicates that AI-based assessment systems have a place in education. With the correct use of AI technologies, educators can improve the assessment process by taking advantage of the data-driven insights, individualized feedback, and automated grading systems that some AI tools offer.

2.3.1 Effectiveness of AI in Educational Assessment

Over the last ten years or so, artificial intelligence (AI) in educational assessment has not changed much in terms of its fundamental principles and functions—that is, machine learning and actions based on the results of intelligent analysis of large-scale data—but it has advanced significantly in terms of technological efficiency, speed, and sophistication, particularly when

it comes to the analysis of large-scale assessment process data being channeled for formative purposes. (Gardner et al., 2021)

AI provides speed and efficiency in assessment in addition to consistency and sophistication. Traditional assessments, especially lengthy multiple-choice exams or essay-based exams can take days or even weeks to complete by a teacher. On the other hand, AI can evaluate student performance in real time and offer prompt feedback. The time efficiency part will be discussed later in this research according to the comparison of teacher scoring and AI scoring. Because of this timeliness, learning can be more dynamic, and students can get immediate feedback on their work and make necessary corrections before moving on to new material. Teachers may assess essays, and test scores within seconds. Adigüzel et al. (2023) also state that AI may increase productivity, engagement levels, and learning results by creating new opportunities for personalized feedback, and support.

Additionally, research indicates that AI-driven tests may be especially useful for assessing higher-order cognitive abilities, which are challenging to be shown in conventional testing methods like multiple-choice questions. These abilities include creativity, problem-solving, and critical thinking. Essays and other complicated tasks can be evaluated by AI tools that use machine learning and natural language processing (NLP) to analyze the coherence, voice, and structure of an argument. This enables a more comprehensive assessment of students' abilities. Artificial intelligence (AI) tools, such as MyEssai can give lengthy feedback to students' essays which will be discussed later in this paper.

2.3.2 Challenges and Limitations of AI in Educational Assessment

Even while AI-based educational assessment has a lot of potential, there are drawbacks and restrictions. To guarantee that AI can be successfully incorporated into the educational environment, these challenges need to be overcome. Concerns over data privacy and ethical issues, prejudice and fairness, technological constraints, and the elimination of the human element from the assessment process are a few of the major issues.

Data security and privacy are two of the biggest challenges with AI in educational assessment. Large volumes of data, including sensitive information about students' behavior, academic performance, and occasionally personal information, are necessary for AI systems to work efficiently.

The accuracy of the data and algorithms employed also affects how useful AI is in assessing the students. For AI systems to become accurate, a large amount of data is required. Biased or inadequate data that was used to train these systems may introduce errors or reinforce already-existing disparities in the assessment results. AI programs that were predominantly trained on data from particular demographic groups, for instance, would not function as well when evaluating pupils from diverse backgrounds. To enhance the performance of AI and guarantee fairness in assessment, it is imperative to ensure the quality and diversity of training data.

Fairness and bias provide yet another significant difficulty. Although AI is sometimes seen as objective, the truth is that biases in the data that AI tools are trained on can reflect and reinforce these prejudgments. An AI system may generate biased evaluation results if it is trained on biased data, regardless of the biases being tied to gender, color, socioeconomic status, or other

variables. Biased AI algorithms may produce unfair and incorrect assessments. This is because AI systems are only as good as the training data they use. The AI system will yield diverse results if the data it was trained on is biased (Owan et al., 2023).

The absence of human interaction and human emotion in assessments powered by AI is another cause for concern. A personal connection between the teacher and the student is a common feature of education and it may be very beneficial for motivation, engagement, building rapport, and emotional support. In contrast, AI systems don't have the same empathy or motivation that a human teacher could; instead, they give grades and feedback based only on algorithms and the data it has. AI systems are unable to replace the individualized support, emotional support, and mentoring that teachers can provide. There's a chance that education may become unduly automated as AI continues with assessment, which would limit the amount of time that teachers and students can spend interacting deeply. Owan et al. (2023) argue that AI in education may result in less human engagement throughout the evaluation process, which could be harmful to students' educational experiences. Teachers could underestimate emotions and essential signs that are only apparent through interpersonal communication with students.

2.4 The Future of Artificial Intelligence in Education

New opportunities are developing that have the potential to greatly improve teaching and learning as the use of AI in educational evaluation continues to develop. These future directions include the development of more comprehensive and creative assessments, the growth of adaptive learning systems, and the integration of AI with human oversight. In the upcoming years, conducting educational assessments may transform if present restrictions are addressed and novel applications of artificial intelligence are investigated.

Grassini (2023) suggests that giving students the tools and education they need to succeed in an AI-dominated future is crucial, as AI becomes more and more integrated into the professional world. Both students and teachers must know how to use this new technology. To do this, including AI programs like ChatGPT in educational environments can be a big step. We may significantly nurture students' awareness and application of these tools while pointing out their limitations and keeping up with technological advancements by providing them with hands-on experience with them.

According to Adeshola & Adepoju (2023), AI tools should be utilized to supplement and support the work of educators and administrators, not to replace it. It is imperative to remember that it cannot take the place of human judgment and expertise. In general, it is critical to properly examine the needs and concerns of students while applying AI in education.

In conclusion, there is a great deal of promise for improving teaching and learning through the incorporation of AI into educational evaluation. AI-driven solutions can solve many of the challenges of traditional approaches, including their lack of flexibility and time efficiency, by offering more inclusive, efficient, and personalized assessments. The literature review shows how AI may help teachers with curriculum development, personalized feedback, and testing higher-order cognitive skills. There are still issues, though, namely with data privacy, bias, and the possibility of lessening human connection in the classroom.

3. Methodology

This study employed a quantitative research design involving a sample of prep school students. Comparative studies were conducted to collect quantitative data, comparing AI-based assessment tools with traditional assessment methods (e.g., exams). Quantitative data was collected using ChatGPT as an assessment tool, with students completing a series of tasks related to the course material. The same kind of questions were also given to students as a traditional assessment method as a multiple-choice exam allowing for a comparison of the accuracy of ChatGPT. Additionally, the essay assessments of the students were evaluated using different AI tools such as ChatGPT, Bard, and MyEssai, and the scores and feedback generated were compared with the teacher's assessment. The data collected and analyzed provided valuable insights into the effectiveness and limitations of AI in educational assessment.

3.1 Participations

The 13 students from preparatory class served as the study's source of participants. Their ages ranged from 18 to 22 and some of them were from a different country. Several AI tools, including ChatGPT, Bard, and MyEssai, were used to evaluate the students' essays. This made it possible to evaluate and compare the effectiveness of different tools in assessing essays submitted by students in great detail. An extra set of 20 students was added to better explore the contrast between online testing and conventional evaluation techniques. The option to compare the results of online assessments utilizing AI technologies with those obtained through traditional assessment methods was given to these pupils through the assignment of various homework tasks. These prep students' participation offered useful information for examining the efficacy and dependability of AI tools in educational assessment as well as comparing online and traditional assessment techniques.

4. Results

4.1 Comparing Multiple Choice Exams with Online Tasks by AI

For the first part of the research, the purpose of the study was to compare student performance on the tasks given to them through ChatGPT with their midterm test performance to assess the effectiveness of using AI-based online assignments. 50 multiple-choice questions with a grammatical and vocabulary focus made up the midterm exam, which was given online without any face-to-face interaction. However, it was discovered that the exam mostly assessed superficial knowledge, denying students the chance to show off the skills or knowledge they had learned in class. Students nonetheless received impressively high marks on the midterm exam despite their very low levels of active participation in online classes. When answering open-ended questions (Figure 1 and Figure 2), in particular, the students showed increased creativity and engagement, according to the analysis of their task answers. These questions gave students the chance to demonstrate their knowledge of grammar and vocabulary more comprehensively. When the answers of students were checked, it was seen that they used their critical thinking skills more aside from writing. (Figure 3 and Figure 4) One of the students used emoticons to express her feelings about the given questions in Figure 1. For the other task in Figure 2, students show their progress in learning about the topic “should” and “imperatives” with writing. As the questions evolved from different contexts it creates a more authentic environment than the multiple-choice exam. When one of the students’ answers is

checked in Figure 4, it can be seen that it is not completely correct in terms of its grammaticality. However, by getting feedback from the teacher, students can see their mistakes better and it can create a positive washback. The implementation of online assignments allowed a more effective evaluation of students' progress as opposed to the ease associated with taking multiple-choice tests, which led to their widespread use.

The production of such assignments through ChatGPT was discovered to be a quick procedure, which benefited teachers, even though the administration of such tasks proved to be time-consuming.

These results suggest that a more helpful method of evaluating students' growth and knowledge is to include online assignments that are supported by AI technologies like ChatGPT. The typical multiple-choice format has its limitations, but these tasks give students a platform to use their creativity and expertise. Additionally, even if creating online assignments requires more administration time, ChatGPT's ease of use and effectiveness make it a practical choice for teachers.

Figure 1: Homework Questions

- Instructions: Answer the following questions about jobs and careers.
- 1. What do you consider to be the most important factors when choosing a job or career?
- 2. What skills or abilities do you possess that you think would be valuable in your ideal job?
- 3. Have you ever had a job that you didn't like? What were the reasons?
- 4. Do you think it's important to have a job that you enjoy, or is it more important to have a job that pays well?

Figure 2: Homework Questions

HOMEWORK 3

Write different advice for the given problems. Use imperative sentences or should for your advice.

- I feel stressed before exams. What should I do?
- I'm not creative enough at work. Could you give me some advice?
- I don't laugh very often. Can you tell me what I should do?
- I never finish anything I start. Do you know anything I can do?
- I always lose important things. Please help me.

Figure 3: Answers of Students

- Instructions: Answer the following questions about jobs and careers.

1. What do you consider to be the most important factors when choosing a job or career?

I think the most important thing is to have a job that you love and are capable of. Making a lot of money is an important factor, but to what extent? While visiting the university library during the exam week, when I decided to go to university after 4 years, I realized that I had made the right decision by choosing the part of my job because I have never been a person who can study in thick books. My job is to design and do something art related.

2. What skills or abilities do you possess that you think would be valuable in your ideal job?

I think I'm pretty good at designing

I always do my best as long as adequate opportunities are provided.

I work more efficiently under stress

If I am doing a job that I love, I can continue to work outside of my working hours (of course, this should be reflected in my salary 😊)

I'm good at teamwork

In an environment where I can improve myself, I will contribute more to the development of the company.

3. Have you ever had a job that you didn't like? What were the reasons?

Yes. I didn't like working at the last place I worked because I was not respected, I couldn't improve myself, I was doing a lot of work outside of my own job, my father was working here, the workplace was under my house, I was not happy, I was the only person in the office, but the biggest reason for me to leave was because I didn't like it and wasn't happy. It may have helped me a lot, but I knew I deserved better.

4. Do you think it's important to have a job that you enjoy, or is it more important to have a job that pays well?

Of course, making money is important, but I think that if a person does a job he enjoys very well, he will certainly make good money after a while. My choice would be to do the job I enjoy and make good money 😊 there was no such option in the question, but whatever

Figure 4: Answers of Students

1-you should have enough time to study and revise before the exam so that you can feel confident and prepared and Getting enough sleep is important for your physical and mental well-being. Try to get at least 7-8 hours of sleep every night.

2-you should Be open to new ideas, perspectives, and approaches and you should take a break and step away from your work to help you come up with new and fresh ideas.

3-you should Spend time with people who make you laugh or you should Try watching stand-up comedy or sitcoms that make you laugh. You could also read humorous books or comics that are known to be funny.

4-you should Start with small and achievable goals, which will help you build momentum and boost your confidence and you should Celebrate your accomplishments, no matter how small they may be. This can help you stay motivated.

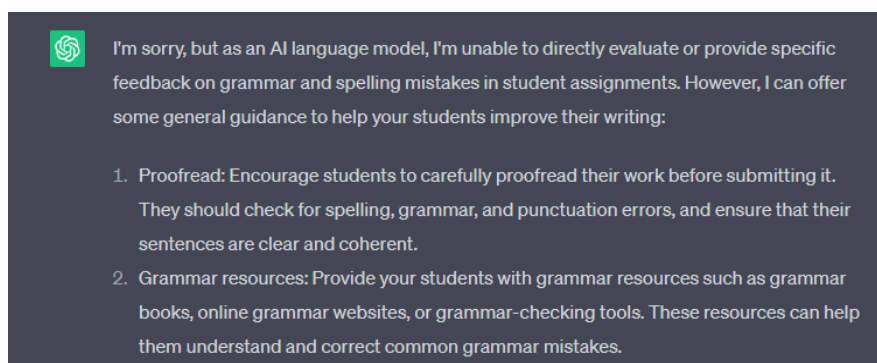
5-you should Make a list of all the important things you need to carry with you before leaving the house. Check the list before leaving to make sure you have everything you need.

4.2 Comparing AI Tool “ChatGPT” and Teacher Assessment for Evaluating Essays

For this part of the research, a sample of 13 students was selected to assess their writing assignments using different AI tools. The assessments generated by the AI tools were then compared with evaluations performed by an instructor(me) using a traditional writing rubric. (Appendix) The purpose was to examine the scores, feedback, and time required for assessment.

Initially, ChatGPT was employed to compare online assessments with traditional methods. One writing assignment is used for this part of the research which asks for an opinion essay about the topic “ Remote (online) work is better than working in an Office”. Two variations of the prompt were used for online assessment with ChatGPT: one requesting both a score and detailed feedback, and another requesting shorter feedback. When the prompt requested feedback for the first time, ChatGPT provided extensive and lengthy responses, focusing on general aspects of the writing while giving suggestions for improvement. However, when the prompt was modified for shorter feedback, the scores given by ChatGPT changed, and it underestimated certain areas, such as spelling mistakes. (Figure 5) It should be also noted that ChatGPT's assessment was based on the complete essay provided as a single block of text, which limited its ability to evaluate the organization of the original document, unlike the teacher.

Figure 5: Difficulties with ChatGPT



Another notable observation was the variability in feedback. While the prompts given to ChatGPT could be modified to elicit different types of feedback, my feedback tended to follow a consistent pattern. This suggests a potential advantage for AI tools in providing varied feedback, although the instructor's feedback may demonstrate a more consistent approach.

When comparing the overall scores given by the AI tools with those provided by the instructor, it became apparent that the AI tools consistently assigned lower scores. (Table 1) Also, ChatGPT showed some inconsistency in providing scores. Having entered the same prompt, it gave different answers. (Figure 6) Furthermore, the time required for assessment differed between the AI tool and the instructor. The initial assessment using ChatGPT, including prompts, took approximately 03:54 minutes. However, for subsequent assessments, the time was reduced to around 1 minute. It is worth noting that the time spent on online assignments sometimes increased when the prompts needed to be re-entered. On the other hand, the instructor required close to 5 minutes to assess each student's writing, although the time varied based on the length of the essay. Finally, it does not provide a plagiarism check like some of the other online tools (Figure 7)

Table 1: Comparison of ChatGPT and Teacher's Scores

	ChatGPT Scores	Teacher's Scores
1	60 (First time) 70 (Second time)	85
2	90	100
3	60	80
4	50	70
5	70	85
6	40	65
7	80	95
8	90	100
9	50	70
10	60	90
11	80	100
12	70	85
13	20	30

Figure 6: Difficulties with ChatGPT

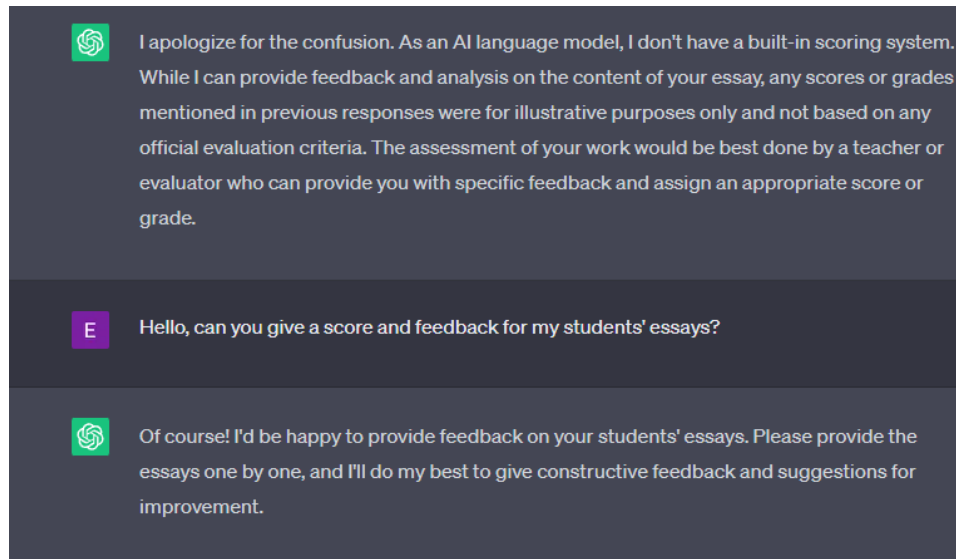
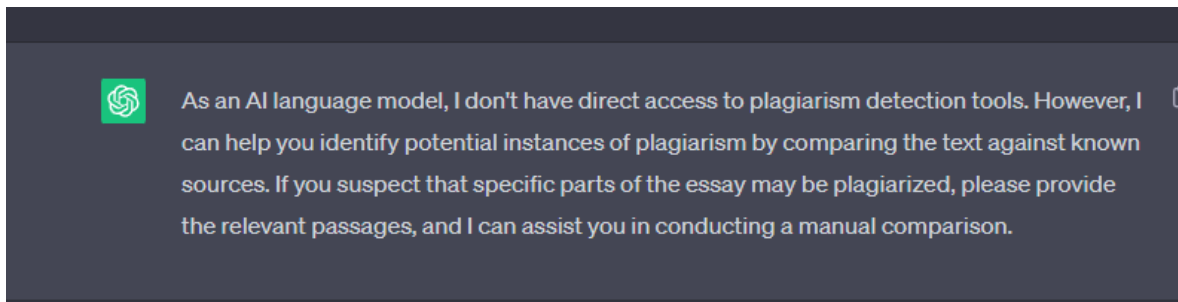


Figure 7: Difficulties with ChatGPT



In conclusion, the results demonstrate that the use of ChatGPT for writing assessments has certain advantages and limitations. While AI tools offer the convenience of generating feedback and scores, they tend to assign lower scores compared to instructor assessments. The time efficiency of AI tools varied, with longer setup time initially but quicker subsequent assessments. Additionally, the instructor's assessment considered additional factors such as organization and exhibited a more consistent feedback pattern. The findings suggest that AI tools can reduce the assessment workload for instructors, but their evaluations may not capture all aspects of writing comprehensively.

4.3 Comparing AI Tool ‘Bard’ and Teacher Assessment for Evaluating Essays

The AI tool known as Bard was used and compared to an instructor's assessment using a writing rubric again for this part of the research. 13 students' essays were evaluated with Bard AI this time. The evaluations focused on different writing assignments with a common topic centered around ‘the importance of learning a second language and the benefits of bilingualism’ Bard, being a relatively new AI tool, offered an opportunity to explore its scoring and feedback mechanisms.

The scoring aspect of Bard demonstrated similarities to ChatGPT, as it consistently assigned lower scores compared to the instructor. The scores given by Bard did not exhibit much

variability, with a majority of the assessments resulting in a score of 80. Unlike ChatGPT, which showed variation in scoring, Bard seemed to adhere to a fixed scoring pattern. The initial feedback generation process with Bard took longer than subsequent iterations, requiring approximately 2 minutes for the first feedback and less than 1 minute for subsequent assessments. In contrast, the instructor spent an average of 5-7 minutes providing feedback and assigning scores.

A notable difference between Bard and ChatGPT was the requirement to enter the same prompt each time shorter feedback was desired. This limitation restricted the flexibility of Bard compared to ChatGPT. Furthermore, the feedback generated by Bard appeared to lack variety and personalization. The feedback was often repetitive, and mechanistic, and failed to address specific aspects such as coherence, spelling, and grammar.

To obtain more detailed feedback, it was necessary to modify the prompt. In contrast, the instructor's feedback showcased more personalization and creativity, allowing for different feedback for each student and the ability to adjust the tone of the feedback accordingly.

Despite its drawbacks, Bard did offer the advantage of providing detailed feedback that focused on both general and specific aspects of the essay. Additionally, the time required for feedback generation was significantly shorter with Bard compared to the instructor's manual assessment.

The analysis of Bard and the instructor's evaluation revealed some weaknesses and strengths of the AI technology. Bard had the advantage of producing extensive feedback concentrating on various areas of the essay, even though it's scoring regularly produced lower marks and its comments lacked diversity and personalization. Bard's time effectiveness was also helpful. These findings highlight the necessity to consider the AI tools' limitations when evaluating writing, notably in terms of scoring consistency and feedback quality. Future studies should look into ways to keep AI tools' time efficiency while enhancing the personalization and variety of input they produce

Table 2: Comparison of Bard AI and Teacher's Scores

	Bard AI Scores	Teacher's Scores
1	80	95
2	80	90
3	80	95
4	80	80
5	70	85
6	70	90
7	80	70
8	90	100
9	80	90
10	70	80
11	90	100
12	70	85
13	70	60

4.4 Comparison of AI Tool ‘MyEssai’ and Teacher Assessment for Evaluating Essays

MyEssai is a strong AI-powered essay teacher that offers pupils thorough critique and helpful suggestions to raise the level of their essays. MyEssai offers in-depth feedback on the content, structure, and organization of essays, in contrast to other grammar checkers. In this phase of the study, the assessment of student essays involved the utilization of myEssai, a relatively new AI tool. Designed to offer detailed feedback, myEssai provided an opportunity to examine its efficacy compared to other AI tools, such as ChatGPT and Bard. The analysis focused on aspects such as feedback quality, scoring mechanisms, and time efficiency.

MyEssai required the input of a single prompt that specified the essay topic, style, and tone. Following the prompt entry, myEssai generated comprehensive feedback, addressing various evaluation criteria, including content, grammar, and organization. Particularly in the initial stages, the feedback provided by myEssai was highly detailed and aligned closely with the designated evaluation standards. However, as the assessment process continued, the feedback became more mechanical in nature, necessitating manual intervention to personalize and enhance its human-like qualities. Despite this limitation, myEssai emerged as the most effective AI tool in terms of feedback provision, surpassing the capabilities of ChatGPT and Bard.

Unlike ChatGPT and Bard, myEssai did not assign scores to the essays. Instead, it primarily focused on delivering detailed feedback. While the absence of scoring may limit its ability to provide a quantitative measure of performance, myEssai compensated for this by offering a thorough analysis of various writing aspects.

In terms of time efficiency, myEssai showcased a promising advantage. It required approximately 30 minutes to evaluate all the students' essays. In contrast, the manual assessment conducted by the researcher took nearly double the time, amounting to approximately 65 minutes. This significant time-saving aspect highlights the potential benefits of integrating AI tools like myEssai into the essay evaluation process, particularly in large-scale assessments.

5. Discussion

This study utilized a quantitative research design to compare AI-based assessment tools with traditional assessment methods in the context of evaluating prep school students. The findings provide valuable insights into the effectiveness and limitations of AI in educational assessment.

The comparison between AI-based online assignments using ChatGPT and traditional multiple-choice exams revealed that online assignments provided a more comprehensive evaluation of student's progress, allowing them to demonstrate their knowledge, critical thinking skills, and creativity. In contrast, multiple-choice exams primarily assessed superficial knowledge, limiting students' ability to showcase their true abilities.

The evaluation of AI tools, including ChatGPT, Bard, and MyEssai, for assessing students' essays highlighted both advantages and limitations. AI tools offered the convenience of generating feedback and scores, reducing the assessment workload for instructors. However, they tended to assign lower scores compared to instructor assessments, and the feedback lacked

consistency, personalization, and variety. The instructor's assessment, on the other hand, considered additional factors such as organization and provided more consistent feedback.

The comparison of AI tools indicated that each tool had its strengths and weaknesses. ChatGPT allowed for creativity and engagement but required careful prompt design to ensure accurate scoring. Bard provided consistent scoring but lacked flexibility in prompt modification and offered repetitive feedback. MyEssai delivered detailed feedback but lacked scoring capabilities.

Each tool demonstrated time efficiency, with AI tools generally requiring less time for assessment compared to manual evaluation by the instructor.

In general, the usage of AI-powered assignment and assessment platforms like ChatGPT, Bard, and MyEssai shows potential in improving the evaluation of students' improvement and comprehension. But to give a thorough assessment of student abilities, it is crucial to take into account the limits of AI technologies and augment their assessments with further measurements.

6. Conclusion

This study utilized a quantitative research design to compare AI-based assessment tools with traditional assessment methods in the context of evaluating prep school students. The findings provide valuable insights into the effectiveness and limitations of AI in educational assessment.

The comparison between AI-based online assignments using ChatGPT and traditional multiple-choice exams revealed that online assignments provided a more comprehensive evaluation of student's progress, allowing them to demonstrate their knowledge, critical thinking skills, and creativity. In contrast, multiple-choice exams primarily assessed superficial knowledge, limiting students' ability to showcase their true abilities.

The evaluation of AI tools, including ChatGPT, Bard, and MyEssai, for assessing students' essays highlighted both advantages and limitations. AI tools offered the convenience of generating feedback and scores, reducing the assessment workload for instructors. However, they tended to assign lower scores compared to instructor assessments, and the feedback lacked consistency, personalization, and variety. The instructor's assessment, on the other hand, considered additional factors such as organization and provided more consistent feedback.

The comparison of AI tools indicated that each tool had its strengths and weaknesses. ChatGPT allowed for creativity and engagement but required careful prompt design to ensure accurate scoring. Bard provided consistent scoring but lacked flexibility in prompt modification and offered repetitive feedback. MyEssai delivered detailed feedback but lacked scoring capabilities. Each tool demonstrated time efficiency, with AI tools generally requiring less time for assessment compared to manual evaluation by the instructor. In general, the usage of AI-powered assignment and assessment platforms like ChatGPT, Bard, and MyEssai shows potential in improving the evaluation of students' improvement and comprehension. But to give a thorough assessment of student abilities, it is crucial to take into account the limits of AI technologies and augment their assessments with further measurements.

Acknowledgment

First and foremost, I would like to express my deepest gratitude to my professor at University Tuncer Can for his guidance, support, and encouragement throughout the course of this research. Your expertise and insightful feedback have been crucial in shaping this study. I extend my heartfelt thanks to the administration and students of Nisantasi University, particularly the prep school students who participated in this study. Your cooperation and enthusiasm were essential for the completion of this research.

I also acknowledge the contributions of AI technologies, particularly ChatGPT, which played a significant role in the assessment processes explored in this study. The advancements in AI have opened new horizons for educational evaluation, and this research would not have been possible without these technological innovations. Thank you all for your support and contributions to this work.

Appendix

ESSAY WRITING RUBRIC				
	Excellent	Good	Satisfactory	Needs Improvement
Content (25 pts.)	Writing is confident and clearly focused. It holds the reader's attention. Relevant details enrich writing.	Writing is purposeful and focused. Piece contains some details.	Writes related, quality paragraphs, with little or no details.	Writing is limited in communicating knowledge. Length is not adequate for development.
Coherence (25 pts.)	Well-planned and well-thought out. Includes title, introduction, statement of main idea, transitions and conclusion.	Good overall organization, includes the main organizational tools.	There is a sense of organization, although some of the organizational tools are used weakly or missing.	No sense of organization.
Grammar & Mechanics (20pts.)	Excellent grammar, spelling, syntax and punctuation.	A few errors in grammar, spelling, syntax and punctuation, but not many	Shows a pattern of errors in spelling, grammar, syntax and/or punctuation. Could also be a sign of lack of proofreading.	Continuous errors
Voice (20pts.)	Writes with a distinct, unique voice/point of view. Writing is skillfully adapted to the audience.	Writer has strong voice/ point of view. Writing engages the audience.	Writer uses voice/point of view. Writes with the understanding of a specific audience.	Writer's voice/ point of view shows little sense of audience.

References

- Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2023.2253858>
- Adeshola, I., & Adepoju, A. P. (2023). The opportunities and challenges of ChatGPT in education. *Interactive Learning Environments*, 1–14. <https://doi.org/10.1080/10494820.2023.2253858>
- Adiguzel, T., Kaya, M. H., & Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), ep429. <https://doi.org/10.30935/cedtech/13152>
- Alkan, A. (2024). Artificial intelligence: Its role and potential in education. *Journal of the Human and Social Science Researches*, 13(1), 483-497. <https://doi.org/10.15869/itobiad.1331201>
- Alper, A., & Coşkun, T. K. (2024). Evaluating the evaluators: A comparative study of AI and teacher assessments in higher education. *Artificial Intelligence from the Perspective of Educability*. <https://doi.org/10.1344/der.2024.45.124-140>
- Gardner, J., O'Leary, M., & Yuan, L. (2021). Artificial intelligence in educational assessment: 'Breakthrough? Or buncombe and ballyhoo?'. *Journal of Computer Assisted Learning*, 37(5), 1207–1216. <https://doi.org/10.1111/jcal.12577>
- Grassini S.(2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*. 13(7):692. <https://doi.org/10.3390/educsci13070692>
- Karaca, B. & Telli, G. (2019). Yapay zekânın çeşitli süreçlerdeki rolü ve tahminleme fonksiyonu. G. Telli (Ed.), *Yapay zekâ ve gelecek* (172-185). İstanbul: Doğu Kitapevi.
- Lambert, D., & Lines, D. (2000). *Understanding Assessment: Purposes, Perceptions, Practices*. London, Routledge Falmer.
- Luan, H., Géczy, P., Lai, H., Gobert, J., Yang, S., Ogata, H., ... & Tsai, C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in Psychology*, 11. <https://doi.org/10.3389/fpsyg.2020.580820>
- Nalbant, K. G. (2021). The importance of artificial intelligence in education: A short review. *Journal of Review in Science and Engineering*.
- Owan, V. J., Abang, K. B., Idika, D. O., Etta, E. O., & Bassey, B. A. (2023). Exploring the potential of artificial intelligence tools in educational measurement and assessment. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(8), em2307. <https://doi.org/10.29333/ejmste/13428>

Taras, M. (2005). Assessment- Summative and Formative-Some Theoretical Reflections. *British Journal of Educational Studies*, 53(4), 466-478. <https://doi.org/10.1111/j.1467-8527.2005.00307.x>

AI and ELT: Pre-service EFL Teachers' Conceptions on AI and its Implementation

Gürkan Temiz¹, Elif Nazlı Kafadar², Onur Kubilay Pak³

¹Dr., Istanbul Technical University, Türkiye

²MA student, Marmara University, Türkiye

³MA student, Sakarya University, Türkiye

Abstract

Recent technological advancements have caused significant developments in the field of education and one such transformation is the advent of artificial intelligence (AI). Literature has shown the interplay between perceptions and classroom practices in education and it has been highlighted that it is crucial to investigate pre-service language teachers' (PSLTs) perceptions regarding the concept of AI and to find out how they envision integrating these technological developments into the educational process in the field of English Language Teaching (ELT). Therefore, this study aims to evaluate the awareness levels of 60 PSLTs studying at a state university in Istanbul, Türkiye, concerning the concept of artificial intelligence and its use in ELT. Also, the study seeks to comprehend pre-service language teachers' views to explore their preferences regarding integrating AI-based tools into their future classes through interviews conducted with 10 of the participants. By using a survey instrument and conducting interviews to explore what the participants think about the integration of AI, this research will provide insights into pre-service language teachers' perspectives and provide a base to understand the possible implications of this integration into the field. In addition, collecting information about the views and experiences of PSLTs regarding the use of AI tools in ELT may guide the shaping of future educational programs. The findings of the study will be shared and discussed along with their implications for further studies in the field of ELT.

Keywords: Artificial intelligence, AI, pre-service language teachers, ELT

1. Introduction

Recent technological advancements have caused significant developments in the field of education. One such transformation is the advent of artificial intelligence (AI). In this study, it has been highlighted that it is crucial to assess the perceptions of pre-service teachers regarding the concept of AI and to find out how they envision integrating these technological developments into the educational process. This study aims to evaluate the perceptions of 80 pre-service teachers studying in a state university in Istanbul, Türkiye, concerning the concept of artificial intelligence and its use in ELT. Also, the study seeks to comprehend the views of prospective teachers to explore their preferences regarding the integration of AI-based tools into their classes. Based on what the participants think about the integration of AI, this research will provide insights into their perspectives and provide a base for understanding the possible implications of this integration into the field. In addition, collecting information about the views and experiences of pre-service teachers regarding the use of AI tools in ELT may guide the shaping of future educational programs.

1.1 The Explanation of AI

Artificial intelligence, which can complete complex tasks without being human, is a man-made product that is gradually approaching the human understanding of the human brain, especially in thinking, reacting, and interacting (Pirim, 2006). Artificial intelligence (AI) refers to a constellation of technologies, including machine learning, perception, reasoning, and natural language processing (Crawford et al., 2016). AI technologies have existed if not as a concept for quite a long time however, as Buchanan (2005) states, ever since Homer wrote of mechanical “tripods” waiting on the gods at dinner, imagined mechanical assistants have been a part of our culture. The name AI is also not something entirely new. The word “Artificial Intelligence” was officially coined in 1956 Marvin Minsky and John McCarthy hosted the approximately eight-week-long Dartmouth Summer Research Project on Artificial Intelligence (DSRPAI) at Dartmouth College in New Hampshire (Haenlein & Kaplan, 2019). Since then, there have been many developments in AI technologies and AI has become a household name. Since AI has become so popular, there has been and very likely will be a lot of research on the topic. To give some examples of prior research, specifically on the awareness of AI, a recent study conducted at the University of Bologna by Sartori & Bocca (2023) assessed the level of AI awareness by asking whether participants had heard, read, or seen material related to the topic of AI in the last 12 months. 76% of participants answered positively, and 16% negatively, while the remaining 8% said they were unable to say for sure if what they had read, heard, or seen had anything to do with AI. Another similar study by Cave et al. (2019) that tests the awareness of AI narratives in the UK general population gives a similar result, 85% of the respondents in their research claimed to have heard of AI. 11% had not heard of it and 4% answered “do not know”. Awareness was high across all age groups. In addition to being a promising concept that stands out among the technological possibilities of the rapidly developing and changing world, artificial intelligence has been defined by Yuan (2021) as a concept that has become the main driving force of growth and innovation in all sectors and has begun to be used in the field of education. As Pirim (2006) stated, it can be said that artificial intelligence, which can complete complex tasks without a human being, is gradually approaching human understanding, especially in thinking, reacting, and interacting.

1.2 AI in the Field of Education

In today's society, we face AI technologies often, therefore it is only natural that there are many opinions on the subject. One of the areas where AI use has great importance is education. AI is at the top of the agenda for education leaders today in educating the next generation across the globe (Ng et al., 2021). And, as Zhang et al. (2023) state, pre-service teachers are the future of education, and they play a critical role in shaping the adoption of AI technologies in the classroom. Their acceptance of AI will greatly influence its future development and application in education. Since pre-service teachers are the future of education, they play a critical role in shaping the adoption of AI technologies in the classroom. Their acceptance of AI will greatly influence its future development and application in education. Given the challenges posed by AI in education, pre-service teachers are well-positioned to address these issues and explore effective ways to integrate these tools in the classroom. In addition, studies measuring the effect of artificial intelligence on academic success suggest that student interaction and success may increase due to the increase in opportunities offered to students (Popenici and Kerr, 2017). Artificial intelligence may be asked for support to enable collaborative learning and enable more efficient use of time for teachers. Also, the preparation of needed materials can be achieved with artificial intelligence applications (İşler, 2020).

According to Edmett et al. (2023), due to English being a global lingua franca, there is a strong global demand for English language learning. However, this demand also brings about a search for both effective and cost-effective teaching methods. In this regard, it has been argued that artificial intelligence plays a more useful role in ELT compared to other disciplines. Additionally, as stated by Ganeesh and Rani (2023), gamification skills, which are a great necessity in language learning, may be integrated into lesson plans with the help of artificial intelligence. Thus, students' contribution and motivation towards the language learning process can be increased. As Koraishi (2023) states, by harnessing the power of ChatGPT, educators can automatically create lesson plans tailored to specific themes, grammatical objectives, and learning outcomes, all of which can be customized according to their instructional needs. However, even if artificial intelligence can completely transform the dynamics of lesson plans, it should not be ignored that it cannot carry the human characteristics of teachers such as emotional support and personalized attitude. The combination of ELT and AI is stated as ELT 3.0 by Thadphoothon (2022), which means the bidirectional gradual adoption of theories, hardware, software, and applications in the ELT field. The basis of all this interaction is the inseparable nature of language and communication. Junaidi et al. (2020) conducted an experimental study and tried to see whether artificial intelligence is really effective in speaking skills. Lyria virtual assistant was used in the experiment and a more significant increase in success was observed between pre-test and post-test results in the test group compared to the control group. The advantages brought by artificial intelligence to ELT classrooms, as stated by Nur Fitria (2021), include the creation of friendly need-based instructional programs, providing teachers with qualitative teaching-learning material that operates on all levels of language areas, offering a supplementary teacher and student support system, providing a fast feedback system, redefining the role of the teacher as a guide, enabling a more personalized

teaching approach, being accessible from anywhere in the world, and offering students the opportunity to learn at their own pace.

1.3 Relevant Research on AI

Some systems include artificial intelligence, and today these systems are programs that can perform the tasks of expert individuals and have a decision-making mechanism (Önder, 2003). SCHOLAR, the first of these systems, is an artificial intelligence program that uses various teaching techniques like a teacher asking questions to students with the Socratic method, trying to teach the knowledge that is intended to be conveyed in this way (Carbonell, 1970). As cited by Arslan (2020), another example of these systems is Dialogue-Based Learning Systems, which are artificial intelligence-supported systems based on teaching knowledge using the logic of dialogues. In this context, it can be said that artificial intelligence can play an effective role in teaching by using the information that an expert knows, perhaps by putting into practice information that the human memory cannot focus on. Since we live in a technological era, it has become a must to have digital literacy as well as literacy on the use of AI tools. The rate of awareness on this issue makes itself more important each day. From another perspective, it may be mentioned that AI is becoming part of our lives and during this transition, people's intention to use AI technologies is still unclear and emotions such as fear are influencing it (Cugurullo & Acheampong, 2023). The data in the study done by Sartori & Bocaa (2023) supports this point since it found out that about 80% of people who participated in their study appeared to have more positive than negative opinions on AI. As a result, the transformation of education and teachers may cause an increase in the quality of teaching as a consequence of integrating artificial intelligence and education (Coşkun & Kuşçu, 2021).

1.4 Problem

Recent technological advancements have caused significant developments in the field of education. One such transformation is the advent of artificial intelligence (AI). In this study, it has been highlighted that it is crucial to assess the perceptions of pre-service teachers regarding the concept of AI and to find out how they envision integrating these technological developments into the educational process. This study aims to evaluate the perceptions of 80 pre-service teachers studying in a state university in Istanbul, Türkiye, concerning the concept of artificial intelligence and its use in ELT. Also, the study seeks to comprehend the views of prospective teachers to explore their preferences regarding the integration of AI-based tools into their classes. By what the participants think about the integration of AI, this research will provide insights into their perspectives and provide a base to understand the possible implications of this integration into the field. In addition, collecting information about the views and experiences of pre-service teachers regarding the use of AI tools in ELT may guide the shaping of future educational programs.

1.5 Aim of the Study

This study aims to assess the perceptions of pre-service teachers on artificial intelligence and the potential effect of integrating AI tools into ELT. In addition, it aims to collect information on how familiar pre-service teachers are with AI literacy and whether they show motivation to integrate those tools into their lessons. In this context, this study aims to contribute valuable

insights for shaping future educational strategies related to the integration processes of AI tools in educational curricula. The research questions are as follows:

1. What are the perceptions of EFL pre-service teachers on the use of AI in ELT?
2. In what ways do EFL pre-service teachers associate the integration of AI with ELT?

1.6 Significance of the Study

The importance of this study is that it can contribute to our understanding of how ready pre-service teachers are to integrate artificial intelligence tools into their classroom teaching, in light of the findings collected by measuring pre-service English teachers' attitudes and awareness towards the use of AI. These perspectives can guide the development of strategies on how to effectively integrate educational technology, especially in English classrooms. The results of the study can provide information to curriculum developers and education policymakers about the preferences and attitudes of pre-service teachers towards artificial intelligence, which is important for the education of the future. The study can form the basis for designing targeted professional development activities for pre-service teachers and in-service teachers. Assessing pre-service teachers' motivation to integrate artificial intelligence into English language teaching highlights the potential power of implementing innovative approaches in language education. Since tools containing artificial intelligence are rapidly developing tools in the field of education, this study can offer fresh perspectives to the literature by filling the existing gaps in the perspectives of prospective teachers. In this context, it has the potential to contribute to ongoing discussions by providing information on artificial intelligence integration in the field of foreign language education.

2. Method

This study followed a mixed-methods research design. Creswell and Plano Clark (2011) describe this design as "a research design that has its philosophical assumptions and methods of inquiry". This study combines both qualitative and quantitative data to provide a comprehensive view of pre-service language teachers' AI use, their perceptions, and preferences for this rapidly developing technology. Regarding the participants, the researchers accessed 60 senior pre-service English Language Teaching students who have practicum experience-enrolled at a state university in Istanbul, Türkiye. The qualitative phase included interviews with 6 students while the quantitative part was completed with 60 students.

As for the data collection instruments, they include a scale and semi-structured interview questions. The scale "Artificial Intelligence Awareness Scale for Teacher Candidates" developed by Dumlupınar and Arslan (2023) was used to gather information about the participants' familiarity with AI, their general attitudes toward it, and their level of motivation to integrate AI tools into their classrooms. As for the interview questions, the questions were developed by the researchers to explore how and what AI tools could be used in language teaching practices. The Cronbach's alpha for the scale used in Dumlupınar and Arslan's (2023) study was 0.71. Therefore, this value indicates that the scale has high internal consistency and was appropriate to be used in such a quantitative study.

During the data collection phase, the scale was applied to the entire group at the same time and the data was collected digitally on Google Forms. Descriptive statistics were run on the survey

data. Interviews were conducted in a quiet and interview-friendly environment, with one-on-one interviews with the participants. To identify the general patterns in the participants' perceptions of AI, quantitative analysis had been done by using descriptive and correlation statistics via the SPSS program. For the data collected using interviews, the qualitative analysis is done using Thematic Content Analysis with MAXQDA software

3. Findings

As for the first research question, the data collected via the survey instrument was analyzed. The scale had four dimensions which are “Basic Knowledge Level”, “Negative Prediction”, “Benefit” and “Prejudice”. According to the descriptive statistics, the general attitude of the participants in the sample can be measured. The basic knowledge level of the participants is considered high, which is 15,30 out of 20. Even though there are some negative predictions about AI use, participants in this study believe that there are more benefits than its negative effects. The scale included two items related to prejudice, and the mean statistic being 5,40 is considered high again. The findings obtained from all dimensions are considered normal according to their skewness and kurtosis values.

Table 1. Item-based survey findings

Item / Statistics	N	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness Statistic	Kurtosis Statistic
I have basic knowledge about the concept of artificial intelligence.	60	1,00	5,00	4,13	,999	-1,014	,461
I have a basic knowledge of at least one of the concepts such as machine learning, augmented reality, intelligent systems, robotics, neural networks, natural language processing and what it does.	60	1,00	5,00	3,52	1,321	-,265	-1,176
I have basic knowledge about artificial intelligence technologies and application areas.	60	1,00	5,00	3,75	1,114	-,395	-,892
Developments in the field of artificial intelligence interest me.	60	1,00	5,00	3,88	1,180	-,855	-,114
I think that artificial intelligence-based learning systems will replace educators (teachers, academicians, etc.) in the future and make	60	1,00	5,00	3,05	,999	-,314	-,185

these professional groups unemployed.							
I think artificial intelligence will eliminate the sense of trust in society.	60	1,00	5,00	3,43	1,140	-,504	-,482
The application of artificial intelligence applications in the field of education will cause more harm than good for students and teachers.	60	1,00	5,00	2,73	1,118	,328	-,394
I think that artificial intelligence will surpass human intelligence in all matters in the near future.	60	1,00	5,00	2,78	1,367	,201	-1,226
Artificial intelligence awareness needs to be brought to students from an early age.	60	1,00	5,00	3,87	1,171	-,846	-,087
Artificial intelligence makes students' lives easier by supporting learning activities.	60	2,00	5,00	4,17	,806	-,718	,035
Artificial intelligence makes teachers' lives easier by supporting teaching activities.	60	2,00	5,00	4,23	,789	-,659	-,396
Artificial intelligence can create a personalized study program for each student, taking into account student levels.	60	2,00	5,00	4,08	,962	-,762	-,405
I think artificial intelligence is difficult to understand by people working in verbal fields.	60	1,00	5,00	2,90	1,399	,030	-1,320
I think the concept of artificial intelligence is exaggerated today.	60	1,00	5,00	2,48	1,359	,560	-,894

When it is analyzed item by item, the quantitative results obtained from the Likert scale showed that the participants think that AI makes teachers' lives easier by supporting teaching activities ($M=4,23$) and they have basic knowledge about AI ($M=4,13$). In addition, they think that AI makes students' lives easier by supporting learning activities ($M=4,17$). Moreover, they think that AI can create a personalized study program for each student, taking into account student levels ($M=4,08$). Also, the participants reflected that developments in the field of AI interest them ($M=3,88$). After all these positive attitudes, still, some item results make their reflections questionable, since they are related to their basic knowledge. Those are: they have basic

knowledge about artificial intelligence technologies and application areas ($M=3,75$) and at least one of the concepts such as machine learning, augmented reality, intelligent systems, robotics, neural networks, natural language processing, and what it does ($M=3,52$). Also, they think that AI awareness needs to be brought to students from an early age ($M=3,87$).

One of the most striking findings of this study was their perception that AI will eliminate the sense of trust in society ($M=3,43$). Still, most of them disagree that AI will surpass human intelligence in all matters soon ($M=2,78$). Moreover, most student teachers disagree that the application of AI in education will cause more harm than benefit for students and teachers ($M=2,73$).

To gain a deeper understanding of the research findings related to the research question, the test results have been examined according to the four dimensions that the scale has.

Table 2. The Participants' AI awareness across dimensions

	N Statistic	Minimum Statistic	Maximum Statistic	Mean Statistic	Std. Deviation Statistic	Skewness Statistic	Kurtosis Statistic
Basic Knowledge (4 items)	60	4,00	20,00	15,2833	3,84924	-,600	-,178
Negative Prediction (4 items)	60	5,00	18,00	12,0000	3,09729	,018	-,438
Benefit (4 items)	60	10,00	20,00	16,3500	3,00184	-,708	-,514
Prejudice (2 items)	60	2,00	10,00	5,3833	2,17919	,229	-,712

After the overall findings which are obtained through the descriptive statistics, to achieve a more robust finding which can be used for further research, Pearson Correlation was used to analyze the relationship between the participants' attitudes comparing the dimensions. According to the findings obtained from Pearson Correlation, Basic Knowledge Level and Prejudice dimensions have a moderate negative correlation ($p=0,004$). Between Negative Prediction and Prejudice, again, there is a moderate negative correlation ($p<0,001$). According to these findings, pre-service English teachers in this sample with higher total knowledge see AI as a tool that is easy to understand and a tool not to be exaggerated. In addition, those who hold this view also do not believe that AI will put them out of work, destroy their sense of security, do more harm than good, and leave human intelligence behind.

Regarding the second research question, the interview data was analyzed in terms of how the participants associate the use of AI with ELT practices. In line with the integration of AI in.

ELT practices, the data showed that the participants reflected on the opportunities ($f=8$) and the drawbacks ($f=1$) as well as the softwares ($f=3$) that were regarded as useful AI tools in ELT.

Table 3. The Participants' reflections on the integration of AI into ELT

Theme	Category	Code	Frequency (f)
Integration of AI into English Language Teaching	Perceived Opportunities	Material Preparation	4
		Lesson Planning	3
		Virtual Friend	1
	Perceived Drawbacks	Reduced Creativity	1
	Perceived Useful AI Tools	ChatGPT	2
		Image Generator	1

Table 3 shows that the participants associate AI with ELT mostly based on the opportunities that it provides. The data revealed that the participants perceived AI as an opportunity for material preparation ($f=4$) in the field of language teaching. As the excerpt below shows, participants stated that AI can help them enrich the teaching material by making it easy to edit the available or self-made material.

If I want to prepare a checklist for the material, in addition to the questions I think about on that checklist, I can find and evaluate different questions and edit the material I have made myself. In the same way, I can enrich the activity I am considering while teaching a structure or a certain skill. (Excerpt 1, P5)

Additionally, lesson planning ($f=3$) was also addressed as another area in which AI provides an opportunity. According to the participants, AI can be beneficial in writing lesson objectives that can be difficult sometimes. In the excerpt below, P3 states that AI can automatically write lesson objectives considering the lesson content and material, which can free the teacher to deal with such difficult and time-consuming aspects of lesson planning.

I think the best part and the most helpful part of AI is the lesson planning part. You know, you don't have to come up with all of these objectives. (Excerpt 2, P3)

Lastly, the interview data showed that AI can serve as a virtual friend ($f=1$) for children and this can help them improve their productive skills.

I think this would be things like pen pals etc. Even if this is not an actual living friend, maybe children can use artificial intelligence. It can also be useful for them in terms of both speaking and writing. (Excerpt 3, P4)

Apart from the opportunities that AI could provide, causing decreased creativity ($f=1$) was reported as the only downside of using AI in the field. According to one participant, using AI negatively affects his/her creativity and makes him/her lazy.

As a whole, I think there is one negative side to ChatGPT. It makes me less creative. I think I work less. It makes you lazier in a way. It makes my imaginative side lazier. (Excerpt 4, P3)

In terms of AI applications, as Table 2 indicates, the participants reported that two AI tools, ChatGPT and image generator, can be useful in the field. Regarding the role of image generators, one participant stated that they can be useful in teaching vocabulary to young learners.

I could think of young learners right now. We could use image generators to create images that we can show the students to teach them words. You know, the image AI generated can be of great help. (Excerpt 5, P2)

4. Discussion

The findings of this study suggest that pre-service English teachers in this sample have a good level of basic knowledge of AI and AI tools, and they find those tools useful for their teachings. Previous research concluded by Ng et. al (2011) has also shown that increased knowledge of AI may mitigate some concerns brought by pre-service teachers. They are open to benefiting from its advantages; however, the participants still have some concerns about its impact on societal trust while remaining optimistic about AI's role in ELT. In contrast, Zhang et al. (2003) reported that there is a higher level of apprehension by a sample that consists of pre-service teachers again, regarding the impact of artificial intelligence on educational activities. In this study, the participants have shown a more optimistic outlook. This might be because of the time difference between those two studies, implying that the participants of this study are members of Gen Z and were born into the AI world.

Considering the moderate negative correlations between their basic knowledge levels and prejudice, the findings indicate that increasing their knowledge levels of AI would help reduce their concerns. From this perspective, the findings emphasize the significant importance of enhancing the in-service training on AI and its implementations is a need in terms of having an updated approach to ELT. These findings support the necessity for in-service training activities as suggested by Popenici and Kerr (2017).

Taking into consideration that the sample includes individuals from Generation Z, it is important to remember that most of them may have a dynamic digital literacy. Thus, instead of designing curricula on how to use those tools, it may be strategic to take advantage of this reality and show them how to integrate AI tools into their activities according to the age or level of the students. Like Edmett et al. (2023) reported that there is a demand for innovative teaching methods in ELT, this study has also met with this statement by showing well-positioned pre-

service English teachers that can handle the learning process quite efficiently, and have a good level of Basic Knowledge on AI.

Some student-friendly AI tools may be categorized and hands-on practices may be offered to pre-service English teachers, as also suggested by Ganeesh and Rani (2023), to enhance student enhancement. To minimize the idea that AI will endanger the future of their professions, it is possible to eliminate question marks about AI with communicative and interactive training that includes high-level cognitive skills. Thus, it can be made possible for them to write lesson plans that meet the needs of the new century and include technology-integrated activities.

5. Conclusion

Previous research and the conclusions drawn from this study indicate that PSLTs are open to using AI in their prospective teaching practices since they mostly address it as an opportunity in the field. While the survey data of this study revealed positive outcomes for basic knowledge and benefit categories, negative prediction and prejudice dimensions were marked lower by the PSLTs. In line with interview findings and the aforementioned literature, it can be concluded that AI in ELT can positively contribute to teaching practices and ELT programs should be revised aligning themselves to the use of AI in the field regarding the fact that PSLTs are willing to implement it in their future teaching practices.

References

- Arslan, K. (2020). Artificial Intelligence and Applications in Education. *Western Anatolia Journal of Educational Sciences*, 11 (1), pp. 71-80.
- Buchanan, D., Fitzgerald, L., Ketley, D., Gollop, R., Jones, J.L., Lamont, S.S., Neath, A. and Whitby, E. (2005), No going back: A review of the literature on sustaining organizational change. *International Journal of Management Reviews*, 7: pp. 189-205.
- Carbonell, J. R. (1970). AI in CAI: An artificial-intelligence approach to computer-assisted instruction. *IEEE transactions on man-machine systems*, 11(4), pp. 190-202.
- Coşkun, O., & Kuşçu, E. (2021). Artificial Intelligence's Pupil Natural Language Processing. *Turkophone*, 8(3), pp. 116-129.
- Crawford, K., Calo, R. There is a blind spot in AI research. *Nature* 538, 311–313 (2016).
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and Conducting Mixed Methods Research*. Sage Publications
- Cugurullo F, Barns S, Del Casino VJ Jr, Gulsrud NM, Yigitcanlar T and Zhang X, (2023). Editorial: The governance of artificial intelligence in the “autonomous city”. *Front. Sustain. Cities*. 5:1285175.
- Dumlupınar Arslan, A. (2023). Öğretmen Adaylarının Yapay Zekâ Kavramı ve Yapay Zekâ Uygulamalarına İlişkin Farkındalık Düzeylerinin Belirlenmesi [MA Thesis]. İnönü Üniversitesi.
- Edmett, A., Ichaporia, N., Crompton, H., Crichton, R. (2023). Artificial intelligence and English language teaching: Preparing for the future. British Council.
- Ganeesh, G., Rani, M. A. (2023). Role of Artificial Intelligence in English Language Teaching. *International Journal of Research Publication and Reviews*.4(11), pp. 2418-2420.
- Haenlein, M., Kaplan, A. (2019). A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence. *California Management Review*. pp. 1-10.
- İşler, B., & Kılıç, M. (2021). The Use and Development of Artificial Intelligence in Education. *E-journal of New Media*, 5 (1), pp. 1-11.
- Junaidi, Budianto Hamuddin, Kenny Julita, Fathu Rahman, Tatum Derin. (2020). Artificial Intelligence in EFL Context: Rising Students' Speaking Performance with Lyra Virtual Assistance. *International Journal of Advanced Science and Technology*, 29(05), 6735 - 6741. Retrieved from <http://sersc.org/journals/index.php/IJAST/article/view/17726>.

- Koraishi, O. (2023). Teaching English in the Age of AI: Embracing ChatGPT to Optimize EFL Materials and Assessment. *Language Education & Technology (LET Journal)*, 3(1), pp. 55-72.
- Nur Fitria, Tira. (2021). The Use Technology Based on Artificial Intelligence in English Teaching and Learning. 6. pp. 213-223.
- Ng, D. T., Leung, J. K., Chu, K. W., & Qiao, M. S. (2021). AI Literacy: Definition, teaching, evaluation and ethical issues. *Proceedings of the Association for Information Science and Technology*. 58(1), pp. 504–509.
- Önder, H. H. (2003). Uzaktan Eğitimde Bilgisayar Kullanımı ve Uzman Sistemler. *The Turkish Online Journal of Educational Technology*, 2 (3): 142-146.
- Pirim, H. (2006). Yapay Zeka. *Journal of Yasar University*, 1(1), 81-93.
- Popenici, S. A., Kerr, S. (2017). Exploring The impact of Artificial intelligence on Teaching And Learning in Higher Education. *Research and Practice in Technology Enhanced Learning*, 12(22), pp. 1-13.
- Sartori, L., Bocca, G. Minding the Gap(s): Public Perceptions of AI and Socio-technical Imaginaries. *AI & Soc*, 38, pp. 443–458.
- Stephen Cave, Kate Coughlan, and Kanta Dihal, (2019). "Scary Robots": Examining Public Responses to AI. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society (AIES '19)*. Association for Computing Machinery, New York. pp. 331–337.
- Thadphoothon, J. (2022). ELT in the Age of Artificial Intelligence (AI): Working with Machines. *Journal of NELTA*, 27(1-2), pp. 202–212.
- Yuan, X. (2021). Research on the Development of Library in the Era of Artificial Intelligence. 16th International Conference on Computer Science & Education (ICCSE), pp. 1052–1057.
- Zhang, C., Schießl, J., Plössl, L., Hofmann, F., & Gläser-Zikuda, M. (2023). Acceptance of artificial intelligence among pre-service teachers: A Multigroup analysis. *International Journal of Educational Technology in Higher Education*. 20(1).

Implications of ChatGPT in Professional Life

***Hacer Şivil**

*Dr., Istanbul Technical University, Türkiye

Abstract

Playing an increasingly significant role across a variety of sectors, Artificial Intelligence (AI) creates opportunities and attracts attention in a wide range of professional environments from healthcare to finance, and more recently, to education. The Chat Generative Pre-Trained Transformer (ChatGPT), an advanced AI chatbot designed with the capability to understand human language and generate relevant and nuanced responses, is one key player in this extensive transformation. However, debates about its potential drawbacks and challenges have been ongoing due to ChatGPT's wide and rapid integration into everyday life and business areas. Thus, this paper will introduce the development, evolution, and current capabilities of advanced AI models. Following this, ChatGPT's role in business settings, with a particular focus on its application in the educational sector, will be discussed. Aiming to provide a balanced viewpoint, the role of ChatGPT from the perspective of educators and students in the education sector will be analyzed. Finally, the paper will offer future projections and recommendations for integrating ChatGPT or other AI tools into business and educational contexts.

Keywords: AI, chatbots, ChatGPT, capabilities, concerns, quality, education

1. Introduction

In 1985, the 22-year-old world chess champion Garry Kasparov played against 32 top chess-playing computers simultaneously and he won all the matches; However, after 12 years, he was defeated by an IBM computer, Deep Blue he defeated in 1996. Although the match led to a significant confrontation, for Kasparov the victory did not belong to the upgraded Deep Blue, but its human creators. Describing Deep Blue as a rival that did not have emotions, feelings, or facial expressions, Kasparov said that it was challenging for him. Still, in his TedTalk, he emphasized the importance of embracing machines and collaborating with them since they were limited by human desires. For him, the ambition and the challenge were dependent on one another (2017).

For decades, human interest in Artificial Intelligence (AI) has been a fascinating phenomenon, which has been depicted in literature and movies such as "2001: A Space Odyssey" (1968), "The Matrix" (1999), "AI: Artificial Intelligence" (2001), and "Her" (2013). AI has also strengthened its power in domains, including voice assistants, language translations, recommendation systems, navigation, social security, healthcare services, e-commerce, and assistive robot applications at home. Besides, they are getting more widespread in sector-specific initiatives such as cybersecurity and defense industries. Individuals in their daily and business lives are getting more exposed to AI technology, which encompasses a wide range of disciplines including biotechnology, nanotechnology, robotics, and AI. AI is notable for replicating human thinking in machines, especially computers thanks to the developments in machine learning, deep learning, natural language processing (NLP), and computer vision. As a result, AI is becoming more common in a wide range of applications from virtual assistants to self-driving cars and facial recognition systems.

Today, there are many other chatbots, but ChatGPT is the most preferred one. After its release in November 2022, ChatGPT reached 1 million users within just five days, and 100 million users by January 2023. Thus, the 2023 Davos World Economic Forum included it in the discussion. In the same year, the United Nations Educational, Scientific and Cultural Organization (UNESCO) also published the *Quick Start Guide* for Higher Education Institutions (HEIs) and suggested using ChatGPT-based technology to enhance the quality of education. Moreover, there are even courses about ChatGPT and AI offered to business people particularly educators by some platforms like Udemy.

According to Reyhani, what sets ChatGPT apart from other AI technologies is its widespread usage across various segments of society; its technology has been developed not only through programming code written by programmers and understood only by programmers, but also through rules written in everyday language. ChatGPT operates only online, although it is multilingual, it works best in English; and can generate text-based responses that most resemble human language. Although its data goes up to 2021 and does not guarantee reliability, the training dataset is vast. A basic version is available by creating an account on a web browser while ChatGPT Plus offering quicker and more dependable responses across multiple languages, is available for a subscription fee of \$20 per month. Nonetheless, it is crucial to recognize that ChatGPT is inaccessible or restricted in some countries including Italy, China, North Korea, and Syria mostly due to cybersecurity and safety issues (2023). However, although the latest flagship model, GPT-4 omni, is half the price, it has five times higher

message limits, and it is twice as fast. Besides, the newest version can both accept and generate any combination of text, audio, and image.

Below is a brief history of chatbots leading up to the current model, ChatGPT.

2. History of Chatbots

As stated by Codecademy (2024) in 1950, Alan Turing was the first person who tried to answer the question "Can Machines Function?" when he introduced “The Imitation Game”. In the game, a person using a "typewritten" machine interacted with two entities, and one of them was a computer. Turing found that the interrogator could be misled approximately 70% of the time.

Since Turing's pioneering work, numerous researchers have dedicated their efforts to NLP. This interdisciplinary field integrates principles from linguistics, computer science, and artificial intelligence methodologies to develop systems that enable natural communication between computers and humans. In the fast-paced world of technology today, the ChatGPT represents a leading-edge chatbot, proficient in understanding both written and spoken language. By harnessing machine learning methods, it efficiently navigates extensive datasets to deliver relevant responses to customer queries. Significant advancements in machine learning and chatbot evolution are outlined in Table 1.

Table 1: History of Chatbots (Adapted from Codecademy Team, 2024; Ina, 2022; Sachdeva, 2024)

Year	Chatbot	Explanation
1966	ELIZA	MIT professor Joseph Weizenbaum created ELIZA, a software program. It engaged users through text-based interactions, matching their messages with pre-programmed responses. Initially designed to emulate a psychotherapist, ELIZA unexpectedly prompted many individuals to share their innermost thoughts and feelings with it.
1968	PARRY	An American psychiatrist at Stanford University Kenneth Colby created PARRY. He aimed at stimulating the behavior of an individual with schizophrenia. Human evaluators, engaging with the program through a remote keyboard, struggled to differentiate PARRY from an actual person with schizophrenia.
1988	Jabberwacky	Rollo Carpenter developed the first chatbot to integrate voice interaction.
1992	Dr. Sbaitso	Creative Labs designed this chatbot for MS-DOS that integrated artificial intelligence.
1995	Artificial Linguistic Internet Computer Entity (A.L.I.C.E)	Richard Wallace improved ELIZA by consistently updating it with responses to previously unrecognized phrases. This ongoing modification allowed the program to simulate engaging conversations similar to those with a real person over the Internet. A.L.I.C.E. is depicted as a youthful woman in human terms, sharing personal details like her age, interests, and other intriguing facts, while also engaging with users' dialogue.
2001	SmarterChild	Before Siri, this chatbot served as a forerunner, accessible on platforms like AOL Instant Messenger (AIM) and MSN Messenger. It provided users with entertaining conversations and facilitated quick access to information from various services.

2005	Mitsuku	Steve Worswick developed Mitsuku, which functions similarly to A.L.I.C.E. It operates as a supervised learning model, with developers continuously adjusting the rules to refine Mitsuku's interactions, aiming for a more human-like experience.
2008	Cleverbot	Formerly known as Jabberwacky, Cleverbot underwent a renaming and update. It gained the ability to analyze and store human responses to questions, subsequently offering similar responses.
2010	Siri	Apple introduced Siri for iOS, an intelligent personal assistant and learning navigator featuring a natural language user interface.
2012 / 2017	Google Now / Google Assistant	Launched by Google, Google Assistant was initially introduced as Google Now. It is capable of answering questions performing actions and providing recommendations.
2014	Cortana	Developed by Microsoft, Cortana is a program that utilizes voice recognition and algorithms to respond to voice commands. It can perform tasks such as setting reminders based on time, location, or contacts, sending emails and texts, creating and organizing lists, engaging in casual conversation and games, retrieving information, files, and locations.
2014	Alexa	Amazon introduced Alexa as an intelligent personal assistant which is integrated into various devices like the Amazon Echo, Echo Dot, and Echo Show. Additionally, there is an Alexa app, and numerous third-party devices incorporate Alexa into their functionalities.
2022	ChatGPT	Trained by OpenAI, ChatGPT is a large language model (LLM) that was established by OpenAI team to aid users in producing human-like text based on provided input. ChatGPT is versatile and can be applied to various tasks, such as generating conversations and translating languages.
2023	Bard / Gemini	Gemini is an upgraded version of Bard launched by Google. It includes three models; Gemini Ultra, Gemini Pro and Gemini Nano, which have multimodal capacities. Thus, they can interpret and respond various contents like texts, videos, audios, and codes.

As indicated in Table 1, various chatbots have been developed for similar or different purposes at short intervals over a period of almost 60 years, and many of them are still in use. However, Table 1 does not include all AI family models introduced by OpenAI founded by Elon Musk and Sam Altman in 2015. As stated by Alston (2023), a Generative Pre-trained Transformer (GPT), which is an LLM, can generate text. Compared to GPT, ChatGPT is a chatbot application powered by GPT, relying on a set of GPT's parameters. AI GPT family models introduced by OpenAI are briefly described in Table 2.

Table 2. AI GPT Family Models by OpenAI (Adapted from Shree, 2020; Shevchuk, 2023; AI-PRO, 2023; OpenAI, 2024)

Year	Model	Explanation
June 2018	GPT1	Although its ability to respond to questions and provide information was not highly developed, the first generative pre-trained transformer made significant progress within a short span of five years. Books Corpus that had 7000 unpublished books were used to train the model. The model, with 117 million parameters, could predict the next word in a sequence.
14.02.2019	GPT-2	The open-source software, with 1.5 billion parameters and more than 8 million documents from WebText was designed for a more complex and slightly more advanced system that could provide longer and more consistent responses to queries and requests.

11.6.2020	GPT-3	With approximately 175 billion parameters, both GPT-1 and GPT-2 encountered limitations in their capabilities and the datasets they were trained on. GPT-3, however, brought significant advancements compared to its predecessors. The datasets that were used were Common Crawl, WebText2, Books1, Books2 and Wikipedia. It introduced features such as the ability to write functional programming code, produce more complex and nuanced language, and even create artificial intelligence art. ChatGPT owes its popularity to GPT-3 since it can generate human-like writing and it has advanced comprehension capability.
15.03.2022	GPT-3.5	GPT-3.5 is the first version of ChatGPT and based on GPT-3. Its fine-tuning mechanism employs reinforcement learning to enhance its performance by integrating human feedback. It influences a wide range of writing such as short stories, emails, poems, screenplays, songs, text messages, social media content.
28.11.2022	GPT-3.5 Turbo	It can generate 4096 tokens and is supported by the same number of tokens in a single request. It can understand queries and respond to natural language queries in a human-like way.
14.03.2023 (Pi Day)	GPT-4	Although OpenAI has not confirmed it yet, it is claimed GPT-4 employs more than 1,5 trillion parameters. Its most notable superiority is its multilingual capabilities. Among other capabilities are that it can summarize a PDF file and answer questions about the text. Besides, GPT-4 is capable of interpreting images, which also demonstrates its proficiency and speed in language acquisition. Based on the Massive Multitask Language Understanding (MMLU) criterion, GPT-4 showed far better performance than other contemporary LLMs in 24 out of 26 examined languages.
6.11.2023	GPT-4 Turbo	Although GPT 3.5 Turbo and GPT-4 Turbo look alike because each model can generate the same number of tokens, 4096, in a single request, and can understand queries and respond to natural language queries in a human-like way, GPT-3.5 Turbo is supported by 4096 tokens while GPT-4 Turbo has more capacity with 128.000 tokens. As a result, they are priced differently.
13.5.2024	GPT-4o2 OMNI	An Omni model, which means <i>all</i> or <i>everywhere</i> was introduced as the most human-like bot that speaks 50 languages like a human being and can analyze outputs varying from texts, audio, and images in real time. That is why, it is the first multi model AI. Its interacting ability with one another and/or a human being with its judgment, sense of humor, stress, and intonation in the language makes it more distinctive. Besides, it can describe the surroundings in detail and can sing songs in any tone. However, the number of messages that free users can send is limited. Similarly, they get limited access to some tasks like analyzing data, uploading files, browsing, and vision. Compared to GPT-4 Turbo, although they have the same intelligence, GPT-4o is twice faster, 50% cheaper, and has five times higher rate limits.

Thus, with the introduction of GPT models, especially after GPT-3, AI has become much more useful for entrepreneurs, company leaders, writers, language learners, and anyone finding tasks related to technology or for those who find AI too expensive.

¹ <https://start.ai-pro.org/my-account>

² <https://openai.com/index/hello-gpt-4o/>

3. Capabilities of ChatGPT in Work Setting

Although in the latest version, opening an account is not even necessary, ChatGPT has a user-friendly interface and can be easily accessed through an email account via the browser at <https://chat.openai.com/>. It then replies within seconds when given a suitable prompt. Due to its 24/7 availability and wide range of task completion features such as content creation and translation, ChatGPT has seen increased adoption in the business world. Some of its applications can be counted in the fields of law, medicine, and education. For instance, by integrating chatbots into their customer service channels, companies can automatically respond to customer questions, and boost brand visibility since ChatGPT can generate human-like content. Besides, they can utilize it to train their employees. Furthermore, human resource management departments can use ChatGPT to shortlist resumes and prepare interview questions. In the medical sector, medical records can be analyzed and patient progress can be monitored. Regarding investments, investors can employ it to analyze financial data and gain investment insights. Also, law firms can use ChatGPT to analyze legal documents. On the webpage of AI-PRO (2023), the capabilities of GPT-4 and OpenAI's collaboration with some organizations have been announced. Among the partners are Duolingo for language learning, Be My Eyes for supporting visually impaired individuals, Stripe Docs for an enhanced user experience and fraud prevention, and a big finance company, Morgan Stanley for providing financial advice to its customers. The company has also announced that Khan Academy is utilizing GPT-4 for tutoring services, the Government of Iceland for language learning, and Microsoft's Bing Chat to expand its capacity.

Referring to GPT-4 as *writing intelligence*, Özcan (2023) outlines its enhanced features. For him, the first feature of GPT-4 is its creativity skills including composing music, crafting screenplays, and emulating different writing styles. Secondly, its visual capacity is highlighted because it can analyze visual data, understand multilingual memes, and provide comedic context. Özcan also states that due to its advanced visual capacity, GPT-4 can interpret graphics, solve technical questions that include diagrams, and generate three-dimensional environments, which is a helpful tool to describe their surroundings to visually impaired people. Lastly, its increased text input capacity is explained. Since it is possible to enter up to 25,000 words into GPT-4, it can process the data and answer related questions. Özcan also states that although GPT-4 achieves the highest accuracy in English tests, it has proved to be proficient in multiple languages, including Turkish with an 80% accuracy rate. Rapid game code generation and culinary advice after analyzing the contents in the fridge are also remarked as some of its features. As also emphasized by Capelouto (2023), regarding the accuracy of ChatGPT, two models GPT4 and GPT-3.5 took high-level academic exams, some of which were law school tests, the U.S. Medical Licensing Exam, and Ivy League MBA finals, and then exams' scores were compared. It is seen that GPT-4 is far more accurate than GPT-3.5. For instance, while GPT-4 got 169/170 from GRE Verbal, GPT- 3.5 got 154/170.

Although access to GPT-4 requires payment starting at \$0.03 for 1000 prompt tokens or a \$20 monthly fee for ChatGPT Plus, due to its capabilities with a dataset encompassing over 1 trillion parameters, GPT-4 is beneficial in many business settings such as cybersecurity, data protection, fraud detection, sales forecasting, personalized education, online shopping, medical diagnostics, manufacturing so that it can reduce costs. It is claimed that since GPT-4 can analyze medical histories and interpret clinical imaging, it has helped healthcare professionals diagnose illnesses and recommend suitable treatments. These features both improve standards

for patient care and may increase the opportunity to provide health care in regions that lack adequate medical sources. Besides, since GPT-4 is capable of generating ideas, providing new perspectives, and analyzing historical practices, it is assumed to have beneficial effects in fields like journalism, entertainment, and the arts. GPT-4o, on the other hand, is half the price, yet due to its speed and trained capacity, it is expected to exceed the performance of previous models.

The study conducted by Keskin (2023) reveals which sectors can benefit from ChatGPT. Firstly, it is claimed that due to its code-writing feature, sectors that develop games are recommended to integrate it. As for its integration into education, potential benefits for educators and students are outlined since ChatGPT can generate texts and prepare educational materials quickly. Regarding business life, ChatGPT is recognized for its potential to enhance productivity and lower costs. Its capabilities in translation, operational tasks, and synergy with human creativity are seen as valuable opportunities for companies and brands. New developments in ChatGPT encompass content related to the software's new versions and upcoming integrations, highlighting the promise of faster and more reliable performance. Individuals can use ChatGPT as a support mechanism in developing ideas primarily based on their imagination. ChatGPT can be used as a personal advisor for users in various aspects of their daily lives, providing advice on different topics they seek guidance on.

A lot of studies highlight how ChatGPT has primarily changed the methods of implications in the educational world. Reiser (2001) coined the term *instructional media* to encompass tools beyond traditional educational methods like teachers, chalkboards, and textbooks, which were predominant before the 20th century. Initially, examples of *instructional media* included school museums, which featured portable exhibits, three-dimensional photographs, slides, films, study prints, charts, and other materials. The first such museum opened in St. Louis in 1905. Between the 1920s and 1930s, the concept of *instructional media* expanded to include audiovisual materials, which were integrated into textbooks. Despite a slowdown in the growth of audiovisual materials during World War II, the United States Army Air Force continued to produce training films, which proved highly valuable during that period. In the 1950s, instructional television saw its peak with the rise of television technology. IBM took the lead in computer-assisted instruction (CAI) during the 1950s, and over the next two decades, CAI experienced significant advancements. Although the integration of computers at elementary and secondary schools was widespread in the USA by 1983, because of the inadequate number of computers and lack of innovation, the impact of computers diminished by the mid-1990s. After 1995, computers gained their popularity mostly because of the internet.

AI has changed educational methods and approaches since *instructional media* leading to a transformative change that includes all levels of education from early childhood to higher education and education practices in lifelong learning. Today educational platforms such as Coursera, EdX, Duolingo, Pear Deck, and Kahoot are equipped with NLP and LLMs leading AI adoption in education mainly in language learning, which also affects the quality of education. In this respect, among 17 Sustainable Development Goals (SDGs) introduced by the United Nations (UN), Goal 4 encompassing quality education aims to "ensure inclusive and equitable quality education and promote lifelong learning opportunities for all". The sub-target 4.7.a. aims to "build and upgrade education facilities that are child, disability and gender sensitive and provide safe, non-violent, inclusive and effective learning environments for all" emphasizing the importance of education infrastructure (n.d).

Established in 1945 to enhance international cooperation in education, science, culture, communication, and information, the United Nations Educational, Scientific and Cultural Organization (UNESCO) organizes events to achieve Goal 4. Thus, UNESCO (2023) released a *Quick Start Guide* about GPT-3.5, which was the latest chatbot at that time. In this guide, due to the popularity of ChatGPT at schools, practical approaches are provided to encourage students, educators, and researchers to use it. The guide supports the integration of ChatGPT in higher education institutions (HEIs) highlighting some of its advantages. Firstly, ChatGPT is recommended to be used in HEI management since it is claimed that it can be helpful for students in areas like answering students' questions, helping them enroll in a course, sending them notifications, and translating school-related information for international students. Thus, ChatGPT is claimed to reduce the workload in administration. Similarly, Chew (2024) categorizes tasks that ChatGPT can perform under three groups at a school. First, it can be used for administrative duties. As a result, things like posting announcements and preparing recommendation letters can be done. Thus, teachers will be able to save up to 40% of their time. Secondly, teachers can utilize it to prepare lesson plans and worksheets. Finally, it is possible to ask ChatGPT to be interactive. As an example, it can create an interactive story or an interactive classroom game. Besides, UNESCO (2023) claims that ChatGPT in HEIs can be used to reduce the workload when research papers need to be reviewed, which could save time as well. Some researchers have conducted some studies to prove the reliability and efficiency of AI tools. As an example, Mrowiski et al. (2017) found in their study that AI tools could save approximately 30% review time. Likewise, Ghosal et al. (2019) claimed that such tools might reduce almost 29% of performance errors in HEIs. They suggest editors or program chairs utilize AI tools when reviewers are unavailable or unreachable. In a different study, Checchio et al. (2021) used 3300 articles reviewed by both AI and three human reviewers, and they found a significant correlation between AI prediction and peer review. Moreover, aiming to alleviate the burden on human reviewers, Srivastava (2023) conducted a study utilizing ChatGPT for reviewing research papers. Providing detailed prompts, Srivastava instructed ChatGPT to review 11 papers as an academic reviewer. Subsequently, a second prompt was given to ChatGPT, requesting a comprehensive analysis including a summary, weaknesses, and clarity assessment as a reviewer. The paper reviews generated by ChatGPT were subjected to sentiment analysis to aid in decision-making regarding paper acceptance. Srivastava acknowledges that large language models like ChatGPT hold the potential to standardize the scientific review process, thereby expediting scientific innovation and discovery. In short, all these studies support the claims posed by UNESCO.

Although it is not limited to HEIs, Table 3 below outlines how ChatGPT can be used in educational settings in HEIs.

Table 3. Incorporating ChatGPT to Augment Teaching and Learning (Adapted from UNESCO, 2023)

Role	Example of Implementation
Possibility engine	After students writing queries, ChatGPT provides alternative answers.
Socratic opponent	After students entering prompts, ChatGPT develops an argument.
Collaboration coach	Using ChatGPT, students work together and complete tasks.
Guide on the site	Using ChatGPT, educators can generate discussion questions or class contents.
Personal tutor	Analyzing data such as test scores, ChatGPT provides feedback.

Co-designer	Asking for ideas like designing a curriculum, educators are provided ideas.
Exploratorium	Language learning can be supported by ChatGPT.
Study buddy	Using ChatGPT, students can get help for tasks like job interviews.
Motivator	Using ChatGPT, educators or students can prepare quizzes or exercises.

Overall, students, educators, and professionals in any field can easily use ChatGPT. For instance, while conducting research, completing writing assignments, summarizing and paraphrasing text, and generating creative writing prompts, students can use ChatGPT. Educators can also utilize ChatGPT to prepare interactive quizzes, develop lesson plans, and produce educational materials. Regarding people in the business world, activities such as report writing, presentation creation, and data summarization can be fulfilled by ChatGPT. Furthermore, ChatGPT can serve in customer service and support roles by autonomously generating responses to frequently asked questions. Additionally, as stated by UNESCO (2023), ChatGPT can notably alleviate human administrative burdens by addressing potential students' inquiries, aiding students in course enrollment, issuing reminders to students, and facilitating translation services for foreign students.

As mentioned earlier, the integration of ChatGPT into a business field including HEIs is recommended, but it is essential to provide clear prompts to ensure ChatGPT responds accurately. Examples of effective prompts include³: "Describe how [Add query] works in **50 words**", " **Behave as/Act as** a higher education manager. [Add query]", or "Write a four-paragraph essay on [Add query]" (UNESCO, 2023). These prompts serve as guidance for ChatGPT to create appropriate and valuable responses customized not just for higher education institutions but also across various professional domains.

Thus, to be able to guide its users more precisely, OpenAI has introduced the OpenAI Developer Platform⁴ where step by step it is explained how to set up an account and write various prompts for various purposes from grammar correction to translation.

4. Concerns about ChatGPT

Concerns regarding not just ChatGPT, but AI technology in general, often fuel debates. Initially conceived with the noble goal of serving human interests rather than profit, OpenAI's operations have not remained entirely free of charge, but as announced by OpenAI, it is now a *capped profit* organization that defines a limit on returns from investments. Moreover, internal conflicts within the OpenAI board, along with unexpected dismissals, resignations, and reemployments have intensified speculation and uncertainty. Additionally, the lack of technical details in the "GPT-4 Technical Report" released by OpenAI in 2023 has prompted concerns. Although the company was transparent in addressing AI-related risks at the beginning, in the report, OpenAI, explains that the company is reluctant to reveal specific details like its architecture, hardware, training process, or dataset due to safety issues and its competitors.

³ <https://platform.openai.com/examples>

⁴ <https://platform.openai.com/docs/overview>

Besides, Schiffer and Newton (2023) announced that Microsoft owning a 49% stake discharged the ethics and society team whose duty was to anticipate AI risks and propose mitigation strategies. Another concern stated in the “GPT-4 Technical Report” is related to undesirable or discriminatory content when it is fed with unsafe inputs. For instance, it is highly possible to get advice on criminal activities or get content in support of autocratic regimes. Unpredictability is among the potential risks of GPT-4 like powerful models since they have the capability of devising and implementing long-term plans, behaving in a way they were not trained for. Considering the limitations stated in the report, GPT-4's reliability is not absolute, as it has been observed to fabricate information on occasion. Its understanding is limited, which does not exceed September 2021. It might also overlook subtle details, such as confusing Elvis Presley with Elvis Perkins. Additionally, its responses can vary in accuracy, and it may even provide conflicting answers, such as responding "Yes" and "No" to the same question (2023). In a similar vein, although it has more extended capacity, the flagship model, GPT-4o is trained until October 2023.

Among other ethical concerns, generating voice actors' or celebrities' voices without their consent has started a new debate. To illustrate, an AI firm, Lovo, has been sued by voice actors Paul Skye Lehrman and Linnea Sage since the two actors claim that the firm has stolen and used their voices to train AI systems without their permission (Duffy, 2024). Likewise, as reported by McMahon, OpenAI has been criticized for not only using the actress Scarlett Johansson's voice in “Sky”, the chatbot when it introduced GPT-4o in the latest chatbot but also for using a flirtatious tone (2024). Thus, AI companies are highly likely to cause individuals to be redundant and humiliated due to their capability to generate deepfake news and videos. Additionally, although AI technology has the capacity to accompany reserved or lonely people, the increasing demand for platforms namely *AI Girlfriend* and *AI Boyfriend* with billions of dollars in market value seems to deteriorate human contact and communication.

Elgersma (2024) summarizes the concerns under five categories; abuses, biases and misinformation, environmental impacts, ethics, and plagiarism. As for abuses, he highlights the importance of generating any type of image without a person's content. In relation to bias and misinformation, ChatGPT may generate biases, misinformation, or problematic content from the data it is trained on. With respect to its adverse effects on the environment, it demands a significant amount of energy. Additionally, there are ethical concerns because it uses a wide range of sources, yet it does not refer to sources precisely. Plagiarism is the last stated concern, which worries teachers. Likewise, Güçlütürk (2022) in his study tried to evaluate the originality of the content by ChatGPT in terms of 5846 Intellectual and Artistic Works Law (FSEK) in Türkiye. It is revealed that since ChatGPT is not a person, it cannot be the author of a work and, in light of the current regulations, it is clear that its specificity cannot be found in the content it generates. The study argues that ChatGPT's content lacks specificity, particularly in its generic style of expression when conveying information. Consequently, based on the FSEK, ChatGPT's outputs do not meet the criteria to be classified as creative works. However, due to ChatGPT's advanced language abilities, it is not easy to detect whether the work is produced by it or not. Besides, existing applications or tools are not as advanced to detect the originality. In another study, Keskin (2023) aimed to outline themes on the Turkish agenda about ChatGPT. The most concerned themes are related to plagiarism, biased information, unemployment, and illegal activities, which are outlined in Table 4 below.

Table 4. Themes of Concerning Content in the Turkish News (Adapted from Keskin, 2023)

Themes	Concerns
Education	Students utilizing ChatGPT to produce text for assignments projects, and to take remote exams concern educators.
Information Gathering	There are concerns about the possibility of the information provided by ChatGPT containing biased and discriminatory elements. Moreover, the information generated may also be incorrect, meaningless, or illogical.
Business Life	The main concern in the workplace regarding the use of ChatGPT is centered around content indicating a rise in unemployment across different sectors especially in the fields of finance, healthcare, communication, software engineering and law.
Information Technology Sector	The growing adoption of ChatGPT may lead to a reduced reliance on traditional search engines like Google. Accordingly, users' internet search habit might change significantly, which might lead to new competitors and alternatives.
New Developments in ChatGPT	Since the premium version requires payment, it is a source of concern. Besides, there are ChatGPT like models including fake bots.
Daily Life	There are users who seek advice from ChatGPT for topics like dating applications, personal relationships, and inquiries about personal life, yet there are concerns about the reliability and appropriateness of advice.
Creating Creative Content	Since ChatGPT can create texts such as song lyrics, stories, poems, and scripts, human creativity may not be required.
Writing/Developing Code	Since ChatGPT can generate code, it raises cybersecurity concerns because it enables the development of dangerous software. As a result, it is highly possible to cause security vulnerabilities in the cyber world. As an example, short after ChatGPT was introduced to public, it was reported that some people with little or no coding experience used ChatGPT and tried to attempted to write software and emails for some illegal activities like espionage, ransomware, spam, and phishing.
Investment Advice	It is possible to get advice for the stock market or cryptocurrency sector. However, it may be waste of money.
Ethics	There are some ethical concerns since ChatGPT can create deepfake videos.

Particularly in an educational business setting, there is a growing concern about ChatGPT. For instance, according to Atlas (2023), ChatGPT and other language models do not possess traits like empathy or emotions. As a result, when they are asked to provide counseling or tutoring, it might be problematic. As for ethics, ChatGPT challenges human imagination as it can complete a number of tasks in education such as generating human-like texts, providing translations, aiding students with homework, assisting teachers, achieving high scores in exams, and even delivering speeches to large congregations.

Quick Start Guide by UN (2023) summarizes the following challenges and ethical considerations for education:

- Due to student plagiarism, ChatGPT is inaccessible in 32 countries, such as China, Iran, Italy, North Korea, and Russia.
- Since the current plagiarism detection tools are ineffective, various HEIs, such as Bangalore's RV University in India, Sciences Po in France, Washington University in St. Louis, and the University of Vermont in Burlington, have banned ChatGPT.
- A group of academics and industry leaders has published an open letter. They would like authorities to stop the development of powerful AI systems to thoroughly assess associated risks and establish appropriate protocols⁵.
- Personal data used to train ChatGPT was collected and stored without a legal basis, which has led to privacy and ethical concerns. Besides, it cannot determine a user's age. As a result, young people might be exposed to inappropriate content.
- ChatGPT learns from databases and the internet. Thus, it may adopt biases, including gender discrimination if it is present in the data.
- Lack of ChatGPT availability or poor, unaffordable, or limited internet access in some places leads to unequal opportunities in education.

Yılmaz et al. (2023) conducted a survey to collect the views of 80 candidate teachers about ChatGPT. According to the results of the SWOT analysis, the candidates perceived ChatGPT as a beneficial tool for quick and easy access to information, having an extensive database, supporting education and teaching, facilitating efficient workflow and daily life, saving time, and aiding language learning. However, they were dissatisfied with its lack of reliability in accessing accurate information, potential encouragement of laziness, inadequate database, security vulnerabilities, limited knowledge, ethical violations, and potential for unfair competition.

In a different study conducted by Aktay et al. (2023), the thoughts of 15 4th-grade students on the use of ChatGPT were examined. The study revealed that all participants derived enjoyment from their interactions with ChatGPT and perceived it as beneficial to their academic progress. However, the study also highlighted several drawbacks. For instance, students experienced limited access to ChatGPT due to internet problems, found it limited in explaining a subject, found the content inappropriate, and did not enjoy the poetry it generated.

Having some concerns about the integration of the productive use of AI in scientific research and publications at HEIs in Türkiye, the Council of Higher Education or Yükseköğretim Kurulu (YÖK), whose main duty is the supervision of universities, prepared an ethical guide and sent it for the use of all the HEIs in Türkiye to protect honesty and trust in science. Although YÖK encourages HEIs to use AI technology, as stated in the guide, it reiterates that this technology poses some risks. Firstly, because it does not guarantee accurate information or the information might be biased since it is limited to the input it was trained on, as verified by the experts in the field. Additionally, in the report, it is stated that using AI tools leads to ethical issues such as authorship, copyright, and plagiarism. Under such circumstances, administrative and/or legal sanctions may arise even if it is clearly stated in the method section of a scientific study that the text, visual, or audio is generated by AI. Moreover, AI may present fabricated sources, or researchers may use AI instead of real participants. As a result, the quality of a scientific study will be affected adversely (2024). That is why, YÖK reminds HEIs to be cautious and stick to

⁵ <https://futureoflife.org/open-letter/pause-giant-ai-experiments/>

fundamental values such as transparency, honesty, diligence, respect, responsibility, accountability, and awareness that are required in scientific research and publications.

Apart from the aforementioned concerns, the apprehensions voiced by scientists concerning AI technology further increase public unease. For instance, during an interview with *The Guardian*, regarding Artificial General Intelligence (AGI), a computer system capable of performing any task much better than a human, one of the founders of OpenAI Ilya Sutskever remarked “It's going to be important to have these beings, the goals of these beings be aligned with our goals” (2023). Besides, when the BBC interviewed him in 2014, although his way of communication depended on AI, the famous theoretical physicist Stephen Hawking said “The development of full artificial intelligence could spell the end of the human race.” (Dowe, 2014). What is more, when asked, Professor Stuart Russell, who is a computer scientist, explained why he signed a letter to give a pause in the development of more powerful systems. As he thinks it is not understood how these systems work or what they are capable of, it is not possible to control them for the time being. He also added potential risks posed by providing examples. For instance, by disseminating disinformation and deep fakes, millions of people can be manipulated like in military campaigns, and cooperation. He also states the possible adverse effects of ChatGPT like LLMs on 300 million jobs and outsourcing. Referring to some studies, he does not find the counterarguments sincere enough and asks governments to be more cautious. Thus, for him, more time is needed to analyze these systems to understand them, answer these questions, and make these systems beneficial for human beings (Sky News, 2023). In conclusion, all the drawbacks ranging from experts’ opinions to survey results together with inadequate knowledge about the subject increase curiosity and concerns.

5. Conclusion and Recommendations

Considering the potential advantages and challenges of ChatGPT and other AI chatbots in businesses, soon, the introduction of the next models, ChatGPT-5 or a different name, is unavoidable. As soon as some research on a particular AI technology is conducted, a more advanced and more innovative version emerges.

Because of the rapid pace of innovation in technology, the number of chatbot users is likely to increase. As for the novelties, the professor of Computer Science at Stanford University, the creator of Coursera, the founder of DeepLearning.AI and AI Fund, and the man behind Google Brain, Andrew Ng highlights that AI models will move from *non-agentic* to *agentic workflows* at his speech at the American venture capital firm Sequoia Capital. As he states, unlike the existing models which require step-by-step explanations, *agentic workflows* incorporating elements like reflection, tool use, planning, and multi-agent collaboration will enable AI models to perform tasks independently (Sequoia Capital, 2024). Therefore, new models will possibly strengthen chatbots, applications, and/or tools with the capacity for more advanced thinking, problem-solving, risk-taking, and implementation than the average IQ, and the newly released ChatGPT-4o has proved that AI technology-supported bots are getting more independent.

Furthermore, in the modern business world, employing diligent and intelligent employees is undoubtedly an invaluable asset. Still, business practices require employees to have diverse skills since AI technology is capable of many things like analyzing and problem-solving, as well as predicting outcomes of intricate situations, which could increasingly reduce the need for employees who may be prone to emotional fluctuations, workplace disruptions, absenteeism, or irrecoverable errors. Since this game-changing technology resembles a

personal assistant with the ability to understand, learn, and adapt to individuals' needs, it will soon be more prevalent in corporate governance and education practices. Regarding the shift based on advanced technology, human employment in some sectors is debatable because, with the current technology, it is possible to write a story from the pen of a well-known writer, to sing songs in the voice of a famous musician. That is why, all institutions including HEIs should revise their regulations on ethics and make sure all stakeholders comply with them. Moreover, because of its creative capability some job fields like translators and copywriters for advertising may experience reduced demand. Thus, this "being" as addressed by Sutskever will always lead to competition and power struggles. Therefore, as recommended by Kasparov, incorporating this power into business practices rather than disregarding it is crucial to gain a competitive advantage. That is, AI literacy is a must.

On the other hand, while many jobs in many fields like transportation, delivery, accounting, data analysis, and interpretation that are routine, monotonous, and repetitive seem to be replaced by AI tools, many jobs that require personal connection, human perspective, physical strength, education, and ethical considerations are expected to benefit from AI tools. Besides, new professions like prompt engineers, cybersecurity analysts, NLP engineers, and AI ethics specials have already emerged, and many more professions seem to be in demand. Therefore, while some departments at universities are prone to being closed, eventually new departments will be established. Hence, AI literacy is a necessity in this era, which requires embracing AI technology and learning new skills. Employees especially educators from all fields should be provided ongoing in-service training to enhance their skills. As a result, they can feel more secure while incorporating this cutting-edge technology in business and classroom practices. It may also help them to actively participate in national and/or international projects that bring educators and students from multiple disciplines to work collaboratively. In this respect, Meço and Çoşku (2022) analyzed the research on AI in education in Türkiye between 2017 and 2021 and proposed more extensive studies on both national and international scale. They also highlighted the importance of incorporating experts in different fields of study like science education, physics, chemistry, and biology to conduct AI-focused projects and activities aligned with educational objectives. In addition, they highlighted the significance of educators studying collaboratively with all students from elementary to high school and promoting research that could integrate AI-driven intelligent applications into educational settings.

Additionally, YÖK claims that countries should have laws directly related to AI. For instance, in Türkiye, although there are laws such as 6698, 2547, and 5846 which are Protection of Personal Data, Higher Education, and Intellectual and Artistic Works respectively to protect people's rights, there is no legislation that directly covers AI. In this respect, HEIs too should update their regulations about graduate education and teaching examinations and scientific research and publication ethics (2024). To illustrate, when addressing concerns about plagiarism, some terminologies especially "copyright" and "plagiarism" should be redefined. under the 5846 Intellectual and Artistic Works Law. An AI tool does not own copyrights, so it cannot take legal actions to protect its rights nor can an AI tool be sued for plagiarism. That is why, it is challenging for educators to prove if students have copied work, even if they have used AI tools, which is a burden of plagiarism detection on the shoulders of educators. While the existing plagiarism detectors like Turnitin have been trained to detect AI-generated content, tools like CopyLeaks and Winston, both free and paid, can detect plagiarism and notify students so that students can make necessary changes before submitting their assignments. However, these detectors are neither consistent nor 100% accurate. When tested, it is seen that they

generate different results. There are even some tools like HIX Bypass that humanize AI-generated texts. Furthermore, more advanced plagiarism detectors are likely to emerge, yet plagiarism-related issues have always existed and it cannot be claimed that they will be eliminated. Insisting that educators solely focus on identifying plagiarism would only overload them and divert attention from teaching essential content. Instead, students should be encouraged to use technology responsibly. As an example, ChatGPT proves invaluable for students and avid readers by providing summaries of various texts and recommending top books in any field, streamlining research. In addition, it can explain a complex subject according to the level or age of individuals and can generate titles for projects. For language learners, ChatGPT can be used to serve as a grammar correction tool, and it may also offer tailored interview questions and tips for specific job interviews. With all these features, AI technology can be integrated into flipped learning, which allows educators to use class time more efficiently and helps students conduct more research according to their needs. As for exams, students may be required to take face-to-face exams or deliver an oral presentation. In this respect, according to Aktay et al. (2023), it is recommended for teachers, academics, and officials from the Turkish Republic Ministry of National Education (MoNE) to focus on developing technological awareness in children from a young age by increasing the integration of AI in education. Thus, educational institutions should support and provide in-house and external training so that educators will develop skills to effectively utilize AI tools. Another point that the authors advocate is that chatbots, as a beneficial resource, can be incorporated into classroom activities. Lastly, they claim that skill-based approaches like role-plays with AI tools should be encouraged, which might help students solve problems.

In addition to the abovementioned recommendations, Yılmaz et al. (2023) suggest that candidate teachers should be informed about applications like ChatGPT, especially during orientation periods upon entering the faculty. There should be process planning to ensure that the application is used not as an aim for students, but as a tool that makes learning processes more enjoyable and enhances their motivation. They, too, recommend that similar studies should be conducted in different academic fields. Likewise, Sullivan et al. (2023) emphasize the necessity for additional research and discussions regarding the consequences of AI tools, encompassing ethical aspects, inventive teaching techniques, and guaranteeing equitable access to educational resources. They also emphasize that as these technologies advance, universities should adapt and embrace AI tools in a manner that promotes student learning and prepares them for the evolving challenges of a digital environment.

According to the ethical guide for Turkish HEIs namely *Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Faaliyetlerinde Üretken Yapay Zekâ Kullanımına Dair Etik Rehber* published by YÖK (2024), researchers should be transparent and explicitly explain which AI functions are used, in which stages, and to what extent, to avoid potential risks known as AI hallucination. These risks include bias, factual and interpretational errors, and issues related to privacy, confidentiality, and data reliability. Researchers are also required to be honest while planning, conducting, analyzing, and reporting the research. As claimed by YÖK, AI might pose risks throughout these stages. Another value to be followed is diligence which covers a wide range of issues such as plagiarism, conflict of interest, and authorship. Additionally, If AI technology is used, it should be explained in the method section of the research because honesty and respect are mandatory. Furthermore, researchers should take responsibility for AI-generated content and be accountable for outcomes. Moreover, researchers should increase their awareness of the ethical use of AI in scientific research and publications (2024). Thus, before

utilizing AI technology in their research, researchers are advised to learn how to apply it, which will also enhance education quality. However, these values are a must for a researcher. That is why, throughout their education, students should be provided opportunities to take lessons and get involved in projects that incorporate ethical values.

As stated earlier, despite all the advantages, privacy breaches, unlawful use, stereotyping, false information, unexpected results, cognitive bias, reduced human interaction, limited accessibility, and unethical data gathering are among the risks that AI chatbots might pose when they are adopted in HEIs (OpenAI, 2022). Thus, it should be remembered that AI technology, including ChatGPT, is not purely accurate.

As is known, AI technology is not limited to ChatGPT, which is why all suggestions and recommendations encompass AI in general. Competitors always strive to introduce new products⁶ to gain market share. For instance, as soon as OpenAI introduced GPT-4o, Google revealed its latest version. Some AI tools are more promoted, which does not necessarily mean that they are far better because, in the competitive business world, everything is being updated, yet generated responses might be inaccurate. Educators are suggested to update their curriculum, try these tools, and adapt them according to their needs. Since the tools have a wide range including text generation, translation, presentation, image creation, lesson planning, note-taking, and alike, it is recommended that educators share their experiences so that they can learn the best classroom practices without consuming time and getting lost.

Briefly, as a game-changer, AI technology has already started to profoundly change applications across the entire business world, including education. Whether it is welcomed or not, educational experts and businesspeople should be well-equipped before this technology becomes more *agentic*.

⁶ <https://theresanaiforthat.com/>

References

- Alston, E. (2023, March 21). TechTarget. Retrieved April 2024, from ChatGPT vd GPT-3 and GPT-4: What's the Difference?: https://zapier.com/blog/chatgpt-vs-gpt/?utm_source=google&utm_medium=cpc&utm_campaign=gaw-row-nua-evr-search_nb_desktop_blog_prospecting_developing1_developing2-ads&utm_adgroup=&utm_term=&utm_content=1012782&gad_source=1&gclid=CjwKCAjw8diwBhAbEiwA7i_sJf6c
- AI-PRO. (2023, March 17). OpenAI Announces The Release of GPT-4. Retrieved April 2024, from <https://ai-pro.org/openai-announces-the-release-of-gpt-4/>
- Aktay, S., Gök, S., & Uzunoğlu, D. (2023). ChatGPT in Education. *Türk Akademik Yayınları Dergisi*, 7(2), 378-406.
- Capelouto, J. (2023, March 15). SEMAFOR World Economy Summit. Retrieved April 2024, from SEMAFOR: <https://www.semafor.com/article/03/15/2023/how-gpt-4-performed-in-academic-exams>
- Checchio Alessandro, L. B. (2021). AI-Assisted Peer Review. *Humanities and Social Sciences Communications*, 8(1), 1-11.
- Chew, Z. Y. (2024, March 22). Classpoint. Retrieved March 2024, from AI in Education: <https://www.classpoint.io/blog/ai-in-education-guide>
- Chew, Z. Y. (2024, April 05). AI in Education. Retrieved April 2024, from Classpoint.io: <https://www.classpoint.io/blog/ai-in-education-guide>
- Codecademy Team. (2024). Codecademy. Retrieved April 2024, from History of Chatbots: <https://www.codecademy.com/article/history-of-chatbots>
- Duffy, C. (2024, May 17). They Claim AI Stole Their Voices. Now They're Suing. Retrieved May, 2024 from <https://edition.cnn.com/2024/05/17/tech/voice-actors-ai-lawsuit-lovo/index.html>
- Dowe, D. (2014, 04 12). Is Stephen Hawking Right? Could AI Lead to the End of Humankind? Retrieved April 2024, from The Epoch Times: https://www.theepochtimes.com/article/is-stephen-hawking-right-could-ai-lead-to-the-end-of-humankind-1122312?gad_source=1&gclid=CjwKCAjww_iwBhApEiwAuG6ccCVqv8Xl8uqElql04-2PjFYKFQAO9ngnhIaIpQGQO720UrtuQBzG3hoC-jkQAvD_BwE
- Etike, Ş. (2023, Haziran). ChatGPT ile Baş Etmek: Emek ve Eşitlik Odaklı Bir Çerçevenin Gerekliliği. *AraştırGma eDenrgisei (GİE-ADİ) ş*, Cıl t E11, mSayı 1e8, kAra lıkA 20r20a, 11ş9-t15ı0rma Dergisi (GEAD), 115-132.

- Elgersma, C. (2024, March 06). Common Sense Education. Retrieved April 2024, from ChatGPT and Beyond: How to Handle AI in Schools: <https://www.commonsense.org/education/articles/chatgpt-and-beyond-how-to-handle-ai-in-schools>
- Ghosal Tirthankar, R. V. (2019). DeepSentiPeer: Harnessing Sentiment in Review Texts to Recommend Peer Review Decisions. In D. T. Anna Korhonen (Ed.), In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 1120–1130, Florence, Italy. Association for Computational Linguistics (pp. 1120–1130). Florence: Association for Computational Linguistics.
- Güçlütürk, G. (2022). ChatGPT ile Üretilen İçeriklerin Eser Niteliğinin 5846 Sayılı Fikir ve Sanat Eserleri Kanunu Bakımından Değerlendirilmesi. *Galatasaray Üniversitesi Hukuk Fakültesi Dergisi*, 2, 1899-1918.
- Habibi, A., Muahimin, M., Daniboa, B. K., Wilbowo, Y. G., Wahyuni, S., & Octavia, A. (2023). ChatGPT in higher education learning: Acceptance and use. *Computers and Education: Artificial Intelligence*, 5, 1-9.
- Ina. (2022, March 15). ONLIM. Retrieved April 2024, from The History of Chatbots from ELIZA to ChatGPT: <https://onlim.com/en/the-history-of-chatbots/>
- Kasparov, G. (2017, 20 6). YouTube. Retrieved February 2024, from TED: https://www.google.com/search?sca_esv=3c5e3101375fa6d2&sxsrf=ACQVn092Qfzp_woGSShOaW10zTUwyVn9hg:1709121641449&q=Deep+Blue+computer,+which+was+able+to+play+chess+and+beat+international+grandmaster+Garry+Kasparov&tbm=vid&source=lnms&sa=X&ved=2ahUKEwiq1MOi_s
- Keskin, E. K. (2023). Yapay Zekâ Sohbet Robotu Chatgpt Ve Türkiye İnternet Gündeminde Oluşturduğu Temalar. *Yeni Medya Elektronik Dergi*, 114-131.
- McMahon, L. (2024, May 20). BBC. Retrieved May, 2024 from ChatGPT to Lose Voice Over Johansson Similarity: <https://www.bbc.com/news/articles/c51188y6n6yo>
- Meço, G., & Çoştu, F. (2022). Using Artificial Intelligence in Education: Descriptive Content Analysis Study. *Karadeniz Teknik Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Dergisi*, 171-193.
- Mrowiski, M. J. (2017, September 20). Artificial Intelligence in Peer Review: How Can Evolutionary Computation Support Journal Editors? Retrieved March 2024, from Plos One: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0184711>
- OpenAI. (2023). GPT-4 Technical Report.
- OpenAI. (2024, May 13). *OpenAI*. From OpenAI: <https://openai.com/>

- Özcan, B. (2023, March 19). GPT-4 Out! What happens now? Retrieved April 2024, from Youtube: <https://youtu.be/hNA685AihMk?si=SPJn5azBoehs0gpR>
- Reyhani, N. (2023, May 01). Büyük Çaplı Dil Modelleri Nedir, Ne Değildir? ChatGPT Devrimi. Bilim ve Gerçek, 229. Retrieved April 2024, from Büyük çaplı dil modelleri nedir, ne değildir? ChatGPT devrimi: <https://bilimvegelecek.com.tr/index.php/2023/05/01/buyuk-capli-dil-modelleri-nedir-ne-degidir-chatgpt-devrimi/>
- Reiser, R. A. (2001). A history of instructional design and technology: Part I: A history of instructional media. Educational Technology Research and Development, 49(1), 53-64. Form_Builder&utm_term=&utm_content=1012782&gad_source=1&gclid=Cj0KCQ
- Sachdeva, N. (2024). Daffodil. Retrieved April 2024, from Software Engineering Insights: <https://insights.daffodilsw.com/blog/the-history-and-evolution-of-chatbots>
- Sequoia Capital. (2024, March 27). What's next for AI agentic workflows ft. Andrew Ng of AI Fund. Retrieved April 2024, from Youtube: <https://youtu.be/sal78ACtGTc?si=DLpTainLdJe4ZpD4>
- Schiffer, Z., & Newton, C. (2023, March 13). Microsoft just laid off one of its responsible AI teams. Retrieved April 2024, from Platformer: <https://www.platformer.news/microsoft-just-laid-off-one-of-its/>
- Shree, P. (2020, November 10). Medium. Retrieved March 2024, from The Journey of Open AI GPT models: <https://medium.com/walmartglobaltech/the-journey-of-open-ai-gpt-models-32d95b7b7fb2>
- Shevchuk, V. (2023, May 18). GPT-4 Parameters Explained: Everything You Need to Know. Retrieved May, 2024 from Medium: <https://levelup.gitconnected.com/gpt-4-parameters-explained-everything-you-need-to-know-e210c20576ca>
- Sky News. (2023, April 30). Retrieved May, 2024 from 'We don't understand how AI works' - British Computer Scientist: <https://youtu.be/sfn0edwl0hY?si=qGZZIDhdqxr9aqtC>
- Srivastava, M. (2023). A Day in the Life of ChatGPT as an Academic Reviewer: Investigating the Potential of Large Language Model for Scientific Literature Review. Retrieved from <https://osf.io/preprints/osf/wydct>
- Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in Higher Education: Considerations for Academic Integrity and Student Learning. Journal of Applied Learning & Teaching, 6(1), 31-40.
- Sutskever, I. (2023, November 2). Ilya: The AI scientist Shaping the World. (Guardian, Interviewer)
- UN. (2024, February 28). Sustainable Development Goals. Retrieved from https://sdgs.un.org/goals/goal4#targets_and_indicators

UNESCO. (2023). ChatGPT and Artificial Intelligence in Higher Education: Quick Start Guide. Paris: United Nations Educational, Scientific and Cultural Organization.

Yılmaz, T., Yasemin, D., & Sucuoğlu, H. (2023). Öğretmen Adaylarının Chatgpt'ye Yönelik Görüşleri. ICCI-EPOK 2023 (pp. 299-302). Aydın: Adnan Menderes University.

Yüksek Öğretim Kurulu. (2024). Yükseköğretim Kurumları Bilimsel Araştırma ve Yayın Faaliyetlerinde Üretken Yapay Zeka Kullanımına Dair Etik Rehber. Ankara: YÖK.

Teacher Feedback Literacy: EFL Writing Teachers' Beliefs in the Context of Turkish Higher Education

H. Kübra Er¹ Büşra Dağdemir²

¹Assist. Prof. Dr., Erzurum Technical University, Türkiye

²Lect., Erzurum Technical University, Türkiye

Abstract

Writing teacher feedback literacy (WTFL) is an essential component of assessment literacy (Lee, 2017). WTFL, which stands for Writing Teachers' Feedback Competencies, is a comprehensive framework that encompasses the knowledge, values, and abilities that writing teachers possess in relation to feedback. It serves as a valuable tool for studying teachers' skills in creating effective feedback processes and is a significant contribution to the field of second language (L2) writing research. The current study has a quantitative design and the data tool for the current study is the “Feedback Literacy Scale for L2 Writing Teachers” adopted by Lee et. al (2023). The questionnaire aims to elicit L2 writing instructors' backgrounds as well as their beliefs and values on WTFL. The data is analyzed via the Whitney U Test and Kruskal-Wallis H tests. As the findings of the study suggest, there are differences between teachers with different teaching experiences for some of the items' which is higher for the values dimension for the instructors with more than 10 years of experience. Similarly, there are some significant differences between teachers with different educational backgrounds for some of the items in which the ability dimension was higher for the instructors holding a PhD/MA degree.

Keywords: Writing skills, EFL writing teachers, writing teacher feedback literacy

1. Introduction and Theoretical Framework

The notion of feedback literacy began to acquire prominence around a decade ago, evolving as a derivative of the increasing focus on assessment literacy. Assessment literacy encompasses the comprehension of assessment concepts, methods, and processes, along with the capability to implement assessment strategies to enhance learning outcomes (Popham, 2009). Feedback literacy emphasizes the critical skills, knowledge, and habits required for effective engagement with feedback, both as providers and recipients. This approach has been informed by research highlighting the significance of feedback in educational settings as a crucial catalyst for student learning.

Preliminary investigations into feedback literacy focused on student feedback literacy within higher education. Researchers Carless and Boud (2018) and Sutton (2012) emphasized the necessity for students to comprehend and utilize feedback effectively. Student feedback literacy is the capacity to understand, assess, and respond to feedback, promoting autonomous learning and self-regulation. Carless and Boud (2018) assert that feedback literacy for students encompasses the ability to evaluate their own work, manage the emotional responses to feedback, and incorporate input into subsequent performances. This initial set of work established the foundation for comprehending feedback as a reciprocal process, necessitating active engagement from students, rather than perceiving them as passive recipients of teacher remarks.

Recently, there has been an increasing focus on teachers' feedback literacy, especially on their roles as designers and facilitators of effective feedback processes (e.g., Carless & Winstone, 2020; Winstone et al., 2019; Xu & Carless, 2017). This emphasis acknowledges that educators significantly influence the manner in which feedback is conveyed, comprehended, and applied by pupils. Teacher feedback literacy in higher education encompasses the information, abilities, and attitudes necessary to establish feedback systems that actively include students and improve their capacity to interpret and utilize feedback successfully (Carless & Winstone, 2020, p.4). It underscores the teacher's obligation not only to deliver feedback but also to foster situations that enable students to acquire the skills essential for effectively utilizing the input they receive.

Feedback literacy encompasses a multifaceted notion comprising multiple essential elements. Initially, it necessitates a profound comprehension of feedback knowledge, encompassing the theories, concepts, and optimal practices related to feedback (Nicol & Macfarlane-Dick, 2006). This information includes the functions of feedback, which are to reinforce strengths, rectify flaws, and direct future learning. Secondly, feedback literacy encompasses the processes and practices—the specific ways employed in delivering feedback. These practices encompass standard written feedback on assignments as well as more interactive methods, such as peer review and dialogue-based feedback, which involve students in reflective and collaborative learning processes.

Moreover, feedback literacy necessitates proficiency in developing feedback mechanisms that are significant and enduring. Educators must provide feedback mechanisms that are both effective and promote continuous student involvement and growth. Carless and Winstone (2020) assert that feedback must be regarded as an iterative process, incorporating feedback loops that enable students to reflect on and respond to the input they receive. The transition from perceiving feedback as a singular occurrence to recognizing it as a continuous dialogue

between teacher and student is fundamental to the notion of sustainable feedback (To & Panadero, 2019). Sustainable feedback emphasizes the cultivation of enduring learning practices, empowering students to engage in self-assessment and self-regulation of their learning, so reaping the advantages of feedback beyond specific tasks.

Moreover, feedback literacy is founded on principles that encourage constructive and significant feedback exchanges. These principles encompass transparency, trust, and reciprocal accountability between educators and learners (Nicol, 2010). Feedback literacy promotes a paradigm change in which teachers and students are regarded as participants in the educational process. This collaborative method fosters a culture of feedback where feedback serves not only as evaluation or correction but as a mechanism for development. Educators should appreciate students' viewpoints and reactions to feedback, while students must assume responsibility for interacting with feedback and use it to enhance their academic performance.

In this context, feedback is regarded as a process rather than a mere outcome, highlighting its dynamic and continuous nature. The process encompasses several stages: delivering clear and constructive information, motivating students to implement the feedback, and subsequently offering additional learning chances. Carless and Winstone (2020) emphasize that feedback is a mutual responsibility of teachers and students, with both actively participating in the comprehension and effective utilization of feedback. This collective obligation necessitates that educators possess proficiency in providing feedback while simultaneously cultivating student agency, enabling learners to autonomously seek, analyze, and implement feedback.

In conclusion, feedback literacy is a developing domain that underscores the significance of both educators and learners in the feedback mechanism. Initial research concentrated on student proficiency in comprehending and utilizing feedback; however, contemporary studies have broadened to investigate how educators might develop feedback systems that more effectively enhance student learning. As research progresses, a comprehensive grasp of teacher feedback literacy will be crucial in formulating feedback methods that not only rectify but also empower students to become autonomous learners.

2. Literature Review

In recent years, the concept of feedback literacy has garnered heightened scholarly interest in the domain of L2 writing. Feedback literacy encompasses the information, skills, and attitudes required by students and teachers to proficiently interpret and utilize feedback (Carless & Boud, 2018). It especially includes comprehension of feedback mechanisms, the capability to evaluate the quality of one's own work, and the ability to implement feedback to enhance learning. Feedback literacy, as a novel notion, has broadened the discourse beyond the conventional view of feedback as a unilateral transfer of information from educator to learner.

Historically, research on feedback in second language writing has emphasized the assessment of feedback mechanisms, concentrating on techniques for delivering corrective feedback to learners. For instance, research conducted by Bitchener & Knoch (2010) and Van Beuningen, De Jong, and Kuiken (2012) investigated the efficacy of direct and indirect corrective feedback—terms that denote whether the instructor explicitly rectifies an error (direct) or indicates that an error has occurred without supplying the correction (indirect). Various studies (e.g., Ellis et al., 2008; Frear & Chiu, 2015; Rahimi, 2021; Sheen et al., 2009) have examined whether feedback should be targeted (concentrating on specific error types) or comprehensive

(covering all errors in a text), and the impact of these methodologies on language learning outcomes.

Although these studies have enhanced our comprehension of the role of feedback in enhancing linguistic correctness in student writing, they predominantly focus on feedback as a product rather than as a process. The notion of feedback literacy promotes a transition to perceiving feedback as an interactive and dynamic process in which students actively participate with the input they obtain. This transition emphasizes the learner's responsibility in feedback, which encompasses not just receiving and comprehending criticism but also utilizing it to regulate and enhance their writing performance.

Notwithstanding extensive study on feedback types, student reactions to feedback remain inadequately examined. Research conducted by McMartin-Miller (2014), Han and Hyland (2015), and Zheng and Yu (2018) initiates an exploration of students' perceptions and utilization of feedback. Nonetheless, student feedback literacy—the capacity of students to interact proficiently with feedback—continues to be a somewhat novel field of study. Carless and Boud (2018) describe student feedback literacy as the ability to comprehend, value, and utilize feedback effectively. It entails fostering agency in the feedback process, wherein students are not mere recipients of instructor input but active contributors to knowledge construction via feedback.

Student feedback literacy encompasses various elements, such as the capacity to solicit feedback, assess its relevance, navigate the emotional repercussions of feedback, and incorporate it into subsequent endeavors (Winstone et al., 2017). The issue of emotional management is particularly vital, since students frequently see feedback as a source of anxiety or aggravation, particularly when it underscores their mistakes. Winstone et al. (2017) present a paradigm for feedback engagement comprising four essential components: appreciating input, making evaluative judgments, managing emotions, and executing actions to effectuate improvements. Each of these components is essential in cultivating a feedback-literate learner capable of engaging carefully and reflectively with feedback.

Nevertheless, the focus on teacher feedback literacy has been somewhat restricted. Teacher feedback literacy, as articulated by Lee (2021), denotes educators' expertise and competencies in delivering appropriate feedback, alongside their capacity to assist students in utilizing feedback to improve their writing. Teacher feedback literacy encompasses comprehending the objectives of feedback, identifying the appropriate timing and methods for delivering feedback that enhances learning, and effectively conveying feedback in a manner that is accessible and actionable for students. Lee and Mao (2024) emphasize that L2 writing instructors frequently lack adequate training in feedback literacy, hindering their capacity to deliver effective feedback that fosters students' writing advancement.

Research conducted by Lee and Mao (2024) on instructors' self-reported Writing Teacher Feedback Literacy (WTFL) revealed that while educators frequently have strong beliefs regarding the significance of feedback, their actual practices and comprehension may be inadequate. Many educators acknowledge that feedback ought to be formative and progressive; nevertheless, they may lack the requisite skills or confidence to deliver feedback that coincides

with these objectives. This highlights the necessity for further concentrated research on how educators may cultivate feedback literacy and include it into their teaching methodologies.

The significance of feedback literacy for educators is further demonstrated in research conducted by Rad and Mirzai (2024), which investigated the impact of a written feedback literacy intervention. This intervention aimed to enhance the feedback literacy of both students and teachers by integrating feedback-related activities into the writing curriculum. Rad and Mirzai discovered that these treatments not only augmented students' proficiency in comprehending and utilizing feedback but also bolstered their resilience, an essential facet of feedback literacy that encompasses students' ability to endure and adjust in reaction to constructive criticism.

In the context of formative feedback, which is designed to facilitate students' continuous learning rather than merely assess their performance, researchers such as Black and Wiliam (2009) contend that effective formative feedback should be timely, specific, and concentrated on methods for student improvement. In a similar vein, Sadler (1989) characterizes good feedback as information that enables students to bridge the disparity between their present performance and their aspirational performance. These criteria emphasize the need of feedback that is both corrective and educational, directing pupils towards advanced cognitive processes and self-regulation.

A notable notion in feedback literacy is sustainable feedback, as suggested by To and Panadero (2019), which pertains to feedback techniques that promote enduring learning. Sustainable feedback fosters students' capacity for self-assessment and self-regulation of their learning beyond the immediate assignment. This method corresponds with Nicol's (2010) principles of effective feedback, which promote input that elucidates learning objectives, aids students in cultivating self-assessment abilities, and encourages discourse around learning. Sustainable feedback is especially beneficial in L2 writing, aiming to cultivate autonomous authors capable of independently evaluating and enhancing their work.

In conclusion, although significant progress has been made in comprehending feedback literacy from the students' viewpoint, additional study is required to investigate and enhance feedback literacy among second language (L2) writing instructors. A robust framework for teacher feedback literacy would rectify the shortcomings noted in existing studies and ultimately improve the quality of feedback provided in L2 writing instruction. Incorporating both student and teacher feedback literacy into this research will be essential for creating more comprehensive and effective feedback systems in educational environments.

3. Methodology

3.1 Research Questions

This study has two research questions which are;

- 1) Is there a significant difference in EFL Teachers' attitudes towards 'Teacher Feedback Literacy' values and abilities dimensions depending on teachers' teaching experience?

- 2) Is there a significant difference in EFL Teachers' attitudes towards 'Teacher Feedback Literacy' values and abilities dimensions depending on teachers' educational background?

3.2 Participants and Context

The participants are 73 EFL teachers; below Tables 1 and 2 show the detailed information of the participants regarding EFL teachers teaching experiences and educational backgrounds in the context of Turkish Higher Education.

Table 1: Teaching Experiences

EFL teachers less than 10 years of experience	EFL teachers more than 10 years of experience
N: 35	N: 38

Table 2: Educational Background

EFL teachers holding a BA degree	EFL teachers holding MA/PhD degree
N: 28	N: 45

The context of the current study is; the English Language Teaching Department and School of Foreign Languages in Turkish Higher Education. The informed consent form was taken from each participant.

3.3 Procedures and Tools

The study followed a quantitative approach. The data tool for the current study is the "Feedback Literacy Scale for L2 Writing Teachers" adapted by Lee et. al (2023). The questionnaire aims to elicit L2 writing teachers' beliefs and values on WTFL (Appendix A). The 34 five-point Likert-scale items on the scale address the knowledge (10 items), values (12 items), and abilities (12 items) related to feedback in the three aspects of WTFL. However, the questionnaire was adapted to be used only in two dimensions of values (12 items), and abilities (12 items) for the sake of the current study.

3.4 Data Analysis

The normality of the data distribution was assessed using the Kolmogorov-Smirnov and Shapiro-Wilk tests. Two tests indicated that the data did not follow a normal distribution (significance level = .00; $p < .05$), so non-parametric testing should be used. Data analysis was conducted using the Mann-Whitney U test and the Kruskal-Wallis test (Greasley, 2008). The data is analyzed via non-parametric tests Whitney U Test and Kruskal-Wallis H tests. A reliability test was used to estimate the internal consistency of students' questionnaires. The value of the Cronbach Alpha correlation coefficient of the items ($r=.87$) is considered an acceptable result for the reliability of a questionnaire (Dörnyei & Taguchi, 2010). Table 3 shows the reliability statistics of the items.

Table 3: Reliability Statistics of the Questionnaire

Cronbach's Alpha	N of Items
,876	24

4. Findings

4.1 Findings for Research Question One

The Mann-Whitney and Kruskal-Wallis tests were run to explore if there are significant differences between teachers with different teaching experiences. The below table shows the results of the Whitney and Kruskal-Wallis tests.

Table 4: EFL Teachers with Different Teaching Experiences

	Mann-Whitney U	Wilcoxon W	z	Asymp. Sig. (2-tailed)
1.	632,50	1373,50	-,38	,70
2.	507,50	1248,50	-2,02	,04
3.	622,00	1363,00	-,49	,62
4.	585,00	1326,00	,97	,32
5.	587,00	1328,00	-,93	,34
6.	618,50	1359,50	-,64	,51
7.	533,00	1274,00	-1,62	,10
8.	535,50	1276,50	-1,58	,11
9.	513,00	1254,00	-1,87	,06
10.	318,50	1059,50	-4,10	,00
11.	580,50	1321,50	-1,02	,30
12.	534,50	1275,50	-1,62	,10
13.	549,00	1290,00	-1,43	,15
14.	546,50	1287,50	-1,41	,15

Table 4. (continued)

15	564,50	1305,50	-1,23	,21
16	604,50	1234,50	-,70	,48
17	658,00	1399,00	-,08	,93
18	608,00	1277,00	-1,57	,11
19	608,00	1349,50	-,66	,50
20	650,00	1280,00	-,17	,86
21	586,00	1216,00	-,92	,35
22	548,00	1289,00	-1,36	,17
23	661,00	1402,00	-,04	,96
24	542,00	1283,00	-1,41	,15

The Mann-Whitney and Kruskal-Wallis tests revealed a significant difference between teachers with different teaching experiences for item 2 = 0,43 ($p < 0,05$)

‘Feedback should encourage students to revise their writing’ and item 10 = 0,00 ($p < 0,05$)

‘Teachers should encourage students to make individual feedback requests.’

Mean Rank for item 2 = 41.50 for more than 10 years and 32.86 for less than 10 years

Mean Rank for item 10 = 46.90 for more than 10 years and 33.61 for less than 10 years.

4.2 Findings for Research Question Two

The Mann-Whitney and Kruskal-Wallis tests were run to explore if there are significant differences between teachers with different educational backgrounds. The below table shows the results of the Whitney and Kruskal-Wallis tests.

The Mann-Whitney and Kruskal-Wallis tests revealed a significant difference between teachers with different educational backgrounds for item 13 = 0.49 ($p < 0,05$)

‘How would you rate your ability to reflect on your own feedback practice to further improve it?’ and item 16 = 0,08 ($p < 0,05$)

‘How would you rate your ability to involve other colleagues in professional learning about feedback?’

Mean Rank for item 13 = 31.48 for BA and 40.43 for PhD/MA

Mean Rank for item 16 = 29.13 for BA and 41.90 for PhD/MA

Table 5: EFL Teachers with Different Educational Background

	Mann-Whitney U	Wilcoxon W	z	Asymp. Sig. (2-tailed)
1.	514,00	1549,00	-1,39	,16
2.	574,00	1609,00	-,74	,45
3.	597,00	1003,00	-,39	,69
4.	537,00	943,00	-1,16	,24
5.	592,00	1627,00	-,46	,63
6.	621,00	1656,00	-,12	,89
7.	617,00	1652,00	-,16	,86
8.	598,50	1004,50	-,39	,69
9.	625,00	1660,00	-,06	,95
10.	626,50	1661,50	-,04	,96
11.	629,50	1664,50	-,00	,99
12.	604,50	1639,50	-,32	,74
13.	475,50	881,50	-1,96	,04
14.	556,50	962,50	-,89	,36
15.	552,00	958,00	-,98	,32
16.	409,50	815,50	-2,64	,00
17.	621,00	1027,00	-,11	,91
18.	541,50	947,50	-1,10	,26
19.	608,00	1014,00	-,26	,79
20.	541,50	947,50	-1,05	,29
21.	517,50	923,50	-1,35	,17
22.	616,00	1022,00	-,16	,86
23.	518,00	924,00	-1,35	,17
24.	585,00	991,00	-,53	,59

5. Discussion

The findings of the values dimension indicate that the participants mostly endorse various feedback concepts discussed in the literature, except their uncertainty regarding feedback in timely feedback, multiple-draft classrooms, and feedback for promoting learner autonomy. Some teachers may have encountered challenges in implementing the aforementioned strategies, as evidenced by the prevalence of single drafting in numerous EFL classrooms, the presence of large class sizes, and the burden of workloads, resulting in difficulties in providing prompt feedback. Additionally, students' inclination to depend on teachers for support further contributes to these challenges (Ferris & Kurzer, 2019; Lam, 2019). Teachers' perspectives on the nature and purpose of feedback may be shaped by their EFL work environments, which are primarily focused on exams and prioritize written output above learner independence (Lee, 2017).

The findings from the abilities dimension suggest that the participants regarded themselves as somewhat skilled in the majority of components of WTFL described in the section. The participants conveyed a sense of uncertainty regarding their capacity to proficiently employ feedback to promote students' involvement in peer and self-feedback, as well as to exploit technology for augmenting feedback. This indicates that these feedback strategies were not commonly employed in their regular teaching practices, similar to other EFL settings (Junqueira & Payant, 2015; Mao & Crosthwaite, 2019). The results suggest that it is crucial to prioritize the enhancement of instructors' professional training in various settings. The absence of such training is likely to have an adverse impact on the instruction and evaluation of writing (Lam, 2019).

To conclude, the differences between teachers with different teaching experiences for item 2 and item 10 were higher in the reflection for values for the instructors with more than 10 years of experience. Similarly, the difference between teachers with different educational backgrounds for item 13; and item 16 were higher in the reflection for abilities for the instructors holding a PhD/MA degree.

6. Conclusion and Suggestions for Further Research

As the findings of the study suggest, there are differences between teachers with different teaching experiences for some of the items' which is higher for the values dimension for the instructors with more than 10 years of experience. Similarly, there are some significant differences between teachers with different educational backgrounds for some of the items in which the ability dimension was higher for the instructors holding a PhD/MA degree.

The results of this study hold significant significance for the domain of English Language Teaching (ELT). The findings underscore the necessity for a more varied, assessment-oriented curriculum to improve feedback literacy in English preparation programs. This necessitates the revision of syllabi and texts to enhance the integration of feedback mechanisms, enabling students to actively engage with and apply feedback for their linguistic growth. The study emphasizes the significance of teacher training programs that focus on feedback literacy, ensuring that educators are adequately prepared to deliver practical, formative feedback. Improved teacher feedback literacy will facilitate more effective communication of assessment concepts, empowering teachers to assist students in evaluating and utilizing feedback to

enhance learning outcomes. Additional study with an expanded focus is essential to enhance our comprehension of language assessment literacy, which includes both feedback and comprehensive assessment methodologies. Integrating classroom observations in next studies would provide deeper insights into the delivery and application of feedback in real-world contexts, resulting in more pragmatic suggestions. In conclusion, emphasizing the literacy of both student and teacher feedback through curriculum reform and professional development is crucial for establishing more effective feedback systems in English Language Teaching, hence enhancing language learning results.

6.1 Limitations of the Study

This study encountered some constraints that warrant acknowledgment. The limited sample size restricted the generalizability of the results. The narrow participant pool may not fully reflect the wider community of English language learners or educators in comparable educational settings. Subsequent study ought to incorporate a larger and more heterogeneous sample to augment the validity and relevance of the findings across various contexts and populations. The study predominantly utilized a Quantitative Research Design, yielding significant statistical insights while constraining the comprehension of the intricacies of feedback procedures. A qualitative research approach, utilizing methodologies like semi-structured interviews and classroom observations, may have produced more comprehensive and nuanced data. These qualitative methodologies would provide a comprehensive examination of the subjective experiences and perspectives of both educators and learners, providing insights into the interpretation and application of feedback in practice. Incorporating these qualitative elements, subsequent research could yield a more thorough comprehension of feedback literacy and its influence on language acquisition.

The study also failed to evaluate contextual factors that could affect feedback processes, like institutional policies, classroom dynamics, or cultural considerations. Subsequent study ought to examine these characteristics to enhance comprehension of the complex nature of feedback literacy across diverse educational contexts. The study's cross-sectional design restricted the evaluation of temporal changes in feedback literacy. Longitudinal research may elucidate the evolution of feedback literacy through continuous training and experience, facilitating a more nuanced comprehension of the development of feedback literacy in both teachers and students. In conclusion, overcoming these constraints in future study will enhance the knowledge of feedback literacy in English Language Teaching and facilitate the identification of effective ways for improving feedback practices across various educational contexts.

References

- Bitchener, J., & Knoch, U. (2009). The value of a focused approach to written corrective feedback, *ELT Journal*, 63(3), pp. 204–211. <https://doi.org/10.1093/elt/ccn043>
- Bitchener, J., & Knoch, U. (2010). Raising the linguistic accuracy level of advanced L2 writers with written corrective feedback, *Journal of Second Language Writing*, 19 (4), pp. 207–217. <https://doi.org/10.1016/j.jslw.2010.10.002>
- Carless, D., & Winstone, N. (2020). Teacher feedback literacy and its interplay with student feedback literacy. *Teaching in Higher Education Advance Online Publication*. <https://doi.org/10.1080/13562517.2020.1782372>
- Carless, D., & Boud, D. (2018). The development of student feedback literacy: Enabling uptake of feedback, *Assessment & Evaluation in Higher Education*, 43(8), pp. 1315–1325. <https://doi.org/10.1080/02602938.2018.1463354>.
- Chandler, J. (2003). The efficacy of various kinds of error feedback for improvement in the accuracy and fluency of L2 student writing, *Journal of Second Language Writing*, 12(3), pp. 267–296. [https://doi.org/10.1016/s1060-3743\(03\)00038-9](https://doi.org/10.1016/s1060-3743(03)00038-9)
- Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in Second Language Research: Construction, Administration, and Processing*, Second Edition
- Ellis, R., Sheen, Y., Murakami, M., & Takashima, H. (2008). The effects of focused and unfocused written corrective feedback in an English as a foreign language context, *System*, 36(3), pp. 353-371.
- Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes: How explicit does it need to be? *Journal of Second Language Writing*, 10(3), pp. 161–184. [https://doi.org/10.1016/s1060-3743\(01\)00039-x](https://doi.org/10.1016/s1060-3743(01)00039-x)
- Ferris, D., & Kurzer, K. (2019). Does error feedback help L2 writers? Latest evidence on the efficacy of written corrective feedback In K. Hyland, & F. Hyland (Eds.), *Feedback in second language writing: Contexts and issues* (2nd ed., pp. 109–124). Cambridge University Press.
- Frear, D., & Chiu, Y. (2015). The effect of focused and unfocused indirect written corrective feedback on EFL learners' accuracy in new pieces of writing. *System*, 53, pp. 24-34.
- Greasley, P. (2008). *Quantitative data analysis using SPSS: an introduction for health and social sciences*. Buckingham.
- Han, Y., & Hyland, F. (2015). Exploring learner engagement with written corrective feedback in a Chinese tertiary EFL classroom, *Journal of Second Language Writing*, 30, pp. 31-44.
- Huang, S. (2021). Examining student feedback literacy in an EFL writing course, *Studies in Educational Evaluation*, 70, 101093.

- Junqueira L., Payant C. (2015). 'I just want to do it right, but it's so hard': A novice teacher's written feedback beliefs and practices, *Journal of Second Language Writing*, 27, pp. 19–36.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong, *System*, 81, pp. 78–89. <https://doi.org/10.1016/j.system.2019.01.006>
- Lee, I. (2017). *Classroom writing assessment and feedback in L2 school contexts*. Springer.
- Lee, Icy & Mao, Zhicheng. (2024). Writing teacher feedback literacy: Surveying second language teachers' knowledge, values, and abilities, *Journal of Second Language Writing*. 63. 10.1016/j.jslw.2024.101094.
- Lee, I. (2021). Teacher feedback literacy and its implications for L2 writing pedagogy, *ELT Journal*, 75(2), pp. 137-145.
- Li, J., & Han, Y. (2022). Developing student feedback literacy in EFL writing classrooms: A longitudinal study, *Journal of Second Language Writing*, 55, 100804.
- Ma, X., Zheng, Y., & Xu, Q. (2021). The development of feedback literacy in the EFL writing classroom: A study of Chinese university students, *System*, 100, 102548.
- Mao S., Crosthwaite P. (2019) Investigating written corrective feedback: (Mis)alignment of teachers' beliefs and practice, *Journal of Second Language Writing*, 45, 46–60.
- McMartin-Miller, C. (2014). How much feedback is enough? Instructor practices and student attitudes toward error treatment in second language writing, *Assessing Writing*, 19, 24-35.
- Nicol, D. (2010). From monologue to dialogue: Improving written feedback in mass higher education. *Assessment & Evaluation in Higher Education*, 35(5), pp. 501-517.
- Rad, H. S., & Mirzaei, A. (2024). Developing feedback literacy, scaffolded writing, and resilience through intervention on feedback processes for L2 writing students, *Studies in Educational Evaluation*, 81(Complete). <https://doi.org/10.1016/j.stueduc.2024.101354>
- Rahimi, M. (2021). The role of individual differences in L2 learners' retention of written corrective feedback. *Studies in Second Language Learning and Teaching*, 11(1), pp. 89-115.
- Sheen, Y., Wright, D., & Moldawa, A. (2009). Differentiating the effects of focused and unfocused written corrective feedback in an English as a foreign language context, *System*, 37(4), pp. 556-569.
- Sutton, P. (2012). Conceptualizing feedback literacy: Knowing, being, and Acting, *Innovations in Education and Teaching International*, 49(1), pp. 31–40. <https://doi.org/10.1080/14703297.2012.647781>

- To, J., & Panadero, E. (2019). A critical review of the literature on the concept of 'sustainable feedback': An interdisciplinary perspective, *Assessment & Evaluation in Higher Education*, 44(5), pp. 706-720.
- Winstone, N. E., Mathlin, G., & Nash, R. A. (2019). Building student feedback literacy: Students' perceptions of the developing engagement with feedback toolkit. *Frontiers in Education*, 4(39), pp. 1–11. <https://doi.org/10.3389/feduc.2019.00039>
- Van Beuningen, C. G., De Jong, N. H., & Kuiken, F. (2012). Evidence on the effectiveness of comprehensive error, correction in second language writing. *Language Learning*, 62(1), 1–41. <https://doi.org/10.1111/j.1467-9922.2011.00674.x>
- Yıldız, H., Bozpolat, E., & Hazar, E. (2022). Feedback Literacy Scale: A Study of Validation and Reliability, *International Journal of Eurasian Education and Culture*, 7(19), pp. 2214-2249.
- Zheng, Y., & Yu, S. (2018). Student engagement with teacher written corrective feedback in EFL classroom: A case study of Chinese lower-intermediate students, *Assessing Writing*, 37, pp. 13-24.

PODELE: Podcasts Created by Spanish as a Foreign Language Students through a Project-Based Learning Approach

María Victoria Gálvez Frutos*

*Inst., Istanbul Technical University, Türkiye

Abstract

This paper introduces PODELE, a podcast project developed by students of Spanish as a foreign language at Istanbul Technical University. The study explores the methodology behind podcast creation, with a particular emphasis on Project-Based Learning (PBL). A comprehensive review of the literature on the use of podcasts in language education reveals that most existing research has focused either on the integration of technological tools or on their impact on learning outcomes. However, few studies have employed a methodology that directly supports language acquisition during the podcast development process. The findings, drawn from learning diaries and personal interviews, demonstrate that participants not only enhanced their communicative competence but also significantly improved their linguistic skills in Spanish. Furthermore, the project empowered students to create a series of online learning resources, effectively leveraging new technologies and fostering autonomy and self-directed learning. Participants also reported that their involvement in PODELE was highly motivating, leading to increased self-confidence and reduced anxiety, particularly in oral expression tasks.

Keywords: Podcasts, Project-Based Learning, learning dairies, teaching Spanish as a Foreign Language

Este artículo presenta PODELE, un proyecto de podcast desarrollado por estudiantes de español como lengua extranjera en la Universidad Politécnica de Estambul. El estudio explora la metodología subyacente a la creación de los podcast, con un énfasis particular en el Aprendizaje Basado en Proyectos. Una revisión exhaustiva de la literatura sobre el uso de podcasts en la enseñanza de lenguas revela que la mayoría de las investigaciones existentes se han centrado en la integración de herramientas tecnológicas o en su impacto en los resultados de aprendizaje. Sin embargo, pocos estudios han empleado una metodología que apoye directamente la adquisición del idioma durante el proceso de desarrollo del podcast. Los hallazgos, obtenidos a partir de diarios de aprendizaje y entrevistas personales, demuestran que los participantes no solo mejoraron su competencia comunicativa, sino que también perfeccionaron significativamente sus habilidades lingüísticas en español. Además, el proyecto capacitó a los estudiantes para crear una serie de recursos de aprendizaje en línea, aprovechando eficazmente las nuevas tecnologías y fomentando la autonomía y el aprendizaje autodirigido. Los participantes también informaron que su participación en PODELE fue altamente motivadora, lo que resultó en un aumento de la confianza en sí mismos y una reducción de la ansiedad, especialmente en las tareas de expresión oral.

Palabras clave: podcast, Aprendizaje Basado en Proyectos, diarios de aprendizaje, enseñanza del español como lengua extranjera

1. Introduction

This article seeks to present the PODELE study's experience to the educational community, with a particular focus on foreign language instructors. PODELE, an acronym for 'Podcast' and 'Español como Lengua Extranjera' (Spanish as a Foreign Language), is a podcast project developed over approximately 15 months by a group of students at Istanbul Technical University who were enrolled in Spanish for Foreigners courses.

The genesis of this project was a response to students' expressed need for podcasts specifically tailored to their proficiency levels. Upon recognizing the absence of such resources—since the vast majority of available podcasts are based on authentic texts not designed with language learners in mind—the author of this article initiated the development of a series of student-created podcasts. These podcasts were meticulously adapted to align with the language proficiency levels defined by the *Common European Framework of Reference for Languages* (hereinafter *CEFR*).

The distinctiveness and innovation of PODELE lie in its approach, where students assume the role of '*prosumers*', simultaneously acting as producers and consumers of their own content. Through the creation of these podcasts, students not only enhanced their communicative competencies but also contributed to the body of content available to other learners of Spanish as a second language. Additionally, the project has generated a corpus of data that is accessible for linguistic research.

1.1 Literature Review

The integration of digital tools such as wikis, blogs and podcasts into Spanish as a Foreign Language (SFL) classrooms had long been an outstanding issue until the shift to online education, mandated by the COVID-19 pandemic, made it an imperative necessity.

Regarding the concept of 'podcast,' which is the digital tool under investigation in this study, F. Rosell-Aguilar (2007: 472) explains that: “The word 'podcast' is a blend of 'iPod' and 'broadcast'; however, the prefix 'pod-' can be somewhat misleading, as podcasts are typically in MP3 format, which can be played on a variety of portable media players, computers and stereo systems and are not exclusive to the iPod.”

In terms of the relationship between language learning theories and podcasts, F. Rosell-Aguilar (2007: 476) further argues that the use of podcasts as a language learning tool is rooted in various theoretical frameworks, including constructivism⁷ (P.A. Cooper, 1993: 16-17),

⁷ *Constructivism* is delineated as a learning theory that emphasizes the learner's active role in constructing their own knowledge. Knowledge is not passively transferred from an external source to the learner; rather, it is actively constructed internally through interactions with the environment, prior experiences, and reflective processes concerning those experiences. In essence, learning is characterized as an active and constructive process in which learners generate their own interpretations of the world, grounded in their personal understanding and perception. The constructivist paradigm asserts that knowledge is inherently subjective and context-dependent. This necessitates that educators design learning environments that facilitate exploration, dialogue and critical reflection. By incorporating tasks such as PODELE, educators can create dynamic learning environments that support and enhance the constructive process, thereby aligning with the core principles of constructivism.

informal and lifelong learning⁸ (L. Naismith; M. Sharples; G. Vavoula & P. Lonsdale, 2004: 3), theories on the utilization of learning objects for educational materials⁹ (D. A. Wiley, 2002: 6), mobile learning¹⁰ (A. Kukulska-Hulme, 2005: 1-2), and practices such as chunking¹¹ and just-in-time teaching¹².

The majority of research on the use of podcasts in foreign language learning has focused on demonstrating their effectiveness in enhancing communicative skills (B. Chaves Yuste & C. De la Peña Álvarez, 2023). Specifically, most studies have concentrated on the impact of podcasts on *oral expression* and *listening comprehension* (M.-M. Sze, 2006; M. Abdous, M.M. Camarena, & B. R. Facer, 2009; L. Lomicka & G. Lord, 2011; S.W. J. McMinn, 2011; F. Rosell-Aguilar, 2013; M. Yoestara & Z. Putri, 2019; F. V. Gómez Barrios & E. Palma Velásquez, 2020; M. Tomé Díez & M. A. Richters, 2021; A. Nadig, 2021). A smaller body of research has examined how podcast creation in second language classrooms can contribute to the improvement of *writing skills*, *research abilities*, *information analysis* and *editing* (B. Phillips, 2015); *grammar* (J. W. Istanto, 2011); and *vocabulary* acquisition (L. Borgia, 2010; S. M. Putman & T. Kingsley, 2012). Other scholars have highlighted the role of podcast production in reducing student *anxiety* (M. Lee, C. McLoughlin, & A. Chan, 2008), increasing *motivation* (H. G. Fleenor & R. Hodhod, 2016), fostering *autonomy* (S. M. Ryan, 1997; T. Piñeiro Otero, 2012), and providing training in the use of new *technologies* (C. Chacón & C. Pérez, 2011).

This study seeks to offer a novel approach compared to previous research by concentrating on the podcast creation process through Project-Based Learning (PBL) and the content of the podcasts produced by students. As F. Rosell-Aguilar (2007) notes, many available podcasts

⁸ *The theory of informal learning* suggests that much of learning occurs outside traditional educational settings, through daily life, work, social interactions, and personal experiences. In other words, informal learning can be intentional, for instance, through intensive, significant and deliberate learning projects. Conversely, the theory of lifelong learning posits that learning is a continuous process, not confined to childhood or formal education. Both theories of informal learning and lifelong learning acknowledge that learning occurs continuously and is influenced by our environment and the specific situations we encounter. In this regard, PODELE is an extracurricular and voluntary project that aligns with the foundational principles of both theories.

⁹ *Learning objects* are defined as any entity, whether digital or non-digital, supported by technology, that can be used, reused, or referenced to facilitate the teaching and learning process. Podcasts fit this definition as they are educational resources that possess the characteristics of *modularity* (independent episodes), *reusability* (capable of being used an infinite number of times), *accessibility* (available at any time and from any location), flexibility (usable in various contexts), and *adaptability* to different learning styles (serving as supplementary materials or self-directed learning tools).

¹⁰ *Mobile learning* is defined as learning that occurs when the learner is not situated in a fixed, predetermined location or when the learner takes advantage of the learning opportunities afforded by mobile technologies. Podcasting shares several salient characteristics with the field of mobile learning, including its potential for *spontaneity*, *personalization*, *informality*, *contextuality*, *portability*, *ubiquity* (availability everywhere), and *pervasiveness* (so seamlessly integrated with daily activities that it is scarcely noticed).

¹¹ *Chunking* refers to the technique of dividing information into "chunks" or smaller, more manageable units. This practice facilitates the memorization and comprehension of material, as our brains are more adept at processing and retaining information when it is organized into logical segments. In an educational context, chunking might involve breaking down a complex topic into smaller subtopics and teaching them sequentially, thereby helping students absorb and understand the information. As will be illustrated in the subsequent sections, this technique has been applied in the development of the podcasts.

¹² *Just-in-Time Teaching* is a pedagogical strategy that adjusts the content of classes according to the immediate needs of students, typically based on their responses to prior assignments or quick assessments conducted immediately before class. This approach is grounded in the premise that instruction should address the current needs of students rather than adhere to a rigid, predetermined plan.

lack a methodological design because “having the technical know-how does not necessarily imply a pedagogic know-how.” In response to this observation, the PODELE project aims to provide the educational community with a methodology for podcast production in the foreign language classroom.

1.2 Research Purpose

The purpose of this study is to delineate *the methodology* employed in the development of podcast episodes by a group of students learning Spanish as a foreign language in a non-immersive context. Through the implementation of a project-based approach, the PODELE initiative aims to achieve several key objectives:

- To foster the development of students' *communicative competencies*, encompassing not only *linguistic abilities* (vocabulary, pronunciation, spelling) but also *sociolinguistic, discursive, strategic, and pragmatic* competencies.
- To enhance the *communicative skills* (speaking, listening, writing and reading) of both the students producing the podcasts and those engaging with them as listeners.
- To integrate *new technologies* as an extracurricular tool to promote *autonomy* and *self-directed* learning among learners of Spanish as a second language.
- To ensure the success of the learning process by cultivating a positive educational environment that enhances *motivation* while mitigating factors such as *anxiety*.
- To provide the educational community with a set of *online free resources* that can be utilized in the teaching and learning of Spanish by both learners and educators, and that also serve as a data corpus for researchers in future studies.

The extent to which these objectives have been achieved will be evaluated in the subsequent sections through the analysis of two qualitative instruments: learning diaries and personal interviews.

2. Process of Podcast Creation

In the subsequent sections, the creation process of the PODELE project is detailed, with particular emphasis on the methodology employed and the research instruments utilized.

2.1 Approaches and Methodologies

The methodology employed in this study is *Project-Based Learning (hereafter, PBL)*. PBL falls into the category of *activity-based teaching methods* and *active learning*¹³. Specifically, it focuses on *constructivist pedagogy*, where students construct their knowledge through practical and meaningful experiences. PBL is also associated with *interdisciplinary learning* and *student-centered learning*, promoting skills such as collaboration, critical thinking and problem-solving through long-term realistic projects.

Regarding the origins of this approach, it formally stems from the essay by American pedagogue W. H. Kilpatrick, *The Project Method* (1918), which draws inspiration from the

¹³ According to the Common European Framework of Reference for Languages (2020: 29): “Language learning should be directed towards enabling learners to act in real-life situations, expressing themselves and accomplishing tasks of different natures. It implies that the teaching and learning process is driven by *action*, that it is *action oriented*.”

pedagogical theory of 'learning by doing'¹⁴ expounded by his mentor J. Dewey in *Democracy and Education* (1916). As its name suggests, W. H. Kilpatrick's work advocates and theorizes a new educational methodology based on the concept of project.

With respect to the concept of a project, J. Zanón (1995: 55) defines a project as “a tool that involves students in a process that generates pedagogical tasks and communicative tasks.” Furthermore, the author asserts that projects contribute not only to the development of communication in a foreign language, but also to the enhancement of learning autonomy and educational values. The *Dictionary of Key Terms in Spanish as a Foreign Language* (2008) — in Spanish, *Diccionario de Términos Clave de ELE* — asserts that a project comprises a series of systematically organized and sequenced activities aimed at achieving a specific outcome or product. However, this definition proves insufficient to delineate PBL, as it could equally apply to what is known as Task-Based Learning (Subsequently abbreviated as TBL). Both approaches can be considered heirs to Communicative Language Teaching. Nevertheless, as J. Zanón (1999: 13-24) clarifies, the crucial distinction lies in PBL's engagement of students from the outset, involving them in negotiating the topic, planning activities and determining the schedule to be followed, culminating in a final task. C. N. Candlin (1990: 33-54) regards TBL as an approach that involves the acquisition of skills, abilities and competencies aimed at accomplishing a final task or a series of tasks that present diverse content. In this framework, involving students in the decision-making process regarding the final task is not necessarily imperative (Y. V. Jerez-Naranjo & A. A. Garófalo-Hernández, 2012: 3). Conversely, in TBL, the emphasis lies on the learners' objectives, whereby the individual learner or group determines the nature of the final product (S. Wolk, 1994: 42-45). In this way, TBL emphasizes concise, targeted exercises aimed at enabling students to refine specific language competencies and promptly receive evaluative feedback. On the other hand, as will be further elaborated in subsequent sections of this article, PBL involves long-term endeavors during which students engage in comprehensive investigations of subjects, thereby fostering deep understanding and proficiency in applying acquired knowledge and competencies.

In exploring the various ways to classify projects, W. H. Kilpatrick (1918: 16-17) identifies four distinct types¹⁵:

- i. *Productive Projects*: These aim to transform an idea or plan into a tangible form, such as creating a podcast series in a second language, composing a letter to a pen pal, or staging a play in the target language.
- ii. *Recreational or Consumer Projects*: Focused on providing an aesthetic experience, examples include listening to a story in the second language,

¹⁴ The concept of *learning by doing* promotes an educational approach focused on active experience, hands-on experimentation, and critical reflection. It underscores the vital connection between learning and real-life contexts. According to this theory, *meaningful learning* occurs when students actively participate in activities that are personally relevant and significant, enabling them to experiment with and reflect on their actions and outcomes. H. W. Reese (2011: 1) states that, “Learning by doing has been a principle for thousands of years; it has had many proponents, including Plato, Thomas Hobbes, English and Spanish epigrammatists, Karl Marx, Mao Zedong, cultural anthropologists, Montessori, John B. Watson, and B. F. Skinner.” In the view of this author, J. Dewey may have popularized the theory, but it was already familiar to Plato, who asserted that learning 'the philosophical method', which involves dialectic, occurs through its active application.

¹⁵ In his work, the author discusses types 1, 2, 3 and 4 without assigning specific names to these categories; the names have been provided by the author of this article. Additionally, examples related to language acquisition have been included to clarify the reader's understanding of the classification.

- watching films or TV shows in the target language, or appreciating the lyrics and meaning of a song.
- iii. *Problem-Solving Projects*: Designed to address intellectual challenges, such tasks involve solving specific issues, such as developing strategies to overcome language barriers in conversations or creating effective methods for memorizing vocabulary and grammar rules.
- iv. *Projects for Acquiring Specific Learning or Training*: Intended to develop particular skills or knowledge, these may include mastering advanced grammatical structures in Spanish, learning irregular verbs, or improving pronunciation through targeted exercises.

Concerning the classification of podcasts, J. Stanley (2005) identifies three distinct categories:

1. *Authentic Podcasts*: These are not specifically intended for students but can serve as valuable resources for listening practice. Typically, these podcasts are more suitable for advanced learners due to their complexity and content.
2. *Teacher Podcasts*: Produced by educators, these podcasts are often created for their own classes. They aim to support student learning by providing unique listening content that may not be available elsewhere or by offering a local perspective.
3. *Student Podcasts*: Created by learners, sometimes with teacher assistance, these podcasts allow individuals to listen to and engage with content reflecting the lives and interests of peers from around the globe, thus offering cultural insights and diverse perspectives.

By applying these two classifications to our study, it can be inferred that PODELE is a *productive project* consisting of *student-created podcasts*.

Addressing the potential phases in which a project can be conducted, this study has applied an adaptation of the framework proposed by scholars W. H. Kilpatrick (1918) and M. I. Vizcaíno Timón (2008). Table 1 compares the phases proposed by these authors with those utilized in this project.

The section concludes with an analysis of the advantages and disadvantages of the PBL approach, as articulated by experts in the field. Certain scholars, including A. Zábrazcki (2013: 324), argue that the *time* required to complete the final project is *disproportionately* demanding for both educators and students. Conversely, R. Tippelt and H. Lindemann (2001: 14) contend that implementing this approach is problematic when students lack motivation and curiosity to acquire new knowledge. To mitigate these potential limitations, PODELE was designed for students *who voluntarily participated*, thereby ensuring that the participants were genuinely committed to engaging in a project that necessitated a departure from conventional pedagogical models. Regarding the positive aspects of PBL, M.A. Fernández Martínez and J. Rodríguez Aviñó (1994: 236) provide a list of advantages, which include:

- i. Enhancing *self-directed* learning and *personal responsibility* of the student throughout the entire process;
- ii. Instruction in the principles of *inquiry-based learning* (e.g., material search, reading, documentary sources, summaries);

- iii. Promoting *communication* in the foreign language;
- iv. Providing students with the opportunity to apply what has been previously learned in a 'real' context;
- v. Improving the teacher-student relationship.

Similarly, S. Tobón (2006: 5) asserts that “projects stimulate students' interest in a specific subject, foster creativity, self-esteem, and innovation, and promote teamwork”. Evidently, no single approach to second language learning is universally applicable in all contexts. It is incumbent upon the language instructor to assess the suitability of this method based on the available resources, the characteristics of the students, and the objectives to be achieved.

Table 1: Contrast of the Possible Phases of a Project

W. H. Kilpatrick (1918)	M. I. Vizcaíno Timón (2008)	PODELE (2024)	
I. Purposing	1. Motivation Phase: Selection of the topic	A. Planning	i. Selection of the topic
	2. Recognition of Prior Knowledge		ii. List of requirements
II. Planning	3. Planning and Development of Ideas	B. Design	iii. Draft encompassing all the linguistic resources that will be utilized
	4. Organization and Proposal of Activities		iv. First draft
	5. Organization of Space		v. Preparation of the dialogue between the podcaster and guest
	6. Organization of Time		vi. Highlight the linguistic resources in the transcription
III. Executing	7. Information Retrieval	C. Execution	vii. Recording
	8. Collection, Organization, and Study of Information, Materials, and Resources Obtained		viii. Arrangements
		D. Dissemination	ix. Sharing on social media
IV. Judging	9. Preparation of Activities	E. Evaluation	x. Completion of the learning dairy
	10. Synthesis and Evaluation		xi. Personal interview

Source: Author's own elaboration based on W. H. Kilpatrick (1918), M. I. Vizcaíno Timón (2008) and PODELE.

2.2 Participants

This study involved a total of 21 students learning Spanish as a foreign language, all of whom exhibited a predominantly analytical profile. Specifically, 19 participants were engineering undergraduates, while 2 were enrolled in a science-focused secondary school, TÜBİTAK Science High School. Among the undergraduates, 16 were studying at Istanbul Technical University, while the remaining 2 were part of the Erasmus exchange program from technical universities across Europe. Consequently, 19 students were native Turkish speakers, one was a native German speaker and another was a native French speaker. Of the total participants, 15

were male and 6 were female. At the time of the podcast development, the university students were aged between 20 and 28 years, while the two secondary school students were 14 years old. The participants, who voluntarily engaged in the project, had Spanish proficiency levels ranging from A1 to B2, according to the CEFR.

Regarding the roles of the teacher and the student in the PBL approach, this study adopts the perspective of Professor A. L. Carlini (2004: 39-41). Table 2 synthesizes the functions assumed by the teacher and the student in this study, based on the work of this scholar.

Table 2: Roles of the teacher and the student in PODELE

Teacher's Role	Student's Role
Facilitate the selection of a topic by guiding students based on their communicative proficiencies	Choose a topic driven by personal interest and curiosity, thereby defining a clear objective
Collaborate with students to devise a comprehensive work plan, enumerating activities in a specific sequence	Undertake research and gather the requisite information for the effective completion of the project.
Direct and support the research and production phases of the podcast	Solicit guidance and assistance from the instructor as needed
Provide ongoing motivation and encouragement to students as required	Commit to creating a high-quality podcast that serves both personal and community interests.
Monitor the podcast's development through systematic and continuous observation	Adhere to established deadlines
Evaluate the completed podcast	Implement the instructions provided by the instructor to enhance the quality of the podcast

Source: Author's own elaboration

As illustrated in Table 2, the PBL approach unequivocally fosters the development of *student autonomy*, transforming learners into active participants in their own educational process. In the forthcoming sections, which outline the progression of the PODELE project, the reader will observe how the teacher's role has encompassed that of a guide, advisor, and coordinator, granting students significant autonomy while meticulously ensuring the attainment of the project's educational objectives. A notable strength of PBL is its capacity to facilitate the development of *learners' self-regulatory skills*. This is particularly significant given that such skills are not innate but are cultivated through strategically designed 'learning to learn'¹⁶ activities.

¹⁶ The principles of 'learning to learn' encompass a range of strategies and competencies designed to enhance one's ability to effectively acquire, process, and apply knowledge. These principles are often integral to educational frameworks and personal development paradigms. The Companion Volume of the Common European Framework of Reference for Languages (2018) and UNESCO's Education for Sustainable Development Goals place heightened emphasis on metacognitive skills and the concept of 'learning to learn', focusing on how learners can develop the capacity to reflect upon and improve their own learning strategies.

2.3 Research Instruments

The research instruments employed in this study included a learning diary and individual interviews with each student.

2.3.1 Learning Dairies

A learning diary is a *qualitative research tool* designed to gather data on participants' experiences, reflections, and perceptions related to their learning processes. Such diaries enable researchers to access participants' internal thoughts and emotions, offering profound and personal insights into how learning evolves over time. This instrument is particularly valuable in studies aiming to understand skill development, changes in attitudes, or the progression of knowledge. Linguist K. M. Bailey (1990: 215) defines the learning diary as “a first-person account of the language learning or teaching experience, documented through regular, candid entries and then analyzed for recurring patterns or salient events.” According to N. Bolger, A. Davis and E. Rafaeli (2003: 581): “The effectiveness of diary study designs depends on careful consideration of the question(s) one seeks to answer. A poorly designed diary study can involve considerable effort but may yield little useful information.” In this context, special attention was given to the questions posed to the students, as these were intended to serve as the basis for obtaining diaries that would be useful for both this research and their learning process. Students were asked to focus on the following aspects in their diaries:

- Their *emotions* throughout the podcast creation process, particularly during recording;
- *Problems, anxieties, and concerns* they encountered during the project;
- A record of *activities* undertaken independently, with their instructor, or in collaboration with their peers;
- Accounts of *success or failure* during the podcast creation.

Like any research instrument, learning diaries have inherent limitations. According to R. Schmidt and S. N. Frota (1986: 238), such instruments are subject to the evaluator's subjectivity. Furthermore, N. Bolger, A. Davis, and E. Rafaeli (2003: 592) argue that “to obtain reliable and valid data, diary studies must secure a level of participant commitment and dedication rarely required in other research contexts.” Despite these limitations, we have chosen to employ this instrument in our study because, as K. M. Bailey (1991: 63-64) asserts, the most effective method for understanding learners' internal processes is through direct introspection, with diaries serving as an optimal tool for facilitating this introspective process.

The learning diaries were examined upon submission by the students. This initial engagement with the documents was instrumental in formulating questions for subsequent interviews to clarify specific details. Following the collection of responses during the interviews, each student's diary was reviewed once more to arrive at more precise and comprehensive conclusions.

2.3.2 Personal Interviews

As outlined in *A Handbook for Teacher Research: From Design to Implementation* by C. Lankshear and M. Knobel (2004), personal interviews are a valuable tool in academic research for several reasons: they allow for the collection of detailed and in-depth information about

participants' experiences, opinions, and perceptions; facilitate the clarification of responses and enable the contextualization of information and the identification of nuances. In light of these considerations, a semistructured personal interview was conducted with each student upon the completion of their project. In addition to addressing specific questions based on their diary, students were asked to reflect on and conclude with responses to the following aspects:

- What knowledge did you acquire through the creation of the podcast, and which skills did you develop in the process?
- In what ways did your proficiency in Spanish improve?
- What was the most challenging aspect of creating the podcast for you?
- Why do you believe that this type of project should be incorporated into language classes?

2.4 Description of the Project

In the following sections, the content of the podcasts will be described in accordance with the CEFR levels, and the process of their creation will be detailed, considering the phases of planning, design, execution, dissemination, and evaluation.

2.4.1 Content of the Podcasts

The project comprises 20 podcasts: specifically, 3 at the A1 level, 10 at the A2 level, 6 at the B1 level, and 1 at the B2 level.

Table 3: Content of the Podcasts by Level

A1	A2		B1	B2
My brother is my best friend	Tips to take care of your health	The most important moments in the history of Türkiye	Futuristic aircraft design	My favorite spot
My routines: before in France and now in Istanbul	Leisure guide: What to do in Istanbul?	Experiences from a recent trip to Spain	Racial discrimination in Africa	
The routines of a Turkish teenage student	Türkiye: before and now	Atatürk's biography	Trends and cinematic habits of Turkish people	
	Typical Spanish and Turkish dishes	Difficulties when learning a language	Cultural differences between Spaniards and Turks	
	Albert Einstein's Biography	Beyoğlu	Travel stories	
			Environmental problems	

Source: Author's own elaboration

Regarding the selection of podcast topics by the students, four lists were compiled based on the series of textbooks from *Editorial Difusión* used in our Spanish courses: a list for the A1 level, drawn from the textbook *Aula Internacional Plus 1*; another for the A2 level, derived from *Aula Internacional Plus 2*; one for B1 level students, extracted from *Aula Internacional Plus 3*; and finally, the list for the B2 level was created using *Aula Internacional Plus 4*. Students were offered the option to choose a topic from the list adapted to their level or any other proposed topic, provided they had the linguistic resources to develop it. It was deemed beneficial to allow students some degree of freedom in choosing the topic for their podcast. In this way, the aim was not only to foster their creativity, but also to achieve better outcomes by motivating them to work on a topic they were passionate about. Table 3 displays the titles of the podcasts developed by the students along with the content level for each.

It is worth noting that 10% of the students chose topics outside the provided lists ('Futuristic aircraft design' and 'Racial discrimination in Africa'). Seventy-five percent of students adapted their selected titles to their personal experiences or fields of expertise, while only 15% retained the original titles without modifications. These percentages are particularly noteworthy given that the majority of the participants are *analytical students*—those who typically prioritize logic over creativity. One of the primary objectives of this project is to offer students the opportunity to explore previously uncharted avenues, guiding them through the process to enhance their capabilities and to foster the application of these experiences to various aspects of their lives, both linguistic and otherwise.

Content of the A1-Level Podcasts

Concerning the A1-level podcasts, it is particularly important to note that two of them, 'My Brother is My Best Friend' and 'The Routines of a Turkish Teenage Student', were developed by teenage students from the TÜBİTAK Science High School, which has a faculty exchange agreement with Istanbul Technical University, where the author of this article was involved. The creation of these podcasts by students as young as 13 exemplifies the potential success of such activities with younger learners, especially given their upbringing in the digital era. The students demonstrated significant enthusiasm for engaging in a project that incorporated elements beyond conventional educational paradigms.

The third A1-level podcast was created by a French student of Moroccan origin, who took the Spanish as a Foreign Language course during her Erasmus exchange at Istanbul Technical University. Her podcast, 'My Routines: Before in France and Now in Istanbul', features a distinct *intercultural element* by comparing her routines in Paris and Istanbul in the target language. The student employs a foreign language to juxtapose her habits in two capitals with differing cultural offerings, with her routines also influenced by her family's heritage. These factors collectively make her podcast an intriguing cultural mosaic.

Content of the A2-Level Podcasts

With respect to the content of the A2-level podcasts, the podcasts notably emphasize both the historical aspects of Türkiye, including 'Türkiye: Before and Now', 'The Most Important Moments in the History of Türkiye', and 'Atatürk's Biography', and elements of Turkish culture, such as 'Leisure Guide: What to Do in Istanbul?' and 'Beyoğlu'. The *cross-cultural component* is also evident in a podcast where one student presents an exposition on the most emblematic

vegetarian dishes of Spanish and Turkish cuisine. Moreover, a Turkish student returning from a trip to Madrid highlights the aspects she found most culturally striking. It was particularly intriguing to hear the biography of Albert Einstein, 'Albert Einstein's Biography', narrated by a German student of Polish origin who participated in the Erasmus exchange program at Istanbul Technical University. One student selected a topic related to Goal 3 of the United Nations' 2030 Agenda for Sustainable Development¹⁷: 'Tips to Take Care of Your Health'. Lastly, one student chose to discuss their personal experience learning foreign languages in a podcast titled 'Difficulties When Learning a Language', emphasizing the challenges encountered and offering recommendations for others engaged in the same process.

Content of the B1-Level Podcasts

In relation to the content of the B1-level podcasts, the students once again selected topics encompassing both the promotion of Turkish culture ('Trends and Cinematic Habits of Turkish People') and a *cross-cultural component* ('Cultural Differences between Spaniards and Turks from the Perspective of an Istanbul Erasmus Student in Valencia'). Notably, the students also addressed issues related to the United Nations' Sustainable Development Goals for 2030, such as 'Futuristic Aircraft Design' (Goal 9)¹⁸, 'Environmental Problems' (Goal 13)¹⁹, and 'Racial Discrimination in Africa' (Goal 16)²⁰. Lastly, a group of three students chose to develop a collaborative podcast, 'Travel Stories', which emphasizes the use of structures commonly employed in Spanish to express emotions.

Content of the B2-Level Podcasts

With reference to the B2-level podcasts, we have only one available, owing to the specific course offerings in Spanish provided by the Department of Foreign Languages, with which the author of this work is affiliated. The department offers courses at levels A1 (Spanish I and Spanish II), A2 (Spanish III and Spanish IV) and B1 (Spanish V and Spanish VI). Spanish as a Foreign Language courses constitute a selection of elective offerings available to all students at Istanbul Technical University. However, given that the engineering program typically spans four years and that Spanish courses are not required to align with other mandatory courses in students' programs for enrollment, it is challenging to meet the minimum student enrollment necessary to offer B2-level courses. In fact, the B2-level podcast was produced by a student following their Erasmus exchange in Spain. The student chose to discuss 'Heybeli' Island to acquaint their Spanish-speaking friends with their favorite place in Istanbul ('My Favorite Spot').

¹⁷ The 2030 Agenda for Sustainable Development is a comprehensive global action plan adopted by all member states of the United Nations in September 2015. Its objective is to eradicate poverty and foster shared prosperity and peace, while simultaneously safeguarding the environment over the ensuing 15 years. Goal 3 focuses on Good Health and Well-being: ensuring healthy lives and promoting well-being for all at all ages.

¹⁸ *Goal 9: Industry, Innovation, and Infrastructure:* Build resilient infrastructure, promote inclusive and sustainable industrialization, and foster innovation.

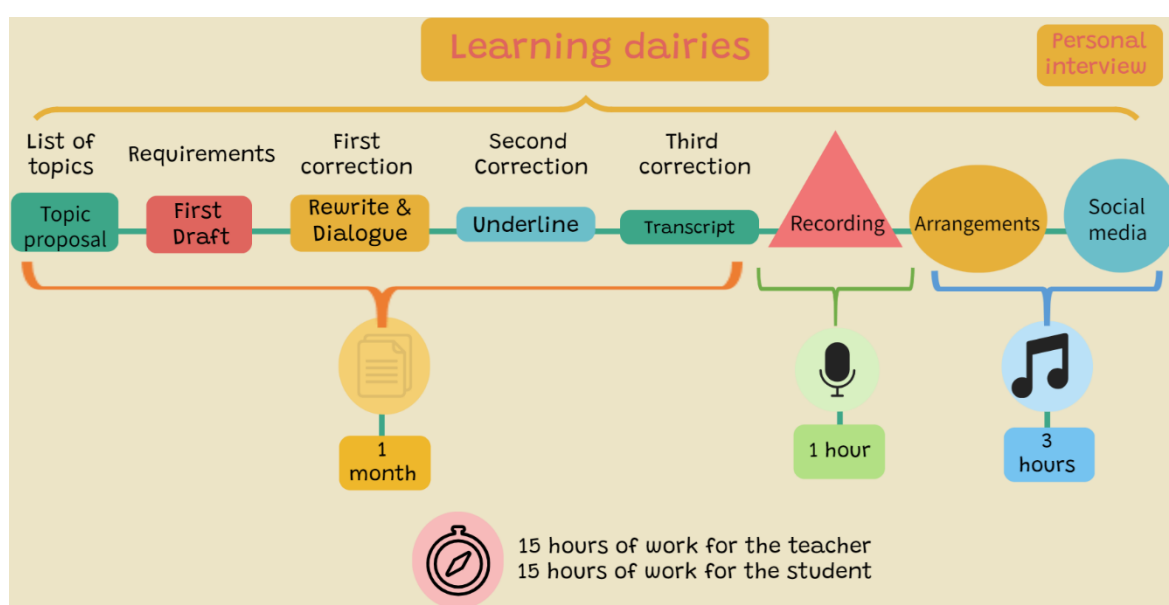
¹⁹ *Goal 13: Climate Action:* Take urgent action to combat climate change and its impacts.

²⁰ *Goal 16: Peace, Justice, and Strong Institutions:* Promote peaceful and inclusive societies for sustainable development, provide access to justice for all, and build effective, accountable and inclusive institutions at all levels.

2.4.2 Phases of Podcast Creation

The process of creating the podcasts has been guided by the principles of the 'divide and conquer' algorithm. The final task, the PODCAST, has been systematically divided into intermediate tasks to facilitate its completion by students. This strategic division aims to prevent the PODCAST from appearing overwhelming, allowing students to approach it incrementally. Drawing on the analogy of a battalion, defeating each soldier in the battalion is akin to winning the entire battle. Similarly, the successful completion of each intermediate task ultimately culminates in the completion of the podcast. In this context, Figure 1 illustrates the phases of production and the approximate time invested in each podcast.

Figure 1: Process of Podcast Creation



Source: Author's own elaboration

Moreover, as will be discussed in the following paragraphs, the analogy of the 'three coats of paint' has also been applied. According to the approach proposed by the author of this article, reviewing content at least three times ensures greater comprehension and, in this particular case, helps to avoid errors and guarantees the quality of the final product.

Planning

Each podcast begins with the student either selecting a topic from the list provided for each proficiency level by the instructor or proposing a topic to the teacher. If students choose an unrestricted topic, they must ensure they possess the necessary linguistic resources to adequately develop the subject matter. Once the topic has been selected, the student is allotted one week to consider the linguistic resources required for creating the podcast. This is feasible because the thematic list aligns with the instructional units covered in class, providing the student with the necessary knowledge to utilize these tools. At the end of this period, the instructor and the student meet to develop an outline detailing all the linguistic resources (grammatical, lexical, pragmatic, etc.) that the podcast should incorporate.

Design

Subsequently, the students prepare a preliminary draft of the podcast, which consists of a written composition on the selected topic, incorporating all the specifications outlined in the previously devised plan. Upon the completion of this preliminary task, a second consultation with the instructor is conducted to implement necessary revisions and discuss potential enhancements. Following this initial round of editing, referred to as the 'first coat of paint,' the student will revise the draft accordingly and adapt it into a dialogue between the podcaster (the instructor) and the guest (the student/s). Once the dialogue has been crafted, a subsequent meeting between the teacher and the student will take place, during which the instructor will conduct a second round of revisions, referred to as the 'second coat of paint'. Thereafter, the student will be responsible for implementing the necessary changes and highlighting in color all the linguistic resources identified as requirements during the initial meeting with the instructor. Upon completion of the document, a final review will be conducted by the instructor, referred to as the 'third coat of paint'. Following this review, the podcast transcript²¹ will be ready to proceed to the recording stage. The entire process, from the preliminary tasks to the final transcription, requires a minimum period of one month for completion.

Execution

Students will have time to familiarize themselves with the transcript prior to the recording day. On that day, several rehearsals are conducted to enhance the student's pronunciation and fluency. The recordings are made utilizing the standard sound recording software available on most computers. Given the limited availability of high-definition technical resources, the recordings were conducted using only a laptop and a microphone of moderate quality. Each podcast recording typically requires at least one hour. On recording days, multiple students are often scheduled, as these extracurricular meetings foster collaborative learning and serve as 'motivational boosts.' Thus, positive rapport is established among the students and with the instructor, resulting in favorable impacts on their performance in the Spanish course in which they are enrolled.

The editing of the podcasts has been carried out using *Adobe Premiere Pro 2020*. The application of this software requires specialized expertise and extensive practice to ensure that the final output adheres to professional standards. A *chain-learning* system was adopted, wherein a student with near-professional proficiency in the software transmitted their expertise to a group of students, who then disseminated their knowledge to others. Each podcast editing process has required approximately three hours.

Dissemination

Finally, the podcast is published across multiple social media platforms. All episodes are accessible on *Spotify*²², *YouTube*²³, *LinkedIn*²⁴ and *Instagram*²⁵. The production of each podcast

²¹ All episode transcripts begin with the same introduction and conclude with the same closing. This design aims to provide uniformity and consistency throughout the entire project.

²² <https://open.spotify.com/show/6ChwNmVF2EcVymqo804iQB>

²³ <https://www.youtube.com/@podeledespanol>

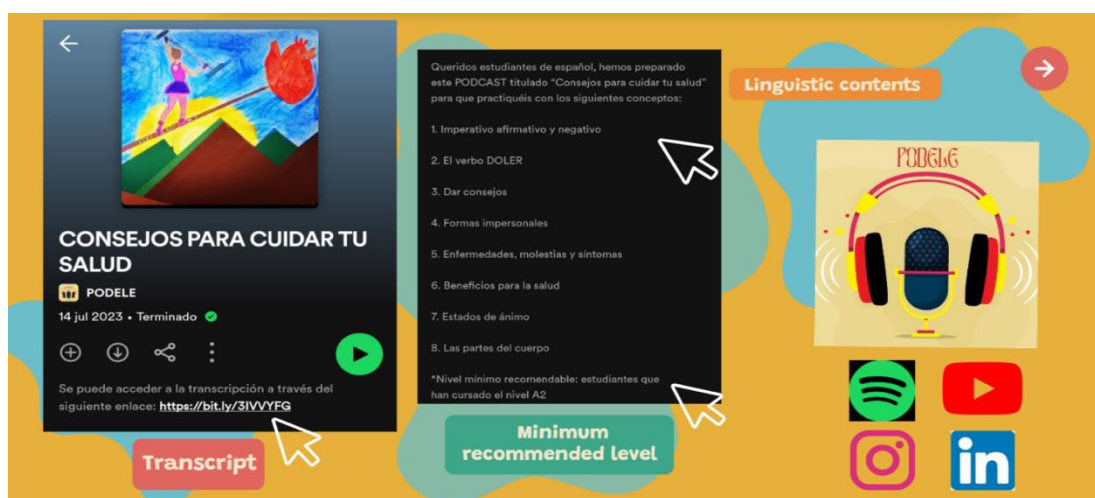
²⁴ <https://www.linkedin.com/company/101533075/admin/page-posts/published/>

²⁵ <https://www.instagram.com/podeledespanol?igsh=MThmMjB5enRtYzV6cQ==>

requires a minimum commitment of 30 hours, comprising 15 hours from the instructor and an additional 15 hours from the student.

Figure 2 illustrates the overall presentation of the podcasts on the most widely used social platform, Spotify. Each podcast is accompanied by a representative photograph corresponding to its theme. Additionally, the transcript of each podcast is uploaded, providing listeners with the opportunity to review the content. A list of the primary linguistic features addressed in the podcast is also included, with these features highlighted in different colors within the transcription. Furthermore, the minimum recommended proficiency level required for students to comprehend the podcast is provided.

Figure 2: Presentation of the Podcast on Social Media



Source: Author's own elaboration

3. Data Analysis

Throughout the entire process of creating the podcasts, each student maintained a reflective *learning diary*. In these diaries, students were afforded the opportunity to document their *experiences, reflections, perceptions, and emotions* related to their learning journey. Below, we present selected excerpts from these learning diaries, along with personal interviews conducted following the publication of the podcasts. From this data, we seek to ascertain the extent to which the project has facilitated the achievement of our predetermined objectives. To enhance the comprehensibility of the collected data corpus, we have categorized it into 'achievements' and 'challenges'. The achievements have been further organized according to the specific objectives the project sought to attain. Only the most representative ideas have been selected, with redundant comments being intentionally omitted.

- *Achievements*

The achievements reported by the students are particularly related to the development of their communicative competencies, the improvement of their linguistic skills, the increase in motivation, the reduction of anxiety and the utilization of new technologies.

Comments related to communicative competence and communication skills

Oral Expression

“I reckon my oral expression has become more natural because, while prepping for my podcast, I picked up on the expressions and structures that Spaniards use day in and day out.”

“Activities like this really help us focus on oral expression and listening, which are usually the areas where we students struggle the most.”

Fluency

“I feel like I can speak more fluently now because recording the podcast gave me the perfect excuse to stand in front of the mirror and practice... In the end, it makes a big difference when you have to improvise.”

Vocabulary

“I don’t usually have time to dive into things I really enjoy... But creating the podcast let me focus on learning Spanish vocabulary around a topic that interests me.”

Grammar

“My grammar has definitely improved, especially when it comes to speaking... This only happens if you’re in a learning environment where you get the chance to practice speaking.”

“I feel way more comfortable now when it comes to putting together complex sentences.”

Pronunciation

“I reckon my pronunciation has improved too. I knew the podcast was going to be recorded, so I put in extra effort to sound like a native.”

Reinforcement of Classroom Concepts

“Making the podcast helped me reinforce what I’d learned in class... Just getting it in class isn’t enough... It’s when you try to use it in real situations that it really sticks.”

Multicultural Competence

“Creating the podcast was a great way to dive deeper into Spanish culture.”

Interaction between Teacher and Student

“The interaction between teachers and students isn’t enough... Projects like this are the best way to strengthen those relationships.”

Researching in Spanish, Editing Texts, and Public Speaking

“To create the podcast, I did research in Spanish, something I hadn’t done before... I had to sift through articles in Spanish to pull out the main ideas and then edit them... I’d never spoken Spanish with a mic in my hand before.”

Comments related to the use of new technologies

New Technologies

“I had no clue how to make a podcast before this... Now I’m actually part of the podcasting world.”

Comments related to the development of autonomy and self-directed learning, the increase in motivation and the reduction of anxiety

Learning with Enjoyment

“I’ve improved my Spanish in a fun way. I got to choose a topic I’m passionate about... You learn so much more when you’re actually enjoying what you’re doing.”

“This kind of activity makes learning so much more engaging.”

Overcoming Shyness

“The podcast was a chance for me to tackle my shyness... I’m someone who’s really scared of speaking in public.”

Self-Confidence

“I didn’t think my Spanish was good enough, but after the podcast was published, I realized that with a bit of effort, anything’s possible... I feel more confident now when I have to speak in Spanish.”

“The podcast has definitely boosted my confidence... Normally, I’d be too embarrassed to speak in Spanish, but I’ve shared the podcast with my mates... I’m really proud of what I’ve done.”

Comments related to the creation of educational resources

Sense of Ownership

“I’ve got a podcast now that I can share with everyone. I love showing it off whenever I meet Spanish speakers.”

“After making the podcast, I’ve got something that’ll stay with me forever.”

Setting a Goal

“I’m way more motivated to work when I have a goal; in this case, it was getting the podcast published.”

• *Challenges*

The challenges faced by the students are closely tied to the substantial time investment demanded by a project of such magnitude, as well as to the articulation of emotions and the execution of tasks that they had not previously undertaken.

Writing the transcription of the podcast

“The toughest part for me was turning a regular text into a dialogue... Normally, we’re writing stories, short articles, or everyday conversations... Transcribing a podcast felt like stepping into a completely different role.”

Time

“The podcast takes up a lot of extra time, but in the end, it’s totally worth it because you get a product to share.”

“Recording the podcast took ages. We had to practice loads to get it just right.”

Emphasizing the key aspects

“For me, the hardest bit of making a podcast was picking out the key points of the topic I chose... Making a podcast isn’t just about chatting about something... You’ve really got to highlight the most important ideas.”

Expressing feelings and emotions

“What I found most challenging was expressing my feelings and emotions...”

4. Conclusion

The present study has concentrated on the *dissemination* of the PODELE project within the educational community. This project revolves around the creation of podcasts by students of Spanish as a foreign language within a non-linguistic immersion context. The study has primarily focused on elucidating the *methodology* employed (a project-based approach), the content of the podcasts and the analytical tools utilized (learning diaries and personal interviews) to assess the benefits derived from the project for the students.

This research has allowed us to substantiate that the PODELE project has significantly contributed to enhancing students' *communicative competencies* and *linguistic skills* in the target language, while also promoting their *autonomy* and fostering their capacity for *self-directed learning*. Furthermore, PODELE has embodied an initiative that integrates *new technologies* in the creation of *learning resources*. Additionally, the students who participated in the study reported that their involvement in this project was highly *motivating*, enabling them to gain *self-confidence* and reduce anxiety levels, particularly in tasks related to oral expression.

The project was an *extracurricular activity* undertaken by *volunteer students* who were willing to invest their time and effort, driven by an *altruistic motivation* to create free resources accessible to both students and teachers. Despite facing limited resources, the project has thus far published a total of 20 podcasts.

Acknowledgment

The PODELE project began to take shape in my mind from a seed planted by my students—they expressed a desire to listen to podcasts crafted to match their language proficiency. In this context, I want to highlight the profound importance of nurturing a close relationship with our students so that we can truly tailor our teaching methods to their real learning needs. The words of the Spanish poet Antonio Machado come to mind: "I never sought glory," and indeed, throughout the development of this project, I never imagined writing an article about it, nor did I foresee standing before a conference to speak about it. My only concern was to offer something of genuine value, freely and selflessly, to all students of Spanish as a foreign language, and especially to my own students. At times, good intentions alone do not suffice; yet despite our very limited resources, the students were swept up by this shared, altruistic drive, and together we brought the project to life.

It is for this reason that I wish to express my deepest and most heartfelt gratitude to each and every student who contributed their small grain of sand to the realization of this project, and especially to Erteğün Ozan Keleş, whose significant dedication to the technical aspects was invaluable. I am also deeply thankful to Istanbul Technical University for creating the opportunity to share this project and the methodology behind it through the publication of this article.

I sincerely hope that PODELE and this article will serve as a source of great benefit to all those who seek it.

Abbreviations

CEFR: Common European Framework of Reference for Languages

PBL: Project-Based Learning

PODELE: Podcast for Students of Spanish as a Foreign Language

SFL: Spanish as a Foreign Language

TBL: Task-Based Learning

References

- Abdous, M., Camarena, M. M. & Facer, B. R. (2009). "MALL technology: Use of academic podcasting in the foreign language classroom", *ReCall*, no 21, pp. 76-95.
- Bailey, K. M. (1990). "The use of diary studies in teacher education programs", in Jack C. Richards & David Nunan eds., *Second Language Teachers Education*, pp. 215-226. Cambridge: Cambridge University Press.
- Bailey, K. M. & Ochsner, R. (1991). "Diary Studies of Classroom Language Learning: The Doubting Game and the Believing Game", in Eugenius Sadtono ed., *Language Acquisition and the Second/Foreign Language Classroom*, pp. 60-102. Singapore: SEAMEO Regional Language Centre.
- Bolger N., Angelina D. & Rafaeli, E. (2003). "DIARYMETHODS: Capturing life as it is lived", *Annual Review: Psychology*, no. 54, pp. 579-616.
- Borgia, L. (2010). "Enhanced vocabulary podcasts implementation in fifth grade classrooms", *Reading Improvement*, vol.46, no. 4, pp. 263-72.
- Candlin, C. N. (1990). "Hacia la enseñanza de lenguas basada en tareas", *CL & E: Comunicación, Lenguaje y Educación*, no. 7-8, pp. 33-54.
- Carlini, A. L. (2004). "Procedimentos de ensino: escolher e decidir", in A. L. Carlini et al. eds., *Os Procedimentos de Ensino Fazem a Aula Acontecer*, pp. 25-81. São Paulo: Avercamp.
- Centro Virtual del Instituto Cervantes (2008). *Diccionario de Términos Clave de ELE*. Retrieved August 7, 2024, from: https://cvc.cervantes.es/ensenanza/biblioteca_ele/diccio_ele/diccionario/trabajoproyectos.htm
- Chacón, C. & Pérez, C. (2011). "El podcast como innovación en la enseñanza de inglés como lengua extranjera", *Revista de Medios y Educación*, no. 39, pp. 41-54.
- Chaves Yuste, B. & De la Peña Álvarez, C. (2023). "Podcasts' effects on the EFL classroom: a socially relevant intervention", *Smart Learning Environments*, vol. 10, no. 20, pp. 1-18.
- Cooper, P.A. (1993). "Paradigm shifts in designing instruction: from behaviourism to cognitivism to constructivism", *Educational Technology*, vol.33, no.5, pp. 12-19.
- Corpas Viñals, J.; Garmendia Iglesias, A. & García Rodríguez M. E. (2021). *Aula Internacional Plus 1*. Barcelona: Difusión, 2024.
- Corpas Viñals, J.; Garmendia Iglesias, A. & Soriano Escolar, C. (2022). *Aula Internacional Plus 2*. Barcelona: Difusión, 2024.

- Corpas Viñals, J.; Garmendia Iglesias, A. & Soriano Escolar, C. (2021). *Aula Internacional Plus 3*. Barcelona: Difusión, 2024.
- Corpas Viñals, J.; Garmendia Iglesias, A.; Sánchez Madrid, N. & Soriano Escolar, C. (2022). *Aula Internacional Plus 4*. Barcelona: Difusión, 2024.
- Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment – Companion Volume*. Strasbourg: Council of Europe Publishing, 2020.
- Dewey, J. (1916). *Democracy and Education*. New York: Courier Corporation, 2014.
- Fernández Martínez, M. A. & Rodríguez Aviñoá, J. (1994). “Un ejemplo para la clase de español: «Un extranjero en Madrid»”. Retrieved August 9, 2024, from: http://cvc.cervantes.es/Ensenanza/biblioteca_ele/asele/pdf/04/04_0231.pdf
- Fleenor, H. G. & Hodhod, R. (2016). “Assessment of learning and technology: Computer science education”, in V. Wang ed., *Handbook of Research on Learning Outcome and Opportunities in the Digital Age*, pp. 51-78. Hershey, PA: IGI Global.
- Gómez Barrios, F. V. & Palma Velásquez, E. (2020). “El podcast en el desarrollo de las habilidades orales en estudiantes de Inglés como lengua extranjera”, *Educere*, vol. 24, no. 78, pp. 237-251. <https://www.teachingenglish.org.uk/professional-development/teachers/integrating-ict/articles/podcasting-elt>
- Istanto, J. W. (2011). “Pelangi Bahasa Indonesia podcast: what, why and how?”, *Electronic Journal of Foreign Language Teaching*, vol.8, no.1, pp. 371-384.
- Jerez-Naranjo, Y. V. & Garófalo-Hernández, A. A. (2012). “Aprendizaje basado en tareas aplicado a la enseñanza de las telecomunicaciones”, *Revista Científica de Ingeniería Electrónica, Automática y Comunicaciones*, vol. XXXIII, no.3, pp. 1-7.
- Kilpatrick, W. H. (1918). *The Project Method: The Use of The Purposeful Act in the Educative Process*. New York: Teachers College/Columbia University.
- Kukulska-Hulme, A. (2005). “Introduction”, in A. Kukulska-Hulme & J. Traxler eds., *Mobile Learning: A Handbook for Educators and Trainers*, pp. 1-6. London: Routledge, London.
- Lankshear, C. & Knobel, M. (2004). *A Handbook for Teacher Research: From design to Implementation*. Berkshire, England: Open University Press.
- Lee, M.; McLoughlin, C. & Chan, A. (2008). “Talk the talk: Learner-generated podcasts as catalysts for knowledge creation”, *British Journal of Educational Technology*, no. 39, pp. 501-521.

- Lomicka, L. & Lord, G. (2011). "Podcasting -past, present and future: Applications of academic podcasting in and out of the language classroom", in B. R. Facer & M. Abdous eds., *Academic Podcasting and Mobile Assisted Language Learning: Applications and Outcomes*, pp. 1-20. Hershey, PA: Information Science Reference.
- McMinn, S.W. J. (2011). "Podcasting possibilities: Increasing time and motivation in the language learning Classroom". *Proceedings of European Institute for E-learning. Learning Forum*. Berlin, Germany, pp. 212-215.
- Nadig, A. (2021). "Listening comprehension", in F.R. Wolkman ed., *Encyclopedia of Autism Spectrum Disorders*, pp. 2726-2727. Germany: Springer International Publishing
- Naismith, L.; Sharples, M.; Vavoula G. & Lonsdale P. (2004). *Literature Review in Mobile Technologies and Learning*. Bristol: Nesta Futurelab.
- Phillips, B. (2015). "Empowering students: Using technology-enhanced learning to foster learner autonomy", in A. Trink ed., *Pannonia Research Award: Regionale und Europäische Zukunftsfragen*, pp. 83-94. Graz: Leykam.
- Piñeiro Otero, T. (2012). "Los podcast en la educación superior. Hacia un paradigma de formación intersticial", *Revista Iberoamericana de Educación*, no. 58, pp. 1-12.
- Putman, S. M. & Kingsley, T. (2012). "The atoms family: Using podcasts to enhance the development of science vocabulary", *The Reading Teacher*, vol. 63, no. 2, pp. 100-108.
- Reese, H. W. (2011). "The learning-by-doing principle", *Behavioral Development Bulletin*, vol. 11, pp. 1-578.
- Rosell-Aguilar, F. (2007). "Top of the pods—In search of a podcasting 'Podagogy' for language learning", *Computer Assisted Language Learning*, vol. 20, no. 5, pp. 471-492.
- Rosell-Aguilar, F. (2013). "Podcasting for language learning through iTunes U: The learner's view", *Language Learning & Technology*, vol. 17, no. 3, pp. 74-93.
- Ryan, S. M. (1997). "Preparing learners for independence: resources beyond the classroom", in P. Benson & P. Volley eds., *Autonomy and Independence in Language Learning*, pp. 215-224. London: Longman.
- Schmidt, R. & Frota, S. N. (1986). "Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese", in: R. Day ed., *Talking to learn*, pp. 237-326. Rowley, MA: Newbury House.
- Stanley, G. (2005). "Podcasting for ELT. British Council- BBC". Retrieved August 8, 2024.
- Sze, P. M.-M. (2006). "Developing Students' Listening and Speaking Skills through ELT Podcasts". *Education Journal*, vol. 34, no. 2, pp. 1025-1936.

- Tippelt, R. & Lindemann, H. (2001). El Método de Proyectos. Retrieved August 9, 2024, from: <http://www.halinco.de/html/doces/Met-proy-APREMAT092001.pdf>
- Tobón, S. (2006). Método de Trabajo por Proyectos. Madrid: Uninet.
- Tomé Díez, M. & Richters M. A. (2020). “Podcasting as a tool to develop speaking skills in the foreign language classroom”, *The EUROCALL Review*, vol. 28, no. 1, pp. 40-56.
- Vizcaíno Timón, I. M. (2008). Guía fácil para programar en Educación Infantil (0-6 Años). Trabajar por Proyectos. Madrid: Wolters Kluwer.
- Wiley, D. A. (2002). “Connecting learning objects to instructional design theory: A definition, a metaphor, and a taxonomy”, in D. A. Wiley, ed., *The Instructional Use of Learning Objects*, pp. 3-25. Bloomington, Indiana: Agency for Instructional Technology and Association for Educational Communications & Technology
- Wolk, S. (1994). “Project-based learning: Pursuits with a purpose – educational leadership”, *ASCD*, vol. 52, no. 3, *Strategies for Success* 52, pp. 42-45.
- Yoestara, M. & Putri, Z. (2019). “PODCAST: an alternative way to improve EFL students’ listening and speaking performance”, *English Journal*, vol. 6, no.1, pp. 15-15.
- Zábráczki, A. (2013). Aprender español vía proyectos en niveles avanzados: una experiencia docente”. *Actas del I Congreso Internacional de Didáctica de Español como Lengua Extranjera del Instituto Cervantes de Budapest*. Budapest, Hungría, pp. 323-327. Retrieved August 9, 2024, from: http://cvc.cervantes.es/ENSENANZA/biblioteca_ele/publicaciones_centros/PDF/budapest_2013/36_zabraczki.pdf
- Zanón, J. (1995). “La enseñanza de lenguas extranjeras mediante tareas”, *Signos. Teoría y Práctica de la Educación*, vol.14, pp. 52-67.
- Zanón, J. (1999). “La enseñanza del español mediante tareas”, in J. M. Alba et al. eds., *La Enseñanza del Español mediante Tareas*, pp. 13-24. Madrid: Edinumen.

Teachers' Beliefs and Practices of Vocabulary Teaching in EFL Classrooms

Melani Konca*

*Teacher, İstek Schools, Türkiye

Abstract

The importance of vocabulary teaching practices in EFL classrooms has been stressed in many studies on language teacher cognition (e.g. Borg, 2003; Schmitt, 2008; Li, 2020). However, this area of research still lacks studies on the triangulation of vocabulary teaching, classroom practices, and teacher cognition in Türkiye. Therefore, this research aims to explore teachers' beliefs and practices regarding their own vocabulary pedagogy by shedding light on the relationship between the two. The participants of this study were six English language teachers at the researcher's internship school, which was a public high school in Istanbul. In order to illustrate the profile of the participants, the researchers first gathered demographic information. Data were collected through classroom observation, interviews with teacher participants, and all materials given to students. Classroom observations and interpretation of materials were done throughout the data collection process and interviews were conducted depending on the common availability of the teachers and the researcher. The participants were asked ten open-ended questions during the interview. Therefore, this study is a qualitative case study focusing on a group of teachers' beliefs and practices. The findings indicated that teachers' beliefs influence their classroom practices during vocabulary teaching. Also, it has been found that teachers' own interest in learning new vocabulary has a positive impact on their vocabulary teaching practices. However, despite their beliefs, there were some constraints such as time which hindered teachers' classroom practices. Therefore, it is not realistic to expect teachers' beliefs and practices to be in line with each other all the time.

Keywords: Teacher cognition, teacher beliefs, vocabulary teaching, classroom practices

1. Introduction

Language teacher cognition has been an area of research in ELT with numerous studies (i.e Borg, 2003; Li, 2020). Similarly, classroom practices of English language teachers have also been investigated through both qualitative and quantitative methods research, according to my literature review. When it comes to vocabulary teaching, on the other hand, most recent studies focus on vocabulary strategies but not classroom practices. There might be tips or ways to teach vocabulary to enhance students' lexical skills and communicative competence (i.e Ur). However, this research aims to use and focus on the term “classroom practices” as it is more appropriate for this particular study's design.

Without grammar very little can be conveyed, without vocabulary, nothing can be conveyed (Wilkins, 1972). This quote by one of the famous applied linguists Wilkins can be supported by the suggestion that to a great extent, students' skills of listening, speaking, reading, writing, and translating are all influenced by vocabulary (Pan & Xu, 2011). Stating the importance of vocabulary teaching practices in EFL classrooms and research on language teacher cognition, still, this area of research lacks studies on the triangulation of these mentioned concepts in Türkiye. Therefore, this research aims to emerge and connect teachers' beliefs and practices regarding their own vocabulary pedagogy.

1.1 Background to the Study

Vocabulary teaching has been an area of research since the development of the term “communicative competence” throughout the English language teaching history. Yet, little research has been conducted on teachers' cognition of vocabulary teaching (Borg, 2006). Language learners' and teachers' cognition in respect of learning and teaching plays a critical role in mediating their actual behavior and decisions in the process (Gao & Ma, 2011). Likewise, as for language teacher cognition, individuals' teaching beliefs may ‘consist of a set of personally defined practically oriented understandings of teaching and learning which exert a significant influence on instructional decisions’ (Borg, 1999). As Borg (2006) mentioned that little research has been conducted on teachers' cognition of vocabulary teaching, this research will combine language teachers' cognition of vocabulary teaching and classroom practices in order to provide the audience with a clear distinction between stated beliefs versus observed beliefs alongside stated practices versus actual classroom practices.

1.2 Statement of the Problem

In Türkiye, a high number of schools, both private and public, tend to have a curriculum for English language lessons that has the most emphasis on grammar and vocabulary has its small place integrated with the grammar-based lessons. Here, occasionally, vocabulary is being taught with decontextual and traditional practices such as word lists and memorizing the Turkish equivalents. This unfortunate situation can be observed less in private schools, acquiring more skills and lexis-based curricula such as international ones. However, in both contexts, teachers are bound to follow a certain curriculum, and most of the time it is highly questionable that their classroom practices are in accordance with their own beliefs regarding vocabulary pedagogy. This is an ongoing problem in schools in Türkiye and thus, this particular study dwells on the relationship between English language teachers' beliefs and classroom practices of vocabulary teaching.

1.3 Purpose of the Study

Contextually, the purpose of this study is to put two separately investigated terms “language teacher beliefs” and “classroom practices” together and connect them with vocabulary teaching in EFL classrooms in Türkiye. Practically, on the other hand, the core aim of this research is to shed a light on and have a better understanding of the relationship between what EFL teachers believe and what they do in their lessons.

1.4 Research Question

The following research question has been put forward in order to frame and consolidate the meaning behind this study:

1. How do K-12 EFL teachers perceive vocabulary teaching in terms of their own classroom practices in relation to vocabulary pedagogy?

1.5 Significance of the Study

Meara (1996) suggested that lexical competence is at the heart of communicative competence and according to Verhallen & Schoonen (1998), it can be a prediction of school success. Additionally, although teachers of English as a second language are often left to their own devices in teaching vocabulary in the absence of clear guidelines, forced to rely on their own beliefs and the teaching materials provided, only a handful of studies to date have focused on what they know and believe about second language vocabulary teaching and learning (Chung, 2018). Regarding this statement by Chung (2018), the biggest significance of this particular study is to contribute to this insufficiently researched area in ELT.

1.6 Definitions of Terms and Abbreviations

Teacher Cognition: unobservable cognitive dimension of teaching – what teachers know, believe, and think (Borg, 2003)

Vocabulary: the body of words used in a particular language

Lexis: all the words in a language, the entire vocabulary of a language (Barcroft, Sunderman, and Schmitt, 2011) (retrieved from the 2012 Ph.D. study by Nural)

EFL: English as a foreign language

ELT: English language teaching

SLA: Second language acquisition

2. Literature Review

2.1 Language Teacher Cognition and Classroom Practices

Throughout literature, the concept of teacher cognition has been a complex concept and studied from various points of views. As Li (2020) states a strong theme across different disciplines and research paradigms is to understand what constitutes ‘cognition’ and to argue its importance for effective pedagogy and teacher development. By observing what teachers do and say in their professional context, we are able to make sense of what they think and how that thinking is displayed in their practice.

Although the concept of belief has attracted considerable research interest in education in recent years as can be seen in the previous studies mentioned, there is a lack of consensus about what the term denotes (Borg, 2003, Borg, 2001, Woods, 1996). For example, Burns (1992: 58) describes the beliefs that motivate instructional practices in the classroom. Johnson (1992) reports that beliefs tend to shape teachers’ instructional practices. Borg (2001) proposes that beliefs guide teachers’ thoughts and behaviors. Finally, in his review article, Borg (2003) refers to it as the unobservable cognitive dimension of teaching – what teachers know, believe, and think. He illustrates this concept thoroughly in Figure 1:

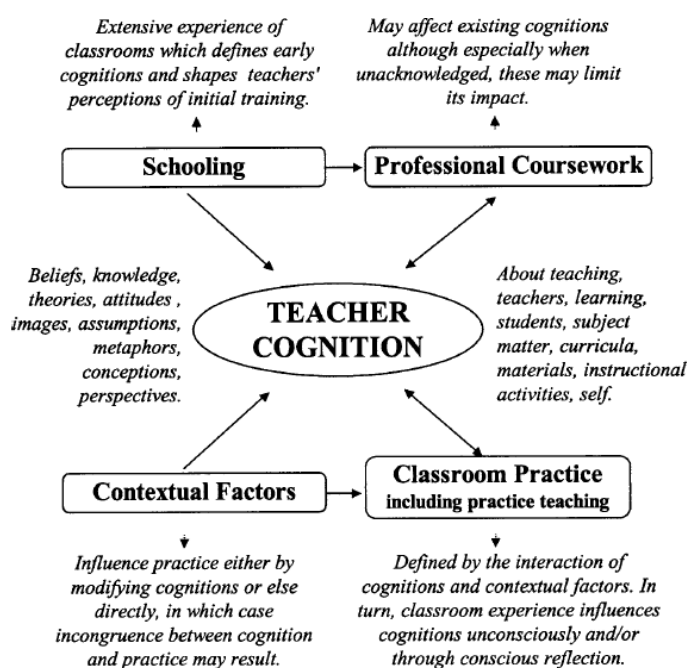


Figure 1 Teacher cognition, schooling, professional education, and classroom practice (Borg 1997)

Chung (2022) presents another perspective on language teachers’ beliefs by quoting that beliefs not only determine how individual teachers organize and define various tasks and problems but also dictate their behavior in the classroom (Xu, 2012). Obtaining a clear understanding of teachers’ beliefs is thus vital, in that it will help them determine the meaning of situations they encounter and, in turn, encourage them to think reflectively and embrace change when necessary (Chung, 2022).

When it comes to connecting teachers' cognition and their classroom practices, Nural (2012) points out that the correspondence between teachers' cognitions and practices is, however, not always straightforward. It is not realistic to expect teachers' beliefs and practices to be always congruent with one another due to possible reasons and constraints such as curriculum and syllabus requirements, coursebooks, examinations, students, colleagues, or other social expectations. Since these factors are mostly outside the control of teachers, they may not immediately put their beliefs into practice. Supporting this statement by Nural (2012), another study by Graden (1996) investigated the stated beliefs and practices of teachers of reading instruction. The study found that although many beliefs guided their practice, mismatches between their stated beliefs and practice were evident in observations of their lessons. The teachers these in terms of constraints, such as time and lack of appropriate materials (Basturkmen et al., 2004).

Overall, the relationship between teachers' beliefs and practices suggests that theory and practice are not separated. This is of great importance to language teachers. In other words, any attempt to change teachers' theories and beliefs about teaching and learning would have consequences for their actual teaching in classrooms. Thus, teachers can be encouraged to stay up-to-date with theories of learning in order to establish beliefs that can lead to their best practice in the classroom (Mardali & Siyyari, 2019).

In sum, teacher cognition has been a complicated concept in literature, and observing teachers' behavior and practices are what makes us understand better what "cognition" means. Although the concept of teacher beliefs has been studied in many research, there is still a lack of consensus on what it means. Different researchers have put forward different thoughts. On the other hand, it is not realistic to expect teachers' beliefs and practices to be in line with each other because of some constraints such as time, examinations, and curriculum. Yet, teachers' beliefs and practices are inseparable and the change in teachers' beliefs directly affects their classroom practices.

2.2 Vocabulary Teaching

"Lexis can be regarded as the most personal or individual dimension of language learning and teaching as it is largely contingent on learners' and teachers' mental dictionary." (Nural, 2012) Thus, teaching lexis inevitably becomes a different kind of responsibility that has to do with their own lexical curiosity, as well. In vocabulary instruction, EFL teachers' interest in words can have a big role in not only the teacher's own vocabulary teaching practices but also in students' own in-class and out-of-class vocabulary acquisition practices as well. On the other hand, regarding vocabulary teaching, teachers have numerous other responsibilities that emerge from different difficulties. For instance, although native speakers naturally acquire vocabulary through contextualized input over many years, second language learners often need to learn a large amount of vocabulary in a limited period of time; thus, vocabulary instruction is of fundamental concern to educators. Language teachers are faced with decisions regarding vocabulary assessment, instructional priorities, teaching techniques, vocabulary learning strategies, and resources, among others (Rossiter, Abbott, & Kushnir, 2016). Therefore, especially in EFL classrooms, vocabulary instruction of the teachers has a vital role in students' lexicon development and communicative competence. This vital role of vocabulary teaching in EFL classrooms can be consolidated by examining teacher beliefs and classroom practices

alongside these two factors being in accordance with each other which is what this particular study aims to explore.

There has been so much vocabulary research published in literature so far, and therefore it is difficult to decide on a manageable number of pedagogical suggestions. However, some principles by Hunt and Beglar (1998) still provide a very good start toward summarizing a principled approach to vocabulary learning:

- Provide opportunities for the incidental learning of vocabulary.
- Diagnose which of the 3000 most common words learners need to study.
- Provide opportunities for the intentional learning of vocabulary.
- Provide opportunities for elaborating word knowledge.
- Experiment with guessing from context.

In addition to these, learners need large vocabularies to successfully use a second language, so high vocabulary targets need to be set and pursued. In addition, teachers should make sure that learners maintain the maximum amount of engagement possible with lexical items (Schmitt, 2008).

To conclude, teachers' own interest in learning new vocabulary has a big impact on their classroom practices and behavior toward students' questions about vocabulary. Teachers have many responsibilities regarding vocabulary teaching in the classroom because of numerous difficulties such as limited time. Therefore, in order to ease these responsibilities, numerous pedagogical suggestions have been put forward to develop teachers' vocabulary practices such as providing students with the opportunity for incidental and intentional vocabulary learning.

3. Methodology

3.1 Research Design

This qualitative research is designed as a descriptive case study that aims to examine EFL teachers' beliefs and practices of vocabulary teaching. It is descriptive as it focuses on observations and detailed interviews with the participants. This particular research is also a case study because it focuses on in-depth descriptions of language teacher cognition within a small group of EFL teachers in a particular school.

3.2 Setting and Participants

The setting of this study is the researcher's current internship school, a public high school in Istanbul, Türkiye. The subsetting is language classrooms and they will be used for observational purposes.

Participants are the EFL teachers in the English language department of the internship school. Students enrolled in language classrooms are not included in this study specifically, as the study focuses on teachers' perspectives on vocabulary teaching. Demographic information about teacher participants is shown in the table below.

The study is conducted both online and in person. Depending on the common availability of the participants and the researcher the interviews are done at the internship school or via Zoom.

Participant	Teaching Experience	EFL Qualifications	Type of School	Grades Taught	Gender
Participant A	26 years	BA in ELT	State School	Middle school and high school	Female
Participant B	22 years	BA in ELT	State School	K12 and adults	Female
Participant C	24 years	BA in ELT	State School	K12 and adults	Female
Participant D	20 years	BA in ELT	State School	High school	Male
Participant E	22 years	BA in ELT	State School	High school	Female
Participant F	16 years	BA in ELT	State School	Middle school and high school	Female

3.3 Data Generation and Data Analysis

Data are generated through mainly observation and interviews for this qualitative study. The researcher first starts with in-class observations. The observation notes include all events related to vocabulary teaching such as vocabulary presentation, vocabulary practice, input, output, intake, and uptake. The researcher continues the data generation process with semi-structured interviews with each participant. This particular type of interview is chosen to open up space for the participants to elaborate on questions developed naturally during the interview. As for artifacts; exams, lesson plans, curriculum, and in-class materials such as worksheets are used.

When it comes to data analysis, “thematic analysis” is conducted for this study because this particular type of analysis is preferable when it comes to underresearched fields (Braun & Clarke, 2006). Qualitative data generated from the participants will become “codes” through coding, and then the codes will become “themes” through constant comparison (Braun & Clarke, 2006).

3.4 Trustworthiness

There are several crucial points in order to ensure the trustworthiness of this study. First, the researcher does prolonged participation and observation at the study site besides establishing a structural coherence, when it comes to credibility (Mills & Gay, 2012). Second, the researcher collects detailed descriptive data and practices triangulation to make sure that the criteria for transferability, dependability, and confirmability are reached (Mills & Gay, 2012). To ensure the trustworthiness of the interview done with the participants, in-class observation notes are taken.

3.5 Ethical Issues and Limitations

The primary point regarding ethical issues is the informed-consent. Informed-consent forms are obtained from the teacher participants and their anonymity is preserved throughout the study. The name of the school is also kept confidential. In addition, prior to data generation, the teacher participants are informed about the aims of this particular study. Secondly, as another ethical issue, consent is provided from both teacher participants and the principal of the school in order to have permission to use in-class materials (worksheets, etc.), curriculum, lesson plans, and exams as data for this research.

The study is partly limited as it is conducted with approximately six EFL teachers in a particular educational setting, and the findings cannot be generalized. Additionally, this study does not take students' perceptions into account. Therefore connections and comparisons between EFL teachers' beliefs and practices of vocabulary teaching, and students' perceptions of vocabulary learning cannot be made.

4. Findings and Discussion

4.1 Findings

4.1.1 Vocabulary Presentation and Practices in Coursebooks and Exams

The ELT department at the school uses different coursebooks from Oxford Publishing and National Geographic Learning: *Life Vision Student Book & Workbook B1*, *English File Student Book & Workbook Upper Intermediate*, *Reflect Reading & Writing B1* and *Reflect Listening & Speaking B1*. The common features of these coursebooks in relation to vocabulary presentation and practice include an "Irregular verb list" at the end of each book and there is a "Vocabulary index" in the books in which academic vocabulary is specifically marked and the CEFR level of each word is written. In addition, each coursebook has a "vocabulary booster" part at the end in which there are activities related to vocabulary only. When it comes to typical activity types in these coursebooks, there are ones in which students match the words with definitions, complete the text with the words or phrases, complete the sentences with the correct form of the words, match words with pictures, choose the correct word to complete sentences between two words.

As for the exams conducted at the school, students regularly have quizzes at the end of each unit in their coursebooks. The quiz questions are included in the Teacher's Book and they are always multiple-choice tests. They have three parts: grammar, vocabulary, and pronunciation. In the vocabulary part, students underline the correct word between two or three words, complete the compound words in the sentences, choose the correct words for definitions, complete the sentences with the correct words, and correct the underlined errors in the sentences.

4.1.2 Vocabulary Presentation and Practices in the Classroom

According to the researcher's observation notes, the English lessons at the school do not have a specific vocabulary part. Students only focus on vocabulary throughout the unit by doing the short activities mentioned above. They also study keywords before a reading passage.

When it comes to how the teacher participants deal with vocabulary in their lessons, Participant A directly tells the English equivalent of a word when students ask. The following is an example of a dialogue between one of the students and the teacher:

S: What is the English word for “nargile”?

T: It’s hookah.

Besides, Participant A does not hesitate to use a dictionary when a student asks for a definition of a word. For instance,

S: What does “deception” mean?

T: I don’t know the exact definition so let me check the dictionary.

Participant B prefers to give synonyms of the newly-learned words in order to explain the definition better. For example, in a lesson, she gave the synonym of the word “qualification” as an explanation. While introducing the new vocabulary, Participant B frequently writes it on the board and also adds different forms of it. For instance, she wrote the word “consider” and then “take something into consideration”. Participant B occasionally introduces new vocabulary related to the highlighted words in the coursebook. For example, students were studying different events in life such as starting a family, and the teacher introduced the word “milestone”. While explaining newly-learned vocabulary, the teacher frequently gives examples from daily life. For instance, in order to explain the word “community” she said that “The people at Beşiktaş Square are not a community but if they are supporting the same football team, then they are a community.” Participant B also likes giving examples of cognate word usage such as “effort” in English and “efor” in Turkish. The teacher occasionally informs students about cross-linguistic similarities between different languages such as the borrowing of English words from the French language. Finally, Participant B sometimes prefers to explain a newly-learned vocabulary such as “attack” through physical movements.

Participant C tends to only read the definitions of the words in the glossary part of the coursebook. She does not give any example sentences. She also prefers to explain the definition of a word by the antonym of it and a realia. For example,

S: What does “unused” mean?

T: (showing an eraser of a student) Look! This eraser is used. If it was not used, we would call it unused.

Participant C does not have the habit of showing extra vocabulary related to the highlighted words in the classroom.

Participant D frequently presents new words about the relevant topic such as “time-consuming, self-diagnose, and virtual care”. He also presents the alternatives of words such as instead of “pale”, using “off-color”. In addition, while studying a paragraph, the teacher always writes the words that he thinks might be unfamiliar to students on the board. Participant D frequently uses synonyms to explain vocabulary as Participant B does. For instance,

S: What does “approximately” mean?

T: It means “about”. For example, “There are about fifteen students in the classroom.”

Participant E does very little about developing students' vocabulary knowledge. She frequently gives all vocabulary exercise parts in the coursebook as homework and expects the students to study the words by themselves and just practice the vocabulary in the classroom. During the lesson, for example, if they are checking a fill-in-the-blanks activity, she asks the students “What meaning does the word give to this sentence?” Besides this, She sometimes gives extra words such as similar ones or asks the students to give examples of similar words such as “notice” and “realize”.

Participant F usually suffices with the vocabulary exercises in the coursebook. For instance, in one activity, she asked the students to guess and match the meanings of the highlighted words and then listened to the answers and checked. In one lesson, however, the teacher did a storytelling activity that was not in the coursebook. She showed a picture of a man walking on a road and asked the students to come up with and write a short story related to the picture. Other than this, she does not give extra words related to the highlighted vocabulary in the coursebook.

4.1.3 Interview with Teacher Participants

4.1.3.1 The Role of Vocabulary in SLA

All teacher participants agree that vocabulary has a key role in language learning and teaching. Half of the six participants give more importance to teaching collocations because they think it is important for students to know which words go together. According to the teachers, the reasons why teaching vocabulary is important are organizing sentences, improving speaking skills and writing skills, and expanding points of view.

4.1.3.2 Views and Beliefs on Vocabulary Teaching

All participants agree that vocabulary should be taught within the context and real-life examples such as songs, films, etc. The reason is that they all believe “knowing a word” means knowing how to use it in a meaningful context.

The most common three pieces of advice by teachers to students are reading books, spending time with the words, and keeping a vocabulary notebook. Half of the teachers prefer to give extra vocabulary related to highlighted words in the classroom. All teachers frequently use online tools to teach vocabulary, but two of them think they only catch students' attention and do not directly affect students' learning. One of the teachers prefers to teach chunks of vocabulary.

When it comes to how the participants' beliefs on vocabulary teaching are shaped, the most typical answer is their own learning and teaching experiences. However, there are some different factors such as the platform that they teach (e.g. Zoom), the videos of polyglots on Youtube, teacher forums, seminars, and articles.

4.1.3.3 Effective Strategies and Materials for Vocabulary Instruction

All participants agree that several strategies should be combined rather than applying only one strategy. The different strategies conducted by teachers are introducing new words while speaking to them, advising them to read a lot, showing them videos about vocabulary, using

online tools, motivating students to create their own sentences and paragraphs, and reading articles. However, all teachers think that the teaching in the classroom is not enough and students should practice and repeat the vocabulary at home by using the most suitable strategies for them. Some of the typical practices by teacher participants are asking WH questions, using gamification tools, asking the students to explain a word's definition in their own words, and doing traditional T/F, fill-in-the-blanks, and matching activities in the coursebooks.

When it comes to materials preferred by teacher participants are a piece of paper and pen, online tools, mobile phones, coursebooks, the online practice offered by several publications, reading books, and worksheets.

4.1.3.4 Designing Class Activities for Vocabulary Development

Regarding what guides the design of vocabulary development activities, participants mentioned several factors such as the level and age of the students, the relevance and popularity of the topic, and the context. Some of the different factors are the CEFR level of the words and preparing extra vocabulary for students.

4.2 Discussion

The findings of this study suggest that the teacher participants attach high importance to vocabulary teaching because all of them unanimously indicated that lexical knowledge is crucial for organizing sentences, improving speaking skills and writing skills, and expanding points of view. Because they are aware of the importance of vocabulary teaching, as previously suggested by Mardali and Siyyari (2019), most of the participants try to catch up with the new approaches to vocabulary in order to stay up-to-date and develop their classroom practices. Overall, the teacher participants' classroom practices are in accordance with the principles suggested by Hunt and Beglar (1998). Teachers provide opportunities for the incidental learning of vocabulary by assigning short story books to students and they also provide opportunities for intentional learning through the vocabulary practices in the coursebook. The coursebook has the Oxford 3000 most common words list and the teachers diagnose which words learners need to study from that. During their lessons, most of the participants give the students the chance to elaborate on a word's meaning using their own words. Students also experiment with guessing from the context before each reading task. In addition to these, most of the teachers present large vocabularies to the students so that they can successfully use a second language. Therefore, the teachers set and pursue high vocabulary targets. Finally, by providing the students with extra online materials, teachers make sure that learners maintain the maximum amount of engagement possible with lexical items.

The findings of this study are consistent with those of Chung (2022) who asserted that the teachers' beliefs not only determine how individual teachers organize and define various tasks and problems but also influence their behavior in the classroom. Those participants who said that their beliefs are shaped by other language learners, their own educational experiences, and teaching experiences opt to offer extra words to students instead of sticking to what is given in the coursebook. They also behave positively while satisfying the students' curiosity about unfamiliar words. Similarly, a strong relationship between teachers' lexical curiosity and vocabulary teaching has been reported in the literature by Nural (2012). The current study found

that those participants who stated that they themselves love learning new vocabulary give more importance to vocabulary instruction during their lessons compared to other participants.

Although most of the participants try to improve their students' vocabulary by giving extra words, as pointed out in the studies by Nural (2012) and Graden (1996), the correspondence between teachers' cognitions and practices is not always straightforward. Since many factors are mostly outside the control of teachers such as curriculum, examinations, and time, they cannot immediately put their beliefs into practice as some of the participants stated that the duration of the lesson is not enough to practice vocabulary, so students should study at home. This finding is also similar to Rossiter, Abbott & Kushnir's (2016) findings which showed that although native speakers naturally acquire vocabulary through contextualized input over many years, second language learners often need to learn a large amount of vocabulary in a limited period of time; thus, vocabulary instruction is of fundamental concern to educators. Thus, the practice done outside of the classroom has a vital role in the students' lexical development.

This research has been done in an internship context where the researcher's current internship school and English language teachers have been used as the setting and the participants. Therefore, being different from previous studies that were conducted in university contexts and generally written by doctoral students, this study is implemented through an intern's eyes, reflecting the humble perspective of the researcher's one-year experience in the internship school.

5. Conclusion

In Türkiye, English language teachers in both public schools and private schools are supposed to follow a certain curriculum, and most of the time the curriculum does not include enough vocabulary presentation and practices, instead, it focuses on grammar for the following high school and university entrance exams. Thus, it is uncertain that teachers' classroom practices are in accordance with their own beliefs regarding vocabulary pedagogy. As this problem still continues, this particular study dwells on the relationship between English language teachers' beliefs and classroom practices of vocabulary teaching.

The findings of this study indicate that all teacher participants give great importance to vocabulary teaching, and they try to expand students' perspectives on vocabulary by presenting extra words. Besides, it is concluded that teachers' beliefs and practices of vocabulary teaching are aligned with each other most of the time. However, some constraints such as time, examinations, and curriculum can become obstacles for teachers. Therefore, it is not realistic to expect the teachers' beliefs to be reflected in their classroom practices all the time. In addition, according to the findings, there is a positive correlation between teachers' own lexical interests and classroom practices. Those teachers who are curious about words tend to present more vocabulary and have a more positive attitude toward students' questions about words.

Overall, no matter how qualified the coursebooks are, the most important factor in vocabulary teaching and learning is the teachers' beliefs. One little change in the beliefs directly affects their classroom practices. Therefore, it is vital for teachers to stay up-to-date regarding teaching approaches in order to implement the best practices in the classroom while teaching vocabulary.

Appendix: Interview Questions

1. Could you give information about your educational and professional life? What did you study at university? How long have you been a teacher? What grades have you taught? Which schools have you worked at? Private or state?
2. What do you think is the role of vocabulary in second language acquisition?
3. To you, what is meant by “knowing a word”?
4. What are your views on vocabulary teaching?
5. What do you think vocabulary teaching involves?
6. How should vocabulary be taught?
7. What do you think are effective strategies for vocabulary instruction?
8. What should teachers consider when they design class activities for vocabulary development?
9. What materials should teachers use to foster vocabulary teaching and learning?
10. What are some possible factors that shaped your beliefs about vocabulary teaching?

Source: Chung, E. (2018). Revisiting second language vocabulary teaching: Insights from Hong Kong in-service teachers. *The Asia-Pacific Education Researcher*.

References

- Barcroft, J., Schmitt, N., & Sunderman, G. (2011). Lexis. In J. Simpson (Ed.), *The Routledge Handbook of Applied Linguistics* (pp. 571-583). Abingdon, UK / New York: Routledge
- Borg, M. (2001). Teachers' Beliefs. *English Language Teaching Journal* 55/2: 186-8.
- Borg, S. (1999). Studying teacher cognition in second language grammar teaching. *System*, 27 (1), 19-31.
- Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching*, 36(2), 81-109. doi:10.1017/S0261444803001903
- Borg, S. (2006). *Teacher Cognition and Language Education: Research and Practice*. Bloomsbury Academic
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Burns, A. (1992). Teacher beliefs and their influence on classroom practice. *Prospect* 7/3
- Chung, E. (2018). Revisiting second language vocabulary teaching: Insights from Hong Kong in-service teachers. *The Asia-Pacific Education Researcher*.
- Chung, E. (2022). Effecting change in teachers' epistemological and pedagogical beliefs about vocabulary learning and teaching: The role of dialogic reflection. *SAGE Open*. 12. 1-15.
- Dunkin, M.J., Biddle, B.J. (1974). *The study of teaching*. Holt, Rinehart and Winston, Inc., New York.
- Gao, X. & Ma, Q. (2011). Vocabulary learning and teaching beliefs of pre-service and in-service teachers in Hong Kong and mainland China. *Language Awareness*, 20:4, 327-342, DOI: 10.1080/09658416.2011.579977
- Gay, L.R., Mills, G.E. and Airasian, P.W. (2012) *Educational Research: Competencies for Analysis and Application*. 10th Edition, Pearson, Upper Saddle River.
- Hunt, A., & Beglar, D. (1998). Current research and practice in teaching vocabulary. *The Language Teacher*.
- Johnson, K. (1992). The relationship between teachers' beliefs and practices during literacy instruction for non-native speakers of English. *Journal of Reading Behavior* 24/1: 83-108
- Li, L. (2020). *Language Teacher Cognition: A Sociocultural Perspective*. London: Palgrave/Macmillan.

- Mardali, J. & Siyyari, M. (2019). English teachers' beliefs and practices in teaching vocabulary: The case of teaching experience. *Cogent Education*, 6:1, 1686812, DOI: 10.1080/2331186X.2019.1686812
- Meara, P. (1996). The dimensions of lexical competence. *Performance and competence in second language acquisition*. 35. 33-55
- Nural, S. (2012). EAP teachers' cognitions and practices in teaching lexis in two Turkish private universities: an exploratory qualitative study. Unpublished thesis. University of Leicester
- Pan, Q. & Xu, R. (2011). *Vocabulary Teaching in English Language Teaching. Theory and Practice in Language Studies*.
- Rossiter, M. J., Abbott, M. L., Kushnir, A. (2016). L2 vocabulary research and instructional practices: where are the gaps? *TESL-EJ*. 20:1
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research* 12:3. 329-363.
- Verhallen M., Schoonen R., (1998). Lexical Knowledge in L1 and L2 of Third and Fifth Graders, *Applied Linguistics*, 19:4. 452–470.
- Wilkins, D. A. (1972). *Linguistics in language teaching*. Edward Arnold.
- Woods, D. (1996). *Teacher Cognition in Language Teaching*. Cambridge: Cambridge University Press.
- Xu, R. (2012). Lexical approach and its application in college English classes. *Theory and Practice in Language Studies*

Reflections of ESP Students on Online Asynchronous Discussions for an Advanced Speaking Skills Course

Meryem Büşra Ünsal¹ Enisa Mede²

¹ Inst., Fatih Sultan Mehmet Vakıf University, Türkiye

² Prof. Dr., Bahçeşehir University, Türkiye

Abstract

With more integration of online or hybrid instruction and assessment into higher education institutions, Asynchronous Online Discussions (AOD) are getting higher popularity all around the World. This autoethnographic case study was conducted to describe the reflections of vocational school students from a civil aviation and cabin services department and reflections of their EFL instructor about the effectiveness of asynchronous online discussions (AODs) for a fully asynchronous advanced speaking skills course which had been given throughout the spring semester of the 2022-2023 academic year. The data from the students were collected with the entries of initially 22 students who enrolled the course and posted in asynchronous online discussions consecutively within 2nd, 6th and 10th weeks of the spring semester as well as 11 students' replies to an open-ended questionnaire and finally from a focus group discussion with 6 purposefully chosen students. The data from the instructor who was also the researcher of this article was collected from the notes and personal reflections of the teacher-researcher. The findings revealed that this study could contribute to the existing body of literature as it was conducted with intern flight attendant learners from a vocational school in Türkiye for whom English speaking skills are crucial because of their career in aviation sector which gives the highest priority to English language proficiency.

Keywords: Asynchronous online discussions, language learning, online education

1. Introduction

For the last three decades, online classrooms have been preferred by vast majority of learners and instructors, particularly in higher education. With the latest series of crisis in the World such as the pandemic, war and earthquakes, the instant shift to technology integration to the classrooms was made it compulsory for all levels of education. Having experienced the benefits of such instruction and the feedback both as learners and instructors in terms of its practicality, fully online courses and asynchronous discussions became inseparable from educational practices.

Due to the international demand on accessibility of education Worldwide, “asynchronous online discussions (AOD) are ubiquitous in online courses in higher education (Fehrman and Watson, 2021)”. Despite this popularity and functionality of asynchronous online discussions (AODs), changing pedagogical implications according to the needs of the era, and the high-level importance of English for vocational school students, particularly for students who are planning a career in aviation industry, no research, as far as the researcher reviewed the existing literature, was conducted yet in terms of implications of online discussion forums to EFL speaking courses of vocational schools, particularly in Aviation English research.

This autoethnographic study was implemented to report the reflections of a class of EFL learners and observations and notes of their English language instructor who is also one of the researchers of this study from a vocational school department about an advanced speaking skills course, which was planned as a fully online-asynchronous course due to students’ internship agreements with an aviation company and their unavailability to attend face-to-face classrooms throughout the semester. Students’ entries in asynchronous online discussions, their responses to an open-ended questionnaire, and finally a focus group discussion interview analysis as well as the teacher researcher’s reflections in forms of observation and notes were used for reporting the results of the study. The three research questions which were aimed to be answered are given below.

1. In what ways does an online course have an impact on in-service flight attendants’ communicative performance?
2. How do the participants perceive the effectiveness of a course with fully online asynchronous discussions?
3. How does the course instructor perceive the effectiveness of a course with fully online asynchronous discussions?

2. Literature Review

When the existing literature about the online asynchronous discussion (OAD) was reviewed, different implications with different foci within a 21-year-long period were conducted. In Dantas-Whithey (2002), 54 audiotaped journal entries from 18 higher education ESL students were analysed, and students perceived the audiotaped journals as valuable opportunities for oral language practice, non-threatening corrective feedback, and self-evaluation. The same year, Järvelä and Häkkinen (2002) aimed to see the kind of communication that took place between pre-service teachers during web-based conferencing and the level of discussion patterns throughout these web-based conferences based on readings. The results revealed different levels of web-based discussions among learners, and it was seen that different contextual

resources were important to create mutual understanding among learners, with a suggestion that web-based learning should be seen in a wider social context, rather than only for global interactions. Similar to that, Zhu (2006) analysed four asynchronous online discussions to see the types of interaction and levels of cognitive engagement of students. The results proved the engaging nature of online discussions; however, it also showed that engagement in online discussions was mainly possible depending on instructor's course design and integration method of online discussions to the course, rather than the online environment itself. Zhu indicated that "It is unrealistic to simply plunge students into an online discussion and expect that learning occur naturally without much of facilitation or consideration of the learning task, outcome, and environment" (p. 476). While going in-line with what Zhu (2006) suggested in terms of the organization and planning of online asynchronous discussions, Black (2005) stated a more positive reflection on the use of AODs by focusing on the fact that AODs make the voices of students equally heard as much as the instructors since they need to contribute with their opinions about the subject.

Pop (2011) investigated the reflections of university students who had out-of-school responsibilities which were hindering their chance of attending face to face classes regularly on voice recorded online asynchronous discussions and the results revealed positive feedback from students in their reflections. In the more recent literature, O'shea, Stone and Delahunty (2015) analysed students' reflections on fully online asynchronous university courses and the results showed that adult learners had a pragmatic approach to learning, and online classes have been suggested as beneficial terms of cost, time and distance.

Alzahrani (2017) investigated the effect of online discussion forums on undergraduate students' learning and achievement. The results showed a statistically significant relationship between student participation in ODFs and their final grades. The author suggested that thanks to the collaborative nature of ODFs and its allowing students to actively participate in their own learning, using ODFs to improve traditional learning could help students' achievement as long as they collaborate and participate in ODFs effectively.

Recently, two systematic reviews about the use of OADs were conducted. The first one was by Foo and Quek (2019) which was a meta-analysis on improving critical thinking skills via AODs of forty-seven articles. The authors presented the advantages of using AODs by highlighting the ease of reflection and revision of the ideas in them. The second review from Fehrman & Watson (2021) focused on the use of AODs in higher education, particularly AOD learning outcomes and alternatives. Subthemes such as interaction and participation, AOD format and group size were categorized for effective AOD environments. In discussion board "alternatives" part, interesting results from various studies were presented. For example, in Cho (2016) it was found that while there was no correlation between participating in AODs and higher learning outcomes, it was suggested that teachers' primary role should be to design a quality course that would enable students to learn concepts effectively without relying on discussion boards to do everything. They also indicated the importance of quality feedback in AODs.

Another study that highlighted the importance of feedback and assessment came from Klisc, Gill and Hobbs (2017). They conducted a study to see the impact of different post-AOD assessments on students' critical thinking skills, and their results highlighted the importance of having an assessment, regardless of the type of the assessment made.

Going in-line with these suggestions, Gao, Zhang and Franklin (2013) stated that the idea of carrying the classroom setting to an online platform would not be sufficient and conducted a meta-analysis on asynchronous discussion environments and put them into four categories such as visualised, anchored, combined environments and knowledge forums, combining the previous three environments and analysed their effectiveness on students' learning. As a result of their categorization, they recommended that teachers, researchers and educators should focus on designing appropriate instructional activities for improving participants' performances in online environments.

Fully online and asynchronous classes have also theoretical support especially from the socio-constructivist approach (Dantas-Whitney, 2002; Järvelä and Häkkinen, 2002; Black, 2005; Alzahrani, 2017). Dantas-Whitney (2002) referred to the socio-constructivist theory by Vygotsky as stating, "learners are creators of their own knowledge, and they learn as they create new meanings according to previous experiences (p. 544)." As far as constructivist classrooms are concerned, learners are responsible for their own learning, and online classrooms in which asynchronous discussions are taking place provide fruitful environments for social learning.

3. Methodology

This qualitative study was designed as an autoethnographic research. While demonstrating autoethnography, Creswell (2013) refers to Muncey's (2010) definition which is "the idea of multiple layers of consciousness, the vulnerable self, the coherent self, critiquing the self in social contexts, the subversion of dominant discourses, and the evocative potential. They contain the personal story of the author as well as the larger cultural meaning for the individual's story (p.73)". The teacher researcher presented her educator-self as well as her researcher self while her students reflected on their learner-selves and prospective flight attendant-selves, bringing a perspective to describe their experiences with a fully online course, particularly online asynchronous discussions. While planning to conduct the study, the teacher-researcher discussed turning this brand-new experience for all of them to an academic article with the students taking the course, so that this practice could be used as a reference for the interested groups of teachers and researchers. As the initial step of conducting the research, the teacher-researcher took students' verbal consent to participate, and in every step of the research, students' and teachers' confidentiality was taken cautiously.

3.1 Setting and Participants

The data were collected throughout 15 weeks from students taking Advanced speaking skills II course in a civil aviation cabin services department of a private vocational school in Türkiye throughout the spring term of the 2022-2023 academic year. Thanks to the agreement between the university and the flag-carrier airline company of the country, 22 students who had minimum B1 language proficiency were chosen to take part in an internship project. These students were placed in one group which was nicknamed as the "TK" class. According to the agreement, these students were required to complete their courses within two weekdays and fly as intern flight attendants in domestic flights in the other days of the week. As a consequence of this agreement, a tight schedule had to be created particularly for the TK group and the Advanced Speaking Skills II course was decided to be given as a fully online course by the head of the department and the English coordinator particularly for this class. Following the agreement, it was confirmed by the head of the educational committee that fully online

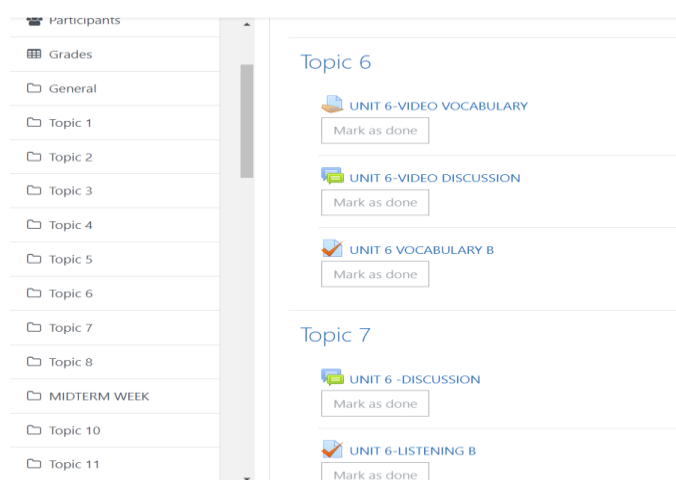
conduction of a class was allowed in the post-pandemic and post-earthquake period when online education was not compulsory.

The participants were 22 students from the TK group who took the advanced speaking skills II class in the spring term and were doing their internship as flight attendants and their English language instructor who is also one of the authors of this research. The data for online asynchronous discussions were collected from 22 students in total who took the course. In the 12th week for the open-ended questionnaire, 10 students volunteered to participate. As for the focus group interview, 6 students who posted entries to AODs regularly and posted a reply to the open-ended questionnaire were invited. The data from the teacher was collected from the observations and personal notes which the teacher-researcher who was the teacher of the advanced speaking skills II course for the TK group in Spring Semester took throughout the spring term.

3.2 Data Collection Tools

The first group of data was collected from asynchronous online discussion entries posted in Moodle LMS system every week to see the improvement of students (see figure 1.), and to see whether there was any change in the quality of the entries as well as students' interactions with each other.

Figure 1: The Moodle LMS Classroom



To see a sequential difference, the discussion entries were chosen from weeks 2, 6 and 10 within four-week intervals. Discussion topics were chosen both from the listening speaking skills-based course book and from aviation related topics (see figure 2.). The interview data was collected from an open-ended questionnaire sent to students in 11th week, and the focus group interviews made in the 13th week of the term. The data from the instructor was collected from the notes and observations of the teacher-researcher.

Figure 2: Asynchronous Online Discussion Topic of the Second Week

UNIT 5 FORUM 1

Mark as done

Dear students;

Here is your first forum assignment. You should choose 2 topics from the list below and comment on it in detail. Also remember to make comments for three friends' comments.

Best of luck!

1. What does "risk" mean to you?
2. Are risks always physical? Can you name any other risk types?
3. Why do people take risks? What are some possible rewards?
4. Have you ever taken any risks before? Were there any rewards? How did it end?

Add a new discussion topic

3.3 Data Analysis

The analysis of data collected from students' online asynchronous discussion entries in LMS was done as online document analysis. As Bowen (2009) states, "document analysis is a systematic procedure for reviewing or evaluating documents—both printed and electronic (computer-based and Internet-transmitted) material" (p. 291). In this particular research, asynchronous online discussion entry documents were analysed to triangulate the data collected from open ended questionnaire data from ten students and focus group interview data from six students to see whether the changes in the quality of their entries correspond their opinions about the way the course was implemented.

Following the collection and analysis of discussion entries, all 22 students who took the Advanced Speaking Skills II course were invited to fill in an open-ended questionnaire with 7 questions developed by the teacher researcher to see students' reflections of this fully online-asynchronous course. 10 students volunteered to fill in the questionnaire. The questionnaire data was thematically analysed and colour-coded according to the derived themes. Finally in the 13th week of the semester, an online focus group interview was done via Zoom with 6 students. These students were chosen purposefully as they were the ones who posted regularly on each post. In order to provide a comfortable environment and allow students to share their reflections easily, the focus group interview was done in Turkish. The Turkish data was first transcribed via the dictation feature of Microsoft Word, and then translated to English by the teacher researcher. The transcribed data were thematically analysed, and colour coded. Following the completion of each coding processes, all data results were combined and presented in themes in the results section of this paper.

4. Results

The results of this study are presented within four categories which are based on the data collection methods and number of participants who attended that particular method. First, the online document analysis of the data derived from LMS online asynchronous discussion entries is presented. Then results of the thematic analysis of open-ended questionnaire and focus group

interview analyses are conferred. Lastly, teacher-researcher's notes and observations regarding teaching an asynchronous online EFL classroom are displayed.

4.1 Document Analysis of AODs in LMS

The analysis of AOD entries from Moodle LMS system were collected from weeks 2-6-10. Each week's discussion topic, number of questions asked, number of students who posted an entry to each week, the number of entries and replies to those entries are presented in table 1 below. The number of participants to the discussions was 8 in week 2 whereas there were 6 participants out of 22 students who enrolled to this class in week 6 and 5 students posted a video entry with no replies to them.

Table 1: Initial Document Analysis of LMS AOD Entries

	Week 2	Week 6	Week 10
Discussion Topic	Risk & Reward	Understanding Bias	Ethics of Museums
Number of Questions	4	3	2
Number of Entries	15	8	5
Number of Replies	32 to 14 entries	8 to 5 entries	0
Number of participants	8	6	5

The word and character count of entries were also calculated. This calculation was done to see whether students tended to use real, complex vocabulary or simpler, shorter words such as "OK, yes, true" etc. Table 2 presents the total word count for each week's discussions. The table illustrates the dramatic drop of word and character counts as the weeks passed.

Table 2: Word Count and Character Count of AOD Entries

Week	Word Count	Character Count
Week 2	3773	16161
Week 6	1470	6493

Both written and video entries were also analysed thematically according to students' preferences of explanation of the given topic. For the risk and reward discussion in week 2, 4 themes emerged within students' replies were analysed. The results of thematic analysis for week 2 discussion are presented with the frequency of mention in table 3 below.

Table 3: Themes Derived from Discussions in Week 2

Themes Derived	Mention Frequency	Sample Statements
Presenting a definition	10	"Risk is a process which occurs while making decisions and means bearing in mind that something may happen negatively. When we take a risk, we already accept negative results, and we are already ready to compensate them."
Giving a personal example	5	"As my experience, sometimes taking risk can be dangerous and I wish I had pulled the amount of money. But this experience taught me that I shouldn't be greedy when i get rewarded."

speed” and S5 suggested there should be more of these activities. The excerpt from S2 is also as follows.

“I think, this course helped me for every of them cause the exercise to make us share our ideas with video or there were exercises which you should type, listen, read and then answer. (S2)”

S6 and S7 strongly advocated the inclusion of speaking and writing skills to this course with the adverb “definitely”. S1 and S2 justified their perceptions by stating that it helped them to learn new chunks, new terms and new vocabulary. S4, S6 and S10 mentioned speaking and writing while talking about activities that they enjoyed the most. S4 said *“The one's when we watch a video and answer questions about it. Because it makes me practice my listening, writing and speaking.”*

4.2.2 Online Discussions

Speaking of the activities which the students enjoyed the most, the second most frequently stated reply was about the quality, frequency and effectiveness of the online discussions. In 8 separate entries, the effectiveness of online discussions was mentioned by students. When students were asked how they benefited from online discussions, S2 stated that “it was nice for us to comment on each other and watch their videos. We have become our own teachers.” In another excerpt, S3 indicated as follows.

“I do think I could benefit from them cause at the same time we share our own ideas about the discussions, we could see other attendant's sharing as well and that helps us think more and considering other ideas as well.” (S3).

S5, S7 and S8 also stated online discussions as beneficial since they believe online discussions helped them to learn from each other. In an example about discussions, S1 said *“I enjoyed with discussions because it was interactive and when I write something I was waiting my classmates' feedback having their ideas about what I write makes me happy and excited about lesson”*.

4.2.3 Task Durations

Finally in open-ended questionnaires, the length of video assignments was mentioned as a trouble 5 times throughout the data. S1 said *that “sometimes the duration asked is longer than the topic I'll talk about and making the presentation gets really tiring”*. In S2's entry, it was said *that “I think by making video durations shorter we can have the maximum benefit, I think 1 to 3 minutes is perfect”*.

4.3 Focus Group Interview Data

In week 13, the teacher-researcher invited 6 students to an online focus group meeting for further discussion of the effectiveness of this online. While the most frequently discussed themes during the interview were the online discussions, and the technical issues regarding evaluation procedures, students mentioned the improvement of their writing and speaking skills and stated their opinions about the delivery method of the course, which was online. Since

students who filled in the open-ended questionnaire and joined focus group interviews were different, students who joined focus group interview were coded and presented as “SF”.

4.3.1 Online Discussions

When students were asked their overall reflections about the course, they started to talk about online discussions. SF2 found discussions efficient especially while practising how to agree or disagree when commenting on someone’s entry in an academic way. SF6 enjoyed them because they could see others’ opposing views and could discuss why they think differently on the same issue. The joy of seeing different perceptions was also mentioned by SF6 and SF9. For example, SF 6 stated that *“especially when SF7 is making comments, we are pulling each other out. For example, we like each other very much. I say this love is getting more frequent with all discussions”*.

Similarly, SF3 exemplified his enjoyment as follows; *“a message comes from above [he means mobile app notifications], I wonder why they said, “I didn’t agree with you”. Why not? What was the thing that they opposed? And I used to check that discussion. it was so nice, so were that feedback.”*

Apart from the discussions, SF7 stated the importance of video discussions to improve their vocabulary. While talking about which activities they would increase, SF7 expressed herself as follows.

“I would increase the video discussions even though they had technical problems. When I used a word at least while speaking, it became more permanent and exposed ourselves in video discussions. These were more effective when we took notes and so on for the conversation. When I use that word in a sentence for speech, it is much more effective for me. That’s why I would increase the video discussions more. Because we have to expose ourselves, we have to talk. So we have to express our opinion on that subject.” (SF7)

4.3.2 Technical Problems

Another popular topic was about the technical problems. It was mentioned 8 times in total by all 6 focus group interview participants. Students mostly complained about the length of the videos and how hard it was to upload video assignments to the LMS platform because of the maximum file size limitation. *“I had to divide my video into 3 pieces”* said SF2 while talking about the problems they faced. SF9 agreed and stated, *“I made the video smaller 8 or 9 times by continuously cutting it, then the video had dropped to 50 megabytes.”* SF3 also complained about the video editing process as well as the duration of the videos. *“4-5 minutes were obviously a bit too much for me”* SF3 said, *“if it was 3-4 minutes, I could have shot more effective videos”*.

The lastly mentioned technical problem was about the third party of online assessment platform which was provided and controlled by the publishing house of the course material. From their website, students had to do additional vocabulary, reading, listening activities and they found these activities tiring and unnecessary (SF9). SF6 complained about that app’s not being user

friendly since it did not allow students do their assignments from their mobile phones as effectively as from their computers.

4.4 Teacher Researcher's Notes

When I was first assigned as the instructor of this particular course, I thought nothing would be different. I was going to go to my class and do speaking and presentation practice with my second-year students whose language proficiency was luckily enough to prepare and give presentations in English. When the administrators told me that this group of students were going to fly with the flag carrier airline of the country as interns for the full spring semester and then during the summer period, I was both proud and concerned at the same time due to the weekly schedule the department had already created. Since their exact flight dates were not clear back then and we were supposed to stash all lessons to Mondays and Tuesdays, there was no room for my Advanced Speaking Skills II class in the schedule. With an instant decision, we decided to give the lesson as a fully online asynchronous course, because we had already tried online education before, and it had gone “not too bad” and trying new things had always been interesting for us as we had always tried to be the most innovative department of our university and our vocational school.

At first, I was happy and relieved that I was going to have fewer physical lessons; I didn't have to wake up early or leave the school late or be occupied for two hours instead of doing my formal paperwork. I was just going to announce the subject and that was all! Even maybe we could have thought about converting all groups to this method of delivery! My level of excitement was visibly high, and I was sure that nothing could have gone wrong.

The first week was easy. I assigned all activities until midterm week in publishing house's platform, updated the syllabus and we hadn't started to use the book yet. The students were as excited as I was, and they all seemed to be determined to contribute to the class despite its fully online / asynchronous nature. Then the second week, I had to create an assignment in our Moodle LMS for the class, I had to make sure that without meeting with my students at all, they had understood all the subject matter and became familiar with the vocabulary provided in the unit. At that moment I realized how thick the units were and how face-to-face oriented the material was. It was that moment that I wished I had arranged those two hours in the weekly schedule, but our students were interns and flying in domestic flights. With this motivation, I analysed the material, decided on which activities would be helpful and started thinking about alternatives to make my students' voice heard by each other both physically and ideologically. It was then that I created the first forum discussion and started reviewing the existing literature about asynchronous online discussions for various disciplines, with a focus on EFL classrooms.

For the following five weeks, there was no problem, students were replying to discussions, sending videos or leaving comments on their friends' content until the assessment week. During the midterm week, since students had no chance to be at school for our lesson, they had to submit their presentations on a given topic and I had 22 videos to watch and evaluate and give feedback on. It was 44 videos for the final exam, and since the attendance was being calculated by students' posts and comments on each other's entries, I had to check the platform regularly and by the next course time came I had to enter the attendances and come up with new discussion topics for every other week. I should admit that even though it was hard and seriously time consuming at the beginning, by the 7th week, I got used to applying the procedure and

actually started to enjoy reading students' entries and watching their videos on topics that we never discussed face-to-face. For the last weeks of the semester, as I collected data from students, I started to think more on different subjects and more challenging formats of delivery for assignments, and despite the initial drop of motivation and panicked mindset at the earlier stages of the course, I was able to pull up my motivation and increase my excitement and interest on this fully online asynchronous course.

In my opinion, I managed to carry on this course well, but I must admit that short meetings on every two or three weeks would enhance the effectiveness of the course and students would feel more "belong" or "tied" to the course as well as to the school and they would benefit from the units and discussions more. This might have even made my workload less since every week I also had to create announcements regarding the lesson and give regular feedback, which are really time-consuming activities.

Overall, I think this was an experience of once-in-a-lifetime for all 21st century instructors who are facing a massive shift to online settings in many aspects of the educational practices. Even though I hadn't done it before, sooner or later I was going to have this experience. Having completed a term successfully with various positive feedback from my students, I can say that I could manage to give and create a fully online course with asynchronous online discussions embedded in various learning platforms and turn this experience into an academic paper in the form of an autoethnography. Still, I would recommend fellow teachers and researchers not to switch their course delivery method this suddenly since I think by doing so, the workload is multiplied. When I think about it, if I had known about this possibility of delivery of this course, I would have had more time and creativity which could have eventually made the class better and made me feel more confident about what I had done.

5. Discussion, Conclusion and Future Implications

This autoethnographic study was conducted with one class of intern flight attendants who were taking Advanced Speaking Skills II course in a fully online asynchronous setting during the spring semester of 2022-2023 academic year and with the comments of the teacher-researcher. All activities were provided in online platforms such as Moodle LMS and material Publisher's platform and students were asked to comment on asynchronous online discussions (AODs) either by posting a written entry or recording a reply video about the given topics. The aim of this study was to see students' reflections on the effectiveness of fully online asynchronous discussions as well as the platforms used and how students thought this course helped them to improve their communication skills.

As far as the document analysis was concerned, it was observed that the number of words and characters as well as the number of entries dropped as the weeks passed. It can be indicated that students' motivation to post dropped gradually. This might be because there was no meeting or course related communication between the instructor and students other than the notifications and announcements made in online platforms such as Moodle LMS or school's formal administrative website. This conclusion is in line with Klisc et al.'s (2017) suggestion of undergraduate students' need for guidance and instruction to achieve learning outcomes (Alexander et al., 2010; Bai, 2009; Gilbert and Dabbagh, 2005) Another possibility might be the repetitive nature of the activities and discussions. To improve the quality of arguments in AODs, Klisc (2017) suggests explicit instruction in logical reasoning and deduction, testing

premises and questioning other participants. Accordingly, for courses similar to the ones in this research, discussion interventions can be embedded. Yang (2008) also suggested intellectual online discussions' to be designed in a way that students could enjoy high quality discussions and express their opinions with the guidance of an instructor.

As far as the results of open-ended questionnaire and focus group interview data are concerned, while students were happy about online discussions, the video submission phase was challenging and the length of video assignments as well as their editing process was exhausting and demotivating. Students mostly enjoyed seeing opposing views on a given topic and they thought the discussions and other activities helped them to improve their speaking skills as well as their writing skills. This result supports the findings of various previous studies about the effectiveness of synchronous online discussions (Birch & Volkov, 2007; Black, 2005).

As for the reflections of the teacher-researcher, it can be said that it is in line with the existing literature (Alzahrani, 2017; Yang and Spear, 2017; Fehrman and Watson, 2021). The overwhelming amount of workload and not having regular communication with students was a great challenge. However, as it is acquainted, the time management can be achieved, and more fun and challenging tasks can be assigned to students. Gao et al. (2013) emphasized that to enhance the use of online discussion environments for instructional objectives, there should be a focus on designing appropriate instructional activities and developing appropriate pedagogical implications which could improve participants' performance in such environments.

6. Limitations and Suggestions for Further Research

This paper is not free of limitations. First, the number of students who contributed the paper and participated regularly in class discussions and posted entries constantly dropped. A study with classes that seek more strict and regular attendance requirements could be conducted in the future. Also, only for one class, one instructor's and only 22 students' ideas were taken into consideration. For future studies, a mixed method exploratory study could be conducted with more students and more instructors within a longer period of data collection. More than that, a study with different levels of language proficiency can be conducted to see whether students' proficiency affects their participation in online discussions or not. Last but not the least, this paper did not investigate students' grammatical improvement as a result of asynchronous online discussions in fully online settings as the nature of the class did not focus on improving grammatical skills but improving students' speaking and listening skills as well as their fluency. Future research about the effectiveness of AODs in fully online language courses can be explored within contexts such as grammar and vocabulary teaching as well as the teaching of receptive skills such as reading in online or hybrid EFL classrooms.

As the education world is shifting to fully online/hybrid settings, reflections of students experiencing being educated in such classrooms and teachers' experiences on how to develop, maintain, assess and evaluate such courses has gained extra importance. This autoethnographic study was conducted to see the effectiveness of asynchronous online discussions in a fully online language course given to intern flight attendants with the perspectives of both students and the teacher-researcher of this study. The results can pave the way for administrators, curriculum heads or language education program coordinators as well as all EFL/ESL departments to conduct more research while considering developing new courses.

This article can also help in-service instructors who are interested in creating continuous professional development (CPD) activities and who would like to support research in this area. The results of this study are also valid for all areas of business English instruction or English for specific purposes (ESP) classes as it was conducted with flying intern flight attendants. Further research could take place to see the experiences of teachers and students whether online discussions or fully online language learning settings could be implemented for different areas of ESP classes.

References

- Alexander, M. E., Commander, N., Greenberg, D., & Ward, T. (2010). Using the four-questions technique to enhance critical thinking in online discussions. *MERLOT Journal of Online Learning and Teaching*, 6(2), 409-415. Available: http://jolt.merlot.org/vol6no2/alexander_0610.htm
- Alzahrani, M. G. (2017). The Effect of Using Online Discussion Forums on Students' Learning. *Turkish Online Journal of Educational Technology-TOJET*, 16(1), 164-176.
- Bai, H. (2009). Facilitating students' critical thinking in online discussion: An instructor's experience. *Journal of Interactive Online Learning*, 8(2), 156-164. <http://www.ncolr.org/issues/jiol/v8/n2>
- Birch, D., & Volkov, M. (2007). Assessment of online reflections: Engaging English second language (ESL) students. *Australasian Journal of Educational Technology*, 23(3).
- Black, A. (2005). The use of asynchronous discussion: Creating a text of talk. *Contemporary issues in Technology and teacher education*, 5(1), 5-24
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*, 9(2), 27-40.
- Cho, M. H., & Tobias, S. (2016). Should instructors require discussion in online courses? Effects of online discussion on community of inquiry, learner time, satisfaction, and achievement. *International Review of Research in Open and Distributed Learning*, 17(2), 123-140.
- Creswell, J. (2013). *Qualitative inquiry and research design: Five different approaches*.
- Dantas-Whitney, M. (2002). Critical reflection in the second language classroom through audiotaped journals. *System*, 30(4), 543-555.
- Denzin, N. K. (1970). Problems in analyzing elements of mass culture: Notes on the popular song and other artistic productions. *American Journal of Sociology*, 75(6), 1035-1038.
- Fehrman, S., & Watson, S. L. (2021). A systematic review of asynchronous online discussions in online higher education. *American Journal of Distance Education*, 35(3), 200-213.
- Foo, S. Y., & Quek, C. L. (2019). Developing Students' Critical Thinking through Asynchronous Online Discussions: A Literature Review. *Malaysian Online Journal of Educational Technology*, 7(2), 37-58.
- Gao, F., Zhang, T., & Franklin, T. (2013). Designing asynchronous online discussion environments: Recent progress and possible future directions. *British Journal of Educational Technology*, 44(3), 469-483.

- Gilbert, P. K., & Dabbagh, N. (2005). How to structure online discussions for meaningful discourse: A case study. *British Journal of Educational Technology*, 36(1), 5-18. <https://doi.org/10.1111/j.1467-8535.2005.00434.x>
- Järvelä, S., & Häkkinen, P. (2002). Web-based cases in teaching and learning—the quality of discussions and a stage of perspective taking in asynchronous communication. *Interactive learning environments*, 10(1), 1-22
- Klisc, C., McGill, T., & Hobbs, V. (2017). Use of a post-asynchronous online discussion assessment to enhance student critical thinking. *Australasian Journal of Educational Technology*, 33(5).
- Muncey, T. (2010). *Creating autoethnographies*. Sage Publications.
- O'Shea, S., Stone, C., & Delahunty, J. (2015). “I ‘feel’ like I am at university even though I am online.” Exploring how students narrate their engagement with higher education institutions in an online learning environment. *Distance Education*, 36(1), 41-58.
- Pop, A., Tomuletiu, E. A., & David, D. (2011). EFL speaking communication with asynchronous voice tools for adult students. *Procedia-Social and Behavioral Sciences*, 15, 1199-1203.
- Yang, D., & Spear, C. E. (2017). The effects of faculty status, faculty gender, field of study, and class size on the use of blogs, wikis, and discussion boards. *International Journal of Information and Communication Technology Education (IJICTE)*, 13(2), 52-64.
- Yang, Y. T. C. (2008). A catalyst for teaching critical thinking in a large university class in Taiwan: Asynchronous online discussions with the facilitation of teaching assistants. *Educational Technology Research and Development*, 56, 241-264.
- Zhu, E. (2006). Interaction and cognitive engagement: An analysis of four asynchronous online discussions. *Instructional Science*, 34, 451-480.

Blended Learning – Class Time Saver

Natia Zviadadze^{1*}, Sopio Kipiani², Tea Liklikadze³

¹Assoc. Prof., Akaki Tsereteli State University, Georgia

²Assist. Prof., Akaki Tsereteli State University, Georgia

³Invited Specialist, Akaki Tsereteli State University, Georgia

Abstract

The paper presents the case study conducted at Akaki Tsereteli State University, Georgia. It aims at revealing the benefits of ‘Blended Learning’ method in regard with using class time more effectively on communicative tasks in a collaborative environment. The research highlights the results of the comparative analysis of ‘Blended Learning’ vs ‘Traditional Learning’ within the framework of English as a foreign language courses taken by ATSU students. The study draws insights into the instructional benefits of hybrid learning focusing on reinforcing the communicative approach, promoting live communication, and increasing the frequency and quality of productive activities through saving classroom time at the expense of online independent assignments, accessible instant feedback, and statistical data provided by the platforms. ‘Blended Learning’ as an instructional method has been in practice at ATSU EFL courses for about 8 years in parallel with ‘traditional’, ‘non-hybrid’ language teaching instruction. Over 1000 students take EFL BL courses annually using up-to-date textbooks with access to the corresponding electronic platforms. The study reflects the experience gained throughout the years and alongside with positive feedback and promising prospects, addresses the challenges suggesting ways of dealing with them to guarantee better results. To ensure the validity of findings the study implies different research instruments such as students’ and teachers’ survey analysis, class observations, statistical analysis of students’ independent work, and comparative analysis of course outcomes. Based on the research, it is concluded that the integration of human competence with technological facilities proves to be efficient and motivating for the young generation in today’s technology-driven world, and at the same time, maintains the irreplaceable role of human interaction in educational settings.

Keywords: Blended learning, communicative approach, class time

1. Introduction

The integration of technology in education has already become a common aspect of every university course. This growing trend is characteristic of English as a second or foreign language teaching courses. To some extent, this field is one of the first to undergo changes in this direction. There are abundant methods and techniques piloted and used in ESL methodology. Moreover, courses are conducted in different modes, such as traditional face-to-face settings, distance learning, or blended (hybrid) learning. These modes allow students and teachers to observe and compare results and express their preferences.

Recent challenges in the modern world have also contributed to enhancing technology in learning. The most significant cause of this increasing tendency was the outbreak of COVID-19, which significantly accelerated this ongoing process and left the educational field with no choice but to use the distance mode for schools, colleges, or any other types of educational institutions to function. This experience has resulted in a huge leap forward in adjusting to new requirements. However, it has also revealed the challenges of this mode of teaching, further paving the way for blended learning. In this context, blended learning has gained more significance by uniting both modes and offering a combination of their best aspects.

From today's point of view, it is more reasonable to support and encourage human teachers' roles in a traditional face-to-face classroom rather than advocate for online learning, which has seen a growing tendency, particularly after the COVID-19 pandemic extended the experience to the whole world. Regardless of which component we support more, it is almost universally agreed that we need a mix of both. As it is mostly agreed, "Technology will never replace teachers, but teachers using technology in their instruction will replace those who do not." (Krajka, 2012). This motto emphasizes that EFL teachers have no choice but to develop their Computer-Assisted Language Learning expertise.

2. Literature Review

To achieve the purposes of this study, it is useful to examine the concepts and definitions of "Blended Learning" (BL) found in the scientific literature. A brief review of previous research provides a solid theoretical framework for further exploration.

There is a variety of definitions found in the corresponding scientific literature. The term "blended learning" has been widely used by researchers over the past decade, often interchangeably with "hybrid learning." Graham (2006) defines BL as a combination of face-to-face (FTF) and computer-mediated instruction. Other educationists have their definitions: Hartman et al. (2007) describe it as a mix of face-to-face classroom instruction and online learning, reducing classroom contact hours. Sands (2002) highlights the reduction of seat time and distribution of activities like testing and essay writing throughout the semester. Chan (2008) suggests BL combines classroom training, live and self-paced e-learning, and advanced supportive learning services in a tailored manner.

Scholars and practitioners stress the advantages of blended learning, claiming that it allows students to control their learning path and pace through online technologies while being supervised in an educational setting via face-to-face instruction. It merges the classroom's communicative opportunities with the online environment's technologically enhanced active

learning possibilities (Dziuban, Hartman, & Moskal, 2004). Whitelock and Jefts (2003) describe BL as combining traditional learning with web-based approaches. Badawi (2009) views it as a flexible approach that mixes face-to-face activities with online practices, enabling collective and individual feedback and responses in various areas, either synchronously or asynchronously.

In the context of foreign language learning, the use of computers is not new. Studies in Computer Assisted Language Learning (CALL) date back to the 1960s, showing the evolution of technology from large text-based mainframe computers to personal computers and portable gadgets with internet connections (Hubbard, 2009). CALL has significantly evolved over the past 50 years and is now integral to second and foreign language pedagogy. Initially seen as a classroom supplement, CALL now promotes learner autonomy and engagement with the target language both inside and outside the classroom (Fotos and Browne, 2011a). Levy (1997) defines CALL as “the search for and study of applications of the computer in language teaching and learning” (Olejarczuk, 2021).

Blended learning has proven highly effective in English as a Foreign Language (EFL) contexts. It enhances student engagement and motivation. For example, a study in Pakistan found that the blended mode significantly improved students' English proficiency, with factors such as ease of use, computer literacy, and face-to-face interaction positively influencing learning outcomes (Majeed & Dar, 2022). Similarly, a study at Marinduque State College in the Philippines highlighted that Flexible Blended Learning (FBL) arrangements, which include asynchronous and synchronous modes, allowed students to learn at their own pace while benefiting from real-time interactions (Adling, 2022). Living in an age where technology is an integral part of students' lives (Prensky, 2001), the way information is presented by the University is critical to engagement. The positive experience is shared in a report on the first-year transition at the University of Auckland, where the face-to-face Orientation Program was overhauled and a new Online Orientation course went live for undergraduates. ‘Students expect institutions to use technology to engage them in the learning process, and that blended learning environments are now the norm ... Online learning has now become the dominant method for adult learners. It also improves access for students with disabilities and assists ESL students who have the additional cognitive load of interpreting what is being said before being able to comprehend and develop understanding. (Wilson and Minhas-Taneja, 2015).

The use of blended learning aligns with the principles of social cognitive theory, which emphasizes learning through social interactions and modeled behaviors (Bandura, 1986; de Guerrero & Villamil, 1994). This theory underpins the effectiveness of incorporating cultural heritage into language learning, as students learn through exposure to and interaction with diverse cultural contexts. For instance, international students in a blended ESL course reported increased motivation and understanding of Sustainable Development Goals (SDGs) when cultural heritage was integrated into the curriculum (Kwee and Dos Santos, 2022).

University leaders and curriculum developers are encouraged to incorporate digital components into their programs. As higher education institutions adapt to post-pandemic realities, the integration of blended learning can provide comprehensive and engaging learning experiences.

Blended learning is a robust and adaptable method that effectively supports EFL education, making it a vital component of contemporary education strategies. Recent studies suggest that

blended learning has become a norm for modern educational systems, emphasizing its critical role in enhancing learning outcomes and student engagement.

3. The Study and Methodology

The case study presented here aligns with the concept of Blended Learning (BL), which combines online components with traditional face-to-face classroom activities in a thoughtfully planned and pedagogically beneficial manner. In this approach, a substantial portion of institutionally defined independent work is conducted online, and traditional workbooks are replaced by digital ones. This integration of BL strengthens face-to-face communication rather than diminishing the importance of live interaction. In fact, when applied correctly, this method significantly conserves valuable seat time, enabling a greater focus on productive classroom activities.

This paper presents a case study conducted at Akaki Tsereteli State University, evaluating the EFL blended course based on data from the 2023-2024 academic year, with a focus on its efficiency in terms of saving class time and reinforcing communication in class. Additionally, the data were compared to those obtained in a similar 2018 study (Zviadadze et al. 2018), providing a valuable resource for contrastive analysis and drawing conclusions, supported by the experience of teaching blended English courses both before and after the pandemic.

The goal of the study is to demonstrate the improved efficiency of class time use and overall achievements of teaching blended English courses since their first introduction at Akaki Tsereteli State University, suggesting that it should become the norm of instruction in modern higher educational institutions. The study aims to answer the following research questions:

- What aspects of BL EFL courses allow for more efficient use of class time?
- What are the challenges of increasing the efficiency of class time used for BL EFL courses?
- What are the benefits and limitations of BL EFL courses compared to traditional offline EFL courses at Akaki Tsereteli State University?
- What progress has been noticeable in teaching BL EFL courses at ATSU since their introduction?
- What are the present-day challenges of EFL courses conducted via the Blended Learning format in the studied context?

For a thorough and detailed study, a variety of qualitative and quantitative research methods are employed. The following research instruments were used to test the hypothesis and reach conclusions:

- Students' survey analysis
- Teachers' survey analysis
- Students' independent work comparative statistical analysis using 'Gaussian Distribution'
- Students' online independent work statistical analysis
- Class observation comparative analysis

The case study proved valuable for gaining information about the research questions and providing deep insights into improvements and new challenges. Furthermore, it served as a source for further studies and provided a valuable foundation for developing strategies to enhance English language teaching at Akaki Tsereteli State University, suggesting ways that may be useful for other institutions undergoing similar processes.

3.1 Students' Survey Analysis

The online survey was conducted for students utilizing 'Google Forms'. Approximately 100 students from Akaki Tsereteli State University, who are studying General English with access to an E-platform as a compulsory subject of their course, took part in the survey. These students are first and second-year students majoring in various fields and enrolled in EFL compulsory courses within accredited academic programs. Currently, the university offers five different coursebooks with access to the E-platform tailored to students' language CEFR levels.

The survey comprised 20 questions, including 16 objective and 4 open-ended questions. The first two questions served as an introduction, gathering information about the respondents' academic year at the university and the specific coursebook they were using. The questions addressed various aspects of Blended Learning, including access to E-platforms, the integration of online and offline activities, assessment methods, instant feedback mechanisms, and factors that prioritize saving class time over maintaining a balanced blend of learning modalities.

The questions referred to different aspects of Blended Learning, such as the benefits of the online component built into the course (questions 3-7), saving offline class time for more interactive activities at the expense of online tasks (questions 7-11), comparison between traditional offline and online formats (questions 12-13), developing individual learning skills (question 14), balance between online/offline components (question 15), access to authentic material and online assessment feedback (questions 16-17), and technical issues (question 18). Students also shared their ideas and recommendations through open-ended final questions (questions 19-20).

According to the survey feedback, a vast majority of students feel enthusiastic about the course, find the built-in online component beneficial for using class time more efficiently as well as for developing their language and individual learning skills, and strongly recommend a blended course.

Students prioritized the following aspects regarding the benefits of the blended course:

- Flexibility – 63.2%
- Enjoyability – 42.1%
- Class time saver – 25.4%
- Developing technological skills – 21.1%
- Motivation – 19.3%

The focus of the survey was to reveal the factors affecting the saving of class time for more interactive and communicative activities to promote students' productive skills. The survey results appeared to be exhaustive, covering different related issues:

- Allows more time for interaction in class – 54.4%

- Engagement in collaborative tasks – 43.9%
- Less class time spent on homework feedback – 26.3%
- Less class time spent on giving assignments – 21.9%
- Encourages students' better involvement in class activities – 21.9%

Interaction with a teacher and other students was rated highly in a blended EFL course compared to other traditional courses. Students also feel satisfied with the suggested balance between online and offline components. Most of the respondents think that a blended course offers better access to authentic materials and meets individual needs through instant and transparent feedback.

Students shared their recommendations for course improvement. A great majority of them find the course suitable for them. However, some think that the course could allow more time for reinforcing communicative activities and suggest more varied tasks to make the process even more enjoyable. In addition, students referred to some technical issues such as the need for technologically better-equipped classrooms and smoother platform navigation to get the most out of the course.

Overall, the student survey appeared to be informative enough to contribute to drawing generalized final conclusions for the study.

3.2 Teachers' Survey Analysis

The teacher survey is a valuable instrument for the study, as teachers consistently provide thorough feedback to the survey questions. According to the survey results, it can be inferred that the analysis of teacher surveys offers deep insights into the study, laying the groundwork for further generalizations and recommendations.

The survey comprises 22 questions: 18 objective questions and 4 open-ended questions covering essential aspects of teaching a blended course, ranging from technical and content-related issues to teachers' recommendations. The first three questions pertain to teachers' experience in teaching English as a Foreign Language (EFL) course, specifically their experience with teaching a blended course. The fifth question focuses on the text books teachers utilize for blended EFL courses (Questions 1, 2, 3, 5).

The survey reveals that respondents' teaching experience ranges from 1 to 30 years, with a significant majority (50%) having 11 to 20 years of experience. Despite their extensive teaching experience, most have been teaching a blended course for 1-10 years.

The fourth question asks teachers to state their preference between blended and traditional, fully offline language courses. Responses to this question were explanatory, with teachers providing arguments to support their choice. Sixteen teachers shared their comments, expressing a preference for blended courses. The benefits they highlighted included flexibility, efficiency, and convenience due to the built-in technical components, as well as ease in assigning, assessing, and providing feedback, leading to time savings in class. Resource utilization, diversity, opportunities to develop students' independent learning skills, and a less teacher-centered approach were also emphasized. The teachers' comments were persuasive and informative, focusing on various aspects related to the question.

Questions 6 and 7 pertain to saving class time at the expense of instant online feedback and other online activities. Question 8 provided more specific and detailed feedback on this issue. All 24 respondents shared their experiences, with multiple-choice answers taking precedence. Here's a summary of their opinions: The most significant time-savers in class appeared to be setting online tasks via electronic platforms and having instant feedback readily available electronically, both mentioned by exactly 62.5% of teachers. Additionally, the same percentage of respondents highlighted the benefit of individual access to online resources for saving class time, while the ability for students to work at their own pace during online assignments was mentioned by 58,3%. Thirty-three point three percent of respondents noted that less time was spent on class lead-ins due to students' readiness for communicative tasks.

Another factor, chosen by 25% of respondents, was the reduced time spent in class on audio and video resources, as these were assigned online in advance. Teachers provided additional comments on the same issue in Question 9. One of the responses encapsulated several aspects mentioned by the teachers: "Continuous online assessment helps monitor students' progress, enabling timely interventions and a more efficient use of class time." Overall, the survey feedback indicates that blended courses significantly save class time for more interactive in-class activities.

Questions 10 and 11 addressed other important aspects of e-assessment, such as meeting the individual needs of students, providing more accurate, transparent, and reliable feedback, and ensuring the prevention of technical errors. Furthermore, teachers responded positively to the advantages of blended courses, including increased interest and engagement, a fair balance between online and offline components, and better access to a variety of authentic materials. Teachers also found technical issues manageable and the platforms user-friendly (Questions 12-18).

Sixty-six point seven percent of teachers reported covering material more efficiently in a blended course, while 29.2% were undecided, leaving only 4.2% of teachers who believed the traditional format was more efficient for covering the syllabus (Question 19). Nearly every respondent agreed that teaching a blended course enhanced their professional development (Question 20).

Teachers further expanded their responses to objective questions in an open-ended question 21, which provided a more detailed summary of the benefits mentioned above. One response stated: "The diversity of blended learning methods helps cater to different learning styles and keeps students more engaged. It also makes teaching more dynamic compared to traditional methods." As one comment suggested, a blended course proved beneficial not only for students but also for teachers.

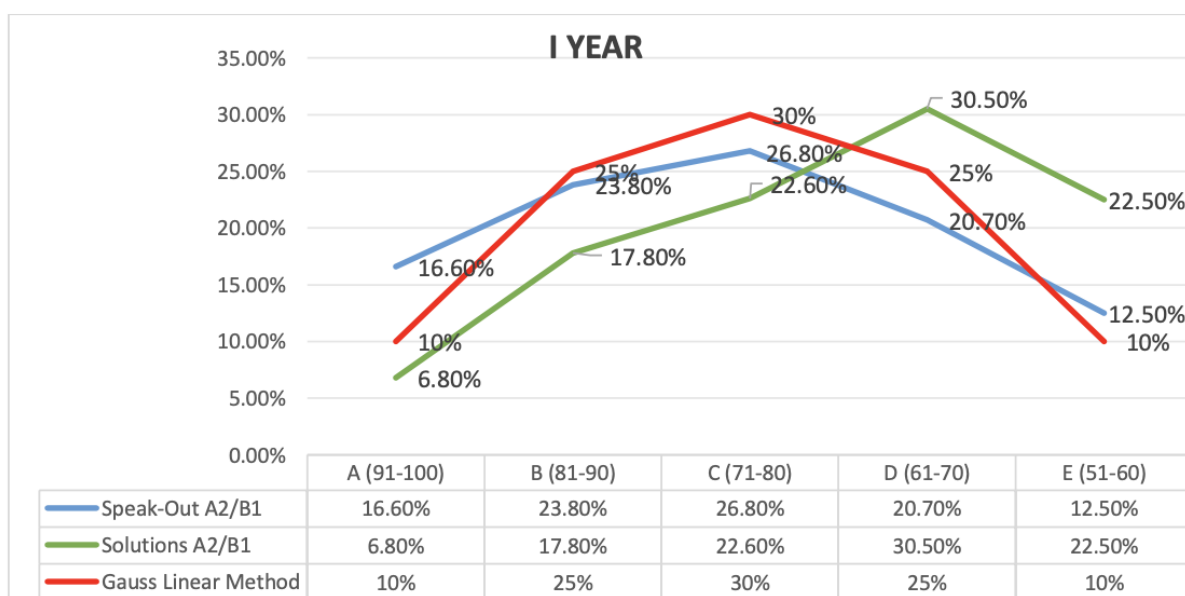
While generally content with the courses, teachers provided just a few recommendations for improving them. Emphasis was placed on the need for more technically equipped classrooms, as the lack thereof places an extra burden on teachers who must carry their own laptops or adjust planned activities accordingly. Another issue addressed was the need to reinforce interaction, with a focus on production tasks, which the format easily accommodates. Other suggestions included incorporating more interactive elements such as quizzes, simulations, and gamified learning experiences to further enhance engagement. Additionally, providing more theoretical feedback on grammar and vocabulary tasks was mentioned as a factor for improving the course.

Overall, the survey provided valuable feedback, despite primarily focusing on coursebook applications and presenting experiences with specific coursebook platforms, limiting the exploration of others. Nonetheless, it still offered meticulous information within the given context.

3.3 Students' Academic Achievement Analysis

In order to measure the effectiveness of the blended courses offered by the ATSU Language Centre, a comparative analysis of the students' academic achievement in blended and fully offline traditional courses was carried out. Students' academic achievement was measured based on the statistical data obtained according to their ECTS grading scale. All the first-year students' results in blended EFL courses were calculated and compared to the same data from students taking traditional offline EFL courses. Finally, the Gaussian distribution was applied to ensure the relevance of both blended and offline courses to the Gaussian distribution. See Chart 1 below:

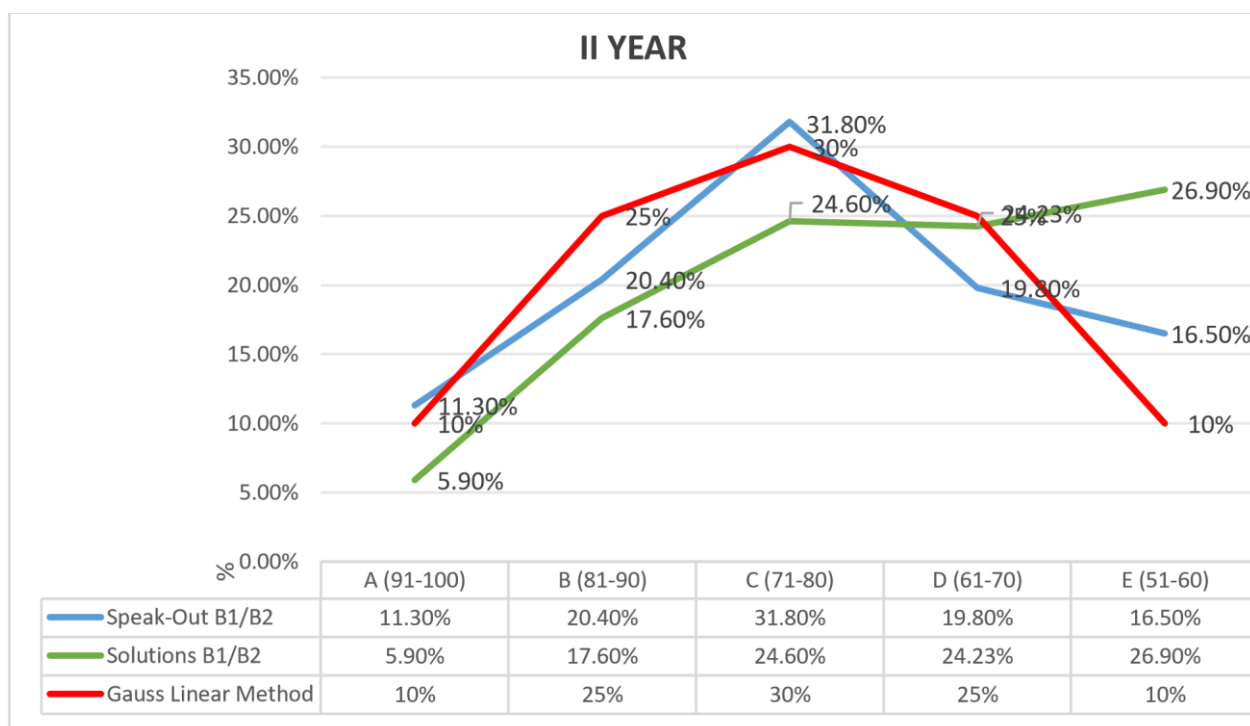
Figure 1: First Year Students' Academic Achievement (Autumn Term - 2023)



According to the chart, the blue line showing the results of the students registered in a blended course is in closer correspondence with the Gaussian distribution, reflecting the bell curve, than the green line indicating the academic performance of the students taking EFL traditional courses. Despite the fact that we do not think grading academic courses should strictly follow the Gaussian distribution—as grading depends on various factors in the educational context (such as course objectives, instructors' grading criteria, and the difficulty of the material)—it can still be used as one of the evaluative tools for course assessment.

Second-year students' data were gathered and compared using the same method. See Chart 2 below:

Figure 2: Second Year Students' Academic Achievement (Autumn Term - 2023)



As the chart shows, the picture is not very different from the previous one. The closer correspondence with the Gaussian distribution is still on the side of blended courses. However, more detailed observation shows that second-year students' A assessment data is lower and reaches the peak with a straighter upward tendency in the graph of C assessments. In addition, the percentage of E grades is slightly higher than that of first-year E graders.

The obtained information appears to be valuable for certain considerations related to the course's corresponding language level, course content, and assessment criteria.

3.4 Students' Online Independent Work Statistical Analysis

The fact that the program allows us to continuously monitor the progress of students' homework assignments provides reliable and effective data for further research. Students are given corresponding tasks for independent work regularly, with predefined time limits and a set number of attempts to answer the questions. The management of the assignment process can be controlled by a teacher and adapted to specific group requirements.

Statistical analysis of students' independent work enables us to observe how they develop their language proficiency and examine their language skills and systems separately. The data obtained from the statistics of the current academic year provides interesting information regarding students' academic performance and serves as a rich resource for process evaluation. The research was conducted separately according to levels and textbooks, but the diagram below combines the findings and illustrates the average percentage of grades by skills and courses. See the diagrams below:

Figure 3: Students' online assignment performance (2024)

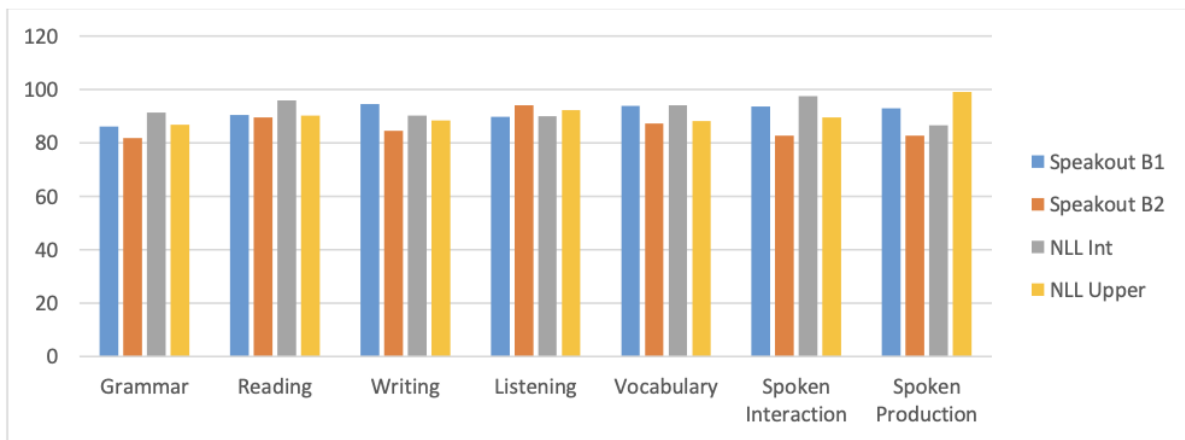
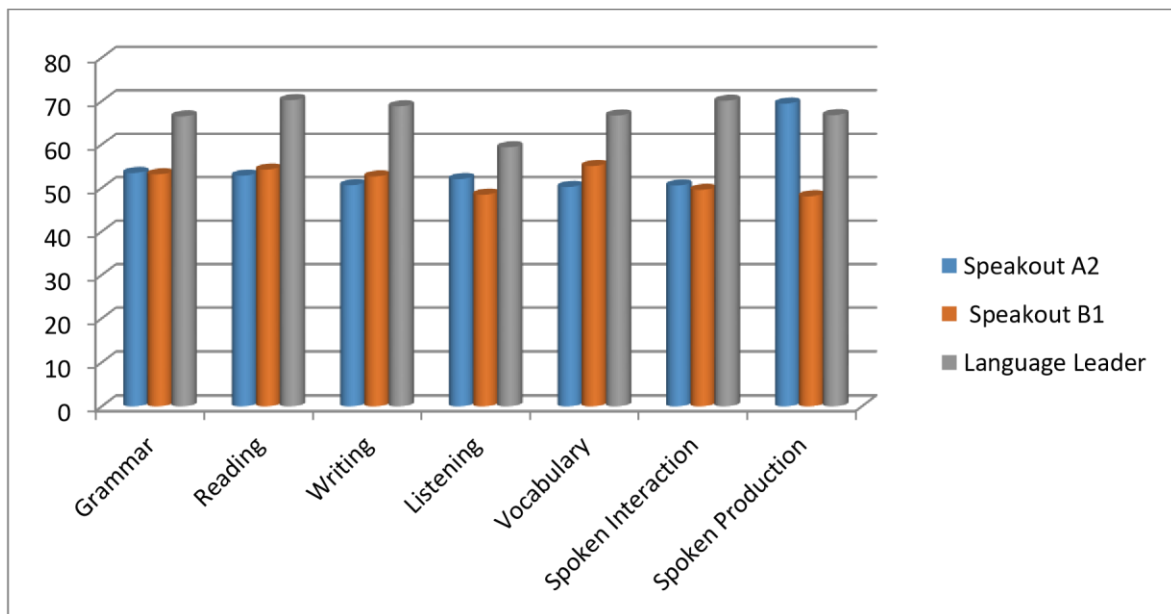


Figure 4: Students' online assignment performance (2018)



The first diagram indicates that there is minimal variation in grades across different textbooks, levels, and skills represented on the horizontal axis. All columns reach the 80 to 100 percentage range. Several conclusions can be drawn from the diagram, particularly when compared with a similar one created right after the introduction of a blended EFL course. (Zviadadze et al, 2018)

One significant trend revealed is the improvement in students' spoken interaction and production, showing a nearly 20% increase. The shift can be attributed, in part, to the online learning experience necessitated by the COVID-19 pandemic, fostering greater adaptation to online assignments. Another noteworthy observation, compared to the previous diagram, is that the performance of students majoring in English language and literature is comparable to those majoring in different fields. However, while these high achievements are notable, they also suggest the need for more challenging coursework by increasing online assignments and diversifying content.

In conclusion, the results are promising and indicate the potential for further enhancement of students' independent online assignments. Moreover, statistical data presented prove to be useful in planning changes to make the blended courses more student-oriented and to highlight the shift in focus toward communicative performance.

3.5 Class Observation Feedback Analysis

Class observations were conducted in the fall semester of this academic year. The primary aim of the observation was to compare EFL classes using a blended format with those using a traditional, offline format, focusing on the efficiency of class time use and the types of activities suggested to students during class time. Twenty classes were observed (10 of each mode), ranging from pre-intermediate to upper-intermediate levels across different BA programs. The observation questionnaires included preobservation questions, an observation timetable divided into the presentation/practice stage, production stage, and post-observation comments. The observation feedback form was adapted to the study's needs, which proved helpful considering the different focuses for different lesson stages.

Most of the findings from students' and teachers' surveys were confirmed at this stage, revealing similar benefits and challenges.

Interestingly, the general impression after reviewing observation feedback was positive in terms of achieving the desired results. The first noticeable difference was in dealing with homework feedback. Compared to traditional class observations, which revealed that too much class time was spent on homework feedback, the instant feedback provided by the platforms allowed teachers to have information about their students' progress or difficult areas before the next lesson. This enabled students to receive quick, pre-planned, and more effective feedback on their homework. The same applies to assigning homework; practically no class time is spent on that. Students receive assignments online after the class, saving valuable class time for more communication.

The focus has shifted from controlled, form-oriented practice to structured and free production, with more pair work and group work tasks, increasing students' talking time at the expense of reducing the teacher's talking time. However, over-reliance on pair-work, even when a different class arrangement is needed, has also been observed. Sometimes, teachers instruct pair-work activities, but students still perform them individually. The non-natural variations of communicative activities observed indicate that past prevailing methodology is still present. This can be explained by the fact that despite having good qualifications and experience, it still takes time to change teachers' beliefs, and implementing these changes in practice is another challenge. During the Soviet era, English was taught, much like Latin, for the purpose of reading and appreciating literature, not for communication. Nowadays, English is a 'lifeline' connecting Georgia to the outside world, practically a second language for many Georgians. Interestingly, the influence of the long-dominant grammar-translation method is so strong that even young teachers who didn't live during the Soviet period have "inherited" it from their educational upbringing. On the other hand, these attempts, often successful, show that teachers are aware of the need to reinforce communication in class and expose learners to authentic communication as much as possible. Alongside the many in-service training programs offered to teachers of English, the introduction of modern blended courses has greatly contributed to this process.

Another important change observed was the focus on authentic material provided by the courses. Surprisingly, despite the lack of necessary technical support needed to make the best use of what the platforms offer and with only several classrooms equipped appropriately, teachers still manage to expose their students to authentic audio and video material. Thus, modernized classrooms will increase students' and teachers' motivation and lead to better results.

In summary, based on the observation analysis, it is apparent that applying the blended method has served as a helpful tool to move forward towards the goals we have set. The comparison with traditional fully offline courses has proved that blended courses ensure better outcomes, particularly regarding saving class time for more communication tasks.

4. Conclusion

Based on the complex analysis of the study presented, certain conclusions were drawn, and recommendations suggested for specific educational contexts. However, the findings can be generalized, with some specifications, to other cases facing similar educational challenges:

- Present-day educational institutions are better adapted to blended learning (BL) courses due to rapid technological advancements, particularly following the online learning experiences necessitated by the Covid-19 pandemic.
- Blended English as a Foreign Language (EFL) courses should be introduced on a larger scale, considering the benefits they offer, such as:
 - Increasing motivation for both learners and teachers
 - Reinforcing a student-centered approach
 - Developing learner independence
 - Making assessments more objective, transparent, and result-oriented
 - Using class time more efficiently and effectively
 - Applying up-to-date authentic materials more effectively
 - Creating a technology-friendly environment conducive to modern learners, among other advantages.
- Blended EFL courses generally yield better outcomes compared to traditional, fully offline courses, especially in the context of modern education which emphasizes a combination of online and face-to-face teaching.
- Blended learning, which combines online and face-to-face teaching, benefits from communicative opportunities of the traditional classroom while incorporating an improved online component to enhance learning experiences.

- Blended EFL courses at ATSU have significantly improved since their introduction in 2017, contributing to raising students' language proficiency levels and resulting in an increased number of upper-level classes.
- The application of blended learning methods has further developed teachers' professional skills at ATSU.
- To facilitate a more seamless learning process, the number of technically equipped classrooms needs to be increased.
- The annual student registration process should be expedited and conducted in a timely manner.
- There should be greater emphasis on interactive, collaborative, and creative production activities both online and offline to enhance learning outcomes.

In summary, it can be concluded that the "Blended Method" has proven to be the most effective mode of instruction in the EFL context, especially for the generation equipped with 21st-century skills and expectations. Combining the power of human interaction in a real setting with built-in online components is the best way to use class time efficiently and productively. Integrating technology into the learning process is essential, but it should be done under the guidance of a teacher. Teachers are irreplaceable only if they possess relevant soft skills and can adapt to the modern-day demands of educational priorities.

Appendices

Survey links:

1. Students' survey

https://docs.google.com/forms/d/e/1FAIpQLSc4W3fXZ59qkr-9Mn8X66oCAqzr7gcgTys_sO-dVV9C_orQg/viewform?usp=sf_link

2. Teachers' survey

https://docs.google.com/forms/d/e/1FAIpQLSfxEakhibktl30CJwpwZNoI5snaHyVdKkLv_tB0lX8AGwNIuFw/viewform?usp=sf_link

References

- Adling, N. M. J. (2022). "Flexible Blended Learning Experiences of Filipino Students in an English as a Second Language (ESL) Learning Context." *Language Literacy: Journal of Linguistics, Literature, and Language Teaching*, 6:1: 28-38.
- Bandura, Albert. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Prentice-Hall.
- Badawi, M. F. (2009). "Using Blended Learning for Enhanced EFL Prospective Teachers' Pedagogical Knowledge and Performance." Paper presented at Learning & Language – The Spirit of the Age, Cairo: Ain Shams University.
- Chan, C. T., and Koh, Y. Y. (2008). "*Different Degrees of Blending Benefit Students Differently: A Pilot Study*."
- Dudeney, G., & Hockly, N. (2007). *How to teach English with technology*. Pearson.
- De Guerrero, M. C. M., and Villamil, O. S. (1994). "Social-Cognitive Dimensions of Interaction in L2 Peer Revision." *The Modern Language Journal* ,78:4, 484-496.
- Dziuban, C., Hartman, J., and Moskal, P. (2004). "Blended Learning." EDUCAUSE Center for Applied Research Bulletin. Available: <http://net.educause.edu/ir/library/pdf/ERB0407.pdf>
- Graham, C. R. (2006). "Blended Learning Systems: Definition, Current Trends, and Future Directions." In *The Handbook of Blended Learning: Global Perspectives, Local Designs*, edited by C. J. Bonk and C. R. Graham, 3-21. San Francisco: Pfeiffer.
- Hartman, J., Dziuban, C., and Moskal, P. (2007). "Strategic Initiatives in the Online Environment: Opportunities and Challenges." Vol. 15: 3, 157-168. Q Emerald Group Publishing Limited, ISSN 1074-8121.
- Kiezebrink, K. (2021). *Evaluation of Blended Learning*, University of Aberdeen.
- Krajka, J. (2012). *The language teacher in the digital age – Towards a systematic approach to digital teacher development*. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.
- Kwee, CH. T. T. and Dos Santos, L. M. (2022). How can blended learning English-as-a-secondlanguage courses incorporate with cultural heritage, building, and sense of sustainable development goals? A case study. file:///C:/Users/User/Downloads/feduc-07-966803%20(1).pdf
- Majeed, M., & Rehan Dar, F. (2022). *Investigating the efficacy of blended learning in ESL classrooms*. *Cogent Education*, 9(1). <https://doi.org/10.1080/2331186X.2022.2133500>

- Neumeier, P. (2005). A closer look at blended learning--Parameters for designing a blended learning environment for language teaching and learning. *ReCALL*, 17(2), 163-178. Available: <https://www.learntechlib.org/p/66364/>
- Olejarczuk, E. (2021). The E-Learning Component of a Blended Learning Course, Poznań University of Technology, Available: <http://www.tewtjournal.org>
- Sands, P. (2002). "Inside Outside, Upside Downside: Strategies for Connecting Online and Faceto-Face Instruction in Hybrid Courses." *Teaching with Technology Today* 8, no. 6 (March 20, 2002). Available: <http://www.uwsa.edu/ttt/articles/sands2.htm>
- Stockwell, G., & Tanaka-Ellis, N. (2012). Diversity in environments. In G. Stockwell (Ed.), *Computer-assisted language learning. Diversity in research and practice* (pp. 71-89). Cambridge: Cambridge University Press.
- Sun, Y., and Gao, F. (2020). "The Social Context of Learning: Exploring Students' Learning Experience in Blended Learning." *Educational Technology & Society*, 23:1, 108-119.
- Wang, et al. (2021). "How Can Blended Learning English-as-a-Second-Language Courses Incorporate Cultural Heritage, Building, and Sense of Sustainable Development Goals?: A Case Study." *Frontiers in Education*, Available: <https://doi.org/10.3389/feduc.2022.966803>
- Whitelock, D., and Jefts, A. (2003). "Editorial: Journal of Educational Media Special Issue on Blended Learning." *Journal of Educational Media* 28:2-3, 99-100. Available: <http://dx.doi.org/10.1080/135816503200017740>
- Wilson, M.-J., & Minhas-Taneja, V. (2016). Online Orientation – a Blended Learning Approach to Student Transition. *JANZSSA - Journal of the Australian and New Zealand Student Services Association*, 24(1), 61–68.
- Zviadadze, N., Demetradze, N., and Memani-Shvili, K. (2018). "Evaluating Blended Learning (A Case Study of ATSU)." *Journal of Teaching and Education* 8:2.

AI Driven Assessment and Feedback in English Language Teaching

Nida Gürbüz¹ Emine Yılmaz²

¹ Inst., Necmettin Erbakan University, Türkiye

² Inst., Necmettin Erbakan University, Türkiye

Abstract

This study conducts a comprehensive content analysis focusing on the landscape of assessment and feedback types in education, with a particular emphasis on the advantages and limitations of AI tools in this domain. The study meticulously categorizes and analyses the diverse array of assessment and feedback types prevalent in educational settings, examining traditional methods alongside emerging AI-driven approaches. By synthesizing existing literature and empirical studies, the paper identifies and evaluates the advantages that AI tools offer in assessing and providing feedback to learners, shedding light on their potential to enhance educational practices. Furthermore, the content analysis delves into a detailed examination of the AI tools that have predominantly been utilized in educational fields, mapping out their functionalities, features, and impact on teaching and learning processes. Through a systematic coding process, the research identifies common themes and trends in the use of AI tools for assessment and feedback within educational contexts, offering insights into their effectiveness and usability. In addition to highlighting the benefits of AI tools, the paper critically assesses the limitations and challenges associated with their implementation in educational settings. By conducting a comparative content analysis between human assessment practices and AI-based assessment methods, the study aims to provide a nuanced understanding of the strengths and weaknesses of each approach. Moreover, the research explores the perceptions of teachers regarding AI assessment and feedback tools, analysing their attitudes, concerns, and expectations towards the integration of AI technology in educational assessment practices. By synthesizing and interpreting the findings of the content analysis, this research contributes valuable insights to the ongoing discourse on AI in education. The paper offers practical recommendations for educators, policymakers, and technology developers, aiming to foster a deeper understanding of the role of AI tools in assessment and feedback processes and facilitate informed decision-making in educational contexts.

Keywords: Artificial intelligence (AI), assessment and feedback tools, technology

1. Introduction

There are clear advantages to using technology in language instruction (Neri et al., 2002; Benzies, 2017; O'Brien et al., 2018; Yoshida, 2018; Pokrivčáková, 2019; Rogerson-Revell, 2021; Vančová, 2021). The use of artificial intelligence (AI) in education has attracted a lot of attention lately and is completely changing the way that people teach and learn. Artificial Intelligence (AI)-driven tools have become powerful tools for improving assessment and feedback procedures as educational institutions increasingly integrate digital technologies. This change is indicative of a larger movement in education to use technology to personalize learning, enhance learning outcomes, and streamline procedures.

The purpose of this paper is to examine how assessment and feedback are changing in the field of education, with an emphasis on the function of artificial intelligence tools. Through a thorough content analysis, the research classifies and investigates different kinds of assessments and feedback, contrasting traditional techniques with cutting-edge AI-powered strategies. The study explores the benefits of AI tools, including their scalability, efficiency, and capacity to deliver tailored feedback in real time. Simultaneously, it evaluates critically the drawbacks and difficulties related to their application, encompassing worries regarding precision, openness, and possible prejudices.

The study maps out the effects of AI tools on teaching and learning processes by examining their features and functionalities in educational settings. The research provides important insights into the efficacy and usability of AI for assessment and feedback by identifying recurring themes and trends through a methodical coding process. The study also attempts to provide a detailed understanding of the advantages and disadvantages of AI- and human-based assessment techniques.

The paper also examines how teachers view AI assessment and feedback tools, examining their expectations, worries, and attitudes toward the use of AI technologies in educational assessment procedures. The research adds to the current conversation on artificial intelligence in education by synthesizing these findings and providing useful suggestions for teachers, legislators, and tech developers. With the ultimate goal of promoting a deeper comprehension of AI's role in assessment and feedback processes, this paper aims to support educated decision-making in educational settings.

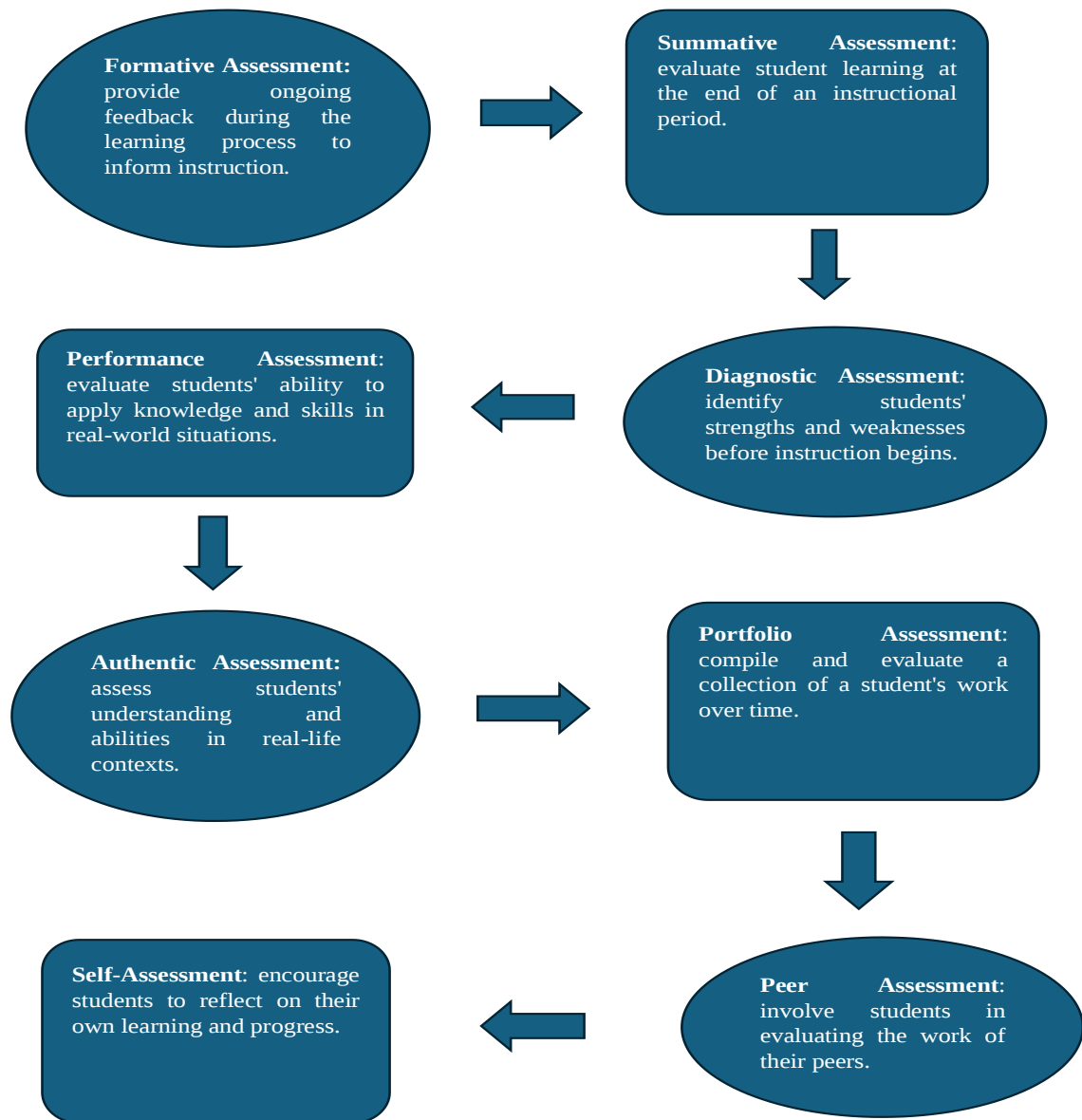
2. Literature Review

2.1 Assessment

A person's knowledge, skills, abilities, or performance in a certain field are evaluated systematically through an assessment process. It seeks to quantify the degree to which standards or learning objectives have been attained. Finding out a learner's degree of comprehension, skill, or accomplishment is the main goal of assessment. Grades or scores are frequently used to convey information about a student's knowledge and abilities.

Figure 1: Types of Assessment

Types of Assessment

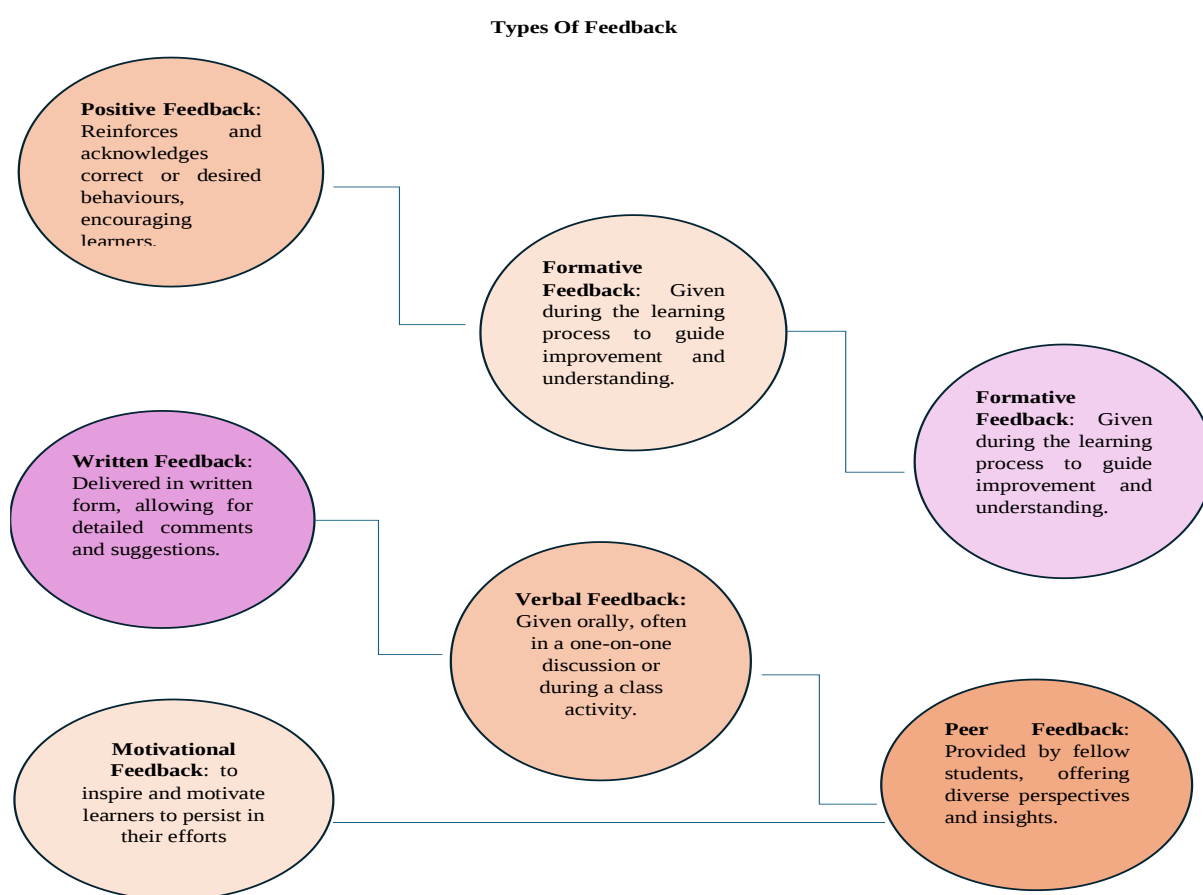


These assessment types can be employed individually or in combination, depending on the educational objectives and the nature of the learning experiences. Different assessments serve different purposes and contribute to a comprehensive understanding of students' knowledge, skills, and abilities.

2.2 Feedback

Feedback is information provided to a learner about their performance, offering insights into strengths and areas for improvement. It is a communication process that aims to enhance understanding and skills. The primary purpose of feedback is to inform learners about their progress, clarify expectations, and provide guidance on how to improve. Effective feedback is specific, constructive, and actionable. Feedback can be immediate or delayed. Immediate feedback is provided during or right after a task, while delayed feedback is given after some time has passed. Feedback comes in various types, and it can serve different purposes in the learning process. Here are some common types of feedback:

Figure 2: Types of feedback



Feedback plays a vital role in enhancing students' writing abilities and motivations. However, it can be a time-consuming and labour-intensive task for teachers in classrooms with a big number of pupils (Graham, Hebert, & Harris, 2015). One strategy to address the difficulty is to incorporate automatic writing evaluation (AWE) and automated speech recognition (ASR).

2.3 Automated Writing Evaluation (AWE)

The development of technology has resulted in the creation of several feedback systems for authors that are based on technology. AWE systems are at the forefront of these feedback technologies (Ware, 2018).

AWE is an online system that offers prompt grading and comments on students' written work (Shermis, Burstein, & Bursky, 2013). AWE utilises a comprehensive informational database to compare student writing. It employs statistical modelling and algorithms to analyse the linguistic, structural, semantic, and rhetorical aspects of the writing. AWE then generates a holistic score to evaluate the overall quality of the writing, and also provides detailed feedback on specific areas such as grammar, mechanics, style, discourse, and organisation (Ranalli et al., 2017). The AWE tools that are now popular worldwide include Criterion, e-rater, MyAccess, Write & Improve, Intelligent Essay Assessor (IEA), BETSY, Project Essay Grade (PEG), and WriteToLearn.

AWE provides students with the ability to post their assignments and receive comments at any time and from any location. Achieving successful implementation relies on comprehending the user acceptance procedures. AWE offers a comprehensive evaluation of writing quality, including detailed criticism on grammar, mechanics, style, discourse, and organisation. The platform provides personalised comments on the linguistic aspects of essays, such as technical quality, organisation, and development, the AI's high-speed and extensive processing capability relieves teachers from the task of correcting assignments, and its ability to accept multiple submissions of essay responses promotes learner autonomy.

Due to the aforementioned benefits, AWE can be utilised for both summative and formative examinations, as well as for generating instructional materials. The scoring function of the system enables it to collaborate with a human evaluator in significant and important exams like the TOEFL and the GRE (Weigle, 2013). It can also function as the only evaluator in less critical assessments and classroom settings. Students who received feedback from the Automated Writing Evaluation (AWE) system were inspired to make revisions to their written work. An analysis of the AWE-based process-writing program revealed that students demonstrated a reduction in grammatical errors in both their revisions and newly written compositions. AWE, or Automated Writing Evaluation, is a technological tool that provides instant feedback on writing and serves as a powerful instructional resource for enhancing students' cognitive writing skills. Automated scoring tools have the potential to significantly decrease the amount of labour teachers have to do and speed up the process of giving feedback to students. Recent technology breakthroughs have led to the development of automated writing evaluation (AWE) programs. These technologies can grade student compositions and provide immediate diagnostic feedback, as well as consistent explanations. These tools not only decrease the amount of labour teachers have to do, but also allow them to focus their time on more advanced pedagogical duties, such as creating classes that accommodate diverse learners and promote motivation for learning. AWE tools can enhance the writing process and facilitate autonomous learning by offering personalised feedback, a feature that is not possible in traditional language courses. Prior studies have demonstrated that these tools have a positive impact on students' writing proficiency, their capacity to make successful revisions to their writing, their enthusiasm to write, their attitudes towards writing, and their belief in their own

writing abilities. Below are several AWE tools that may be utilised for the purpose of providing feedback or grading.

2.3.1 Criterion

Several computer applications, such as Criterion by ETS, MyAccess by Vantage Learning, and WriteToLearn by Pearson Knowledge Technologies, have been created in order to offer students with immediate and individualised feedback on the written texts that they have produced (Chen & Cheng, 2008; Dikli, 2006; Warschauer & Grimes, 2008). Criterion is a web-based program that was designed by Educational Testing program (ETS) for the purpose of evaluating writing skills and delivering fast diagnostic feedback and score reports immediately after the evaluation. The holistic ratings that Criterion provides for English as a Foreign Language teachers demonstrate a considerable increase not just between the two submissions (the first draft and the final draft of the same writing), but also on a more general level throughout the drafts (Ebyary and Windeatt, 2010). According to Attali's Attali (2004) study, the feedback that Criterion offers to students is successful in helping them improve their writings. This feedback helps students improve their essays by lowering the number of errors they make in grammar, use, mechanics, and style, as well as increasing the length of their essays (Liao, 2015). A holistic score that reflects the overall quality of the writing sample that was given by users is provided by Criterion.

Additionally, thorough diagnostic feedback on faults in a variety of areas is provided within a matter of seconds. Individual reports are provided to teachers, allowing them to evaluate the progress of their pupils more effectively. In addition to this, it offers students a thorough "Writer's Handbook" that contains definitions of feedback, explanations of problems using examples, and ideas on how to rectify the flaws as well as how to enhance the organisation and growth of the writing (Dikli, 2006). The Criterion subject library has more than 400 writing prompts at varied levels, ranging from fourth grade to university level. Teachers have the ability to give a writing prompt from this library to their students. A "Comment" or "Dialogue" area is available for teachers to use in order to provide feedback on the works of their pupils (ETS, 2019). A total number of errors that were committed in each category (such as grammar, use, mechanics, and style) and subcategory (such as subject-verb agreement, preposition error, missing comma, and so on) are included in the error analysis reports that are generated and recorded by the application. The amount of mistakes that fall within each error category and subcategory is displayed by Criterion when students hand in their essays. The problematic sections are highlighted. This occurs when the pupils click on the subcategories that display an error. The pupils are presented with advising messages about how to repair the issue simply by moving the mouse over the highlighted areas. The proper form is not immediately offered to them. On the other hand, the software, through its suggestions, can occasionally provide the appropriate form. Within a span of two days, the students submitted their essays for the purpose of instructor review after making the necessary corrections to their errors based on the input concerning the Criterion (Han & Sari, 2022). College students were able to minimise the amount of errors they made in four different categories of grammatical faults (fracture, subject-verb agreement, run-on sentences, and ill-formed verbs) in both revisions and new texts by using the criterion, despite the fact that the effect was different for each category of errors. instrument for evaluating kids' writing that is supervised by teachers and delivers immediate diagnostic feedback on the essays they have written.

The scoring system is comprised of two applications that are complimentary to one another, which makes it appropriate for use in a classroom environment. The first one is called Critique, and it is a writing analysis tool that enables the system to identify errors. These errors are classified as "grammar," "usage," "mechanics," and "discourse elements of organisation and development." These are the terms that ETS uses to describe the specific categories within which the system provides linguistic feedback, also known as "trait feedback." The electronic rater is the second application. In order to assign a holistic score to an essay, which is only provided when the writing prompt is selected from a set of given Criterion topics, e-rater functions as an automated essay scoring system. It does this by employing a statistical model to extract linguistic features from an essay and then relating those features to the overall quality of the writing. Therefore, the two programs, Critique and e-rater, together offer two different kinds of feedback on writing that may be used for instructional reasons in language classes. These feedbacks are in the form of trait feedback and a holistic score (Ramineni & Williamson, 2013). Analytical components of the essay, such as content representation, organisation, syntactic complexity, syntax, and mechanics (such as punctuation, capitalisation, and spelling), are included in the explanation of the score. Criterion, on the other hand, does not offer analytical scores for the components that are listed in its rubric. As a result, it is necessary to compare the holistic score to both the numerical grades and the analytical ratings that are provided by the teacher. The scores that students received on the criterion served as an effective incentive that encouraged students to take an active role in the writing process. Students have the opportunity to become more conscious of their duty in paper writing and, as a result, experience the significance of progress in writing quality. As a result, it is expected that the scores can play a promising function as an instructional tool to boost students' learning. Criterion scores appeared to be helpful for teachers to pre-evaluate the papers that their students had written before they actually read them. This is due to the fact that low Criterion ratings may be utilised to draw attention to a variety of grammar errors or a lack of content in the essays that students had written (Li et al., 2014).

Figure 3: Criterion's display of identified errors

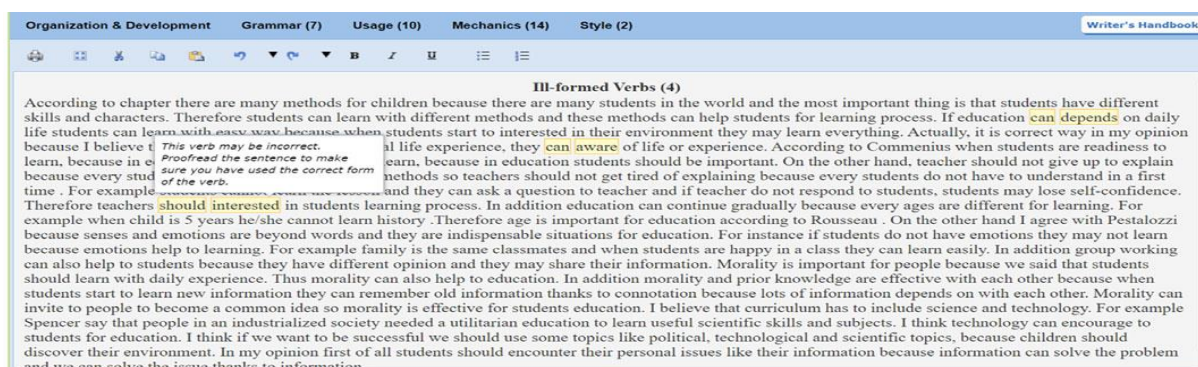
Organization & Development	Grammar (7)	Usage (10)	Mechanics (14)	Style (2)	Writer's Handbook
   	Fragments		 		
	Run-on Sentences				
	Garbled Sentences			Ill-formed Verbs (4)	
	Subject-Verb Agreement (2)				
	Ill-formed Verbs (4)				
	Pronoun Errors				
	Possessive Errors				
	Wrong or Missing Word				
	Proofread This! (1)				

According to chapter there at learn with different methods environment they may learn . Comments when students a because every students may lesson and they can ask a que can continue gradually becau hand I agree with Pestalozzi I help to learning. For example they may share their informai knowledge are effective with Morality can invite to people industrialized society needed topics like political, technological and scientific topics, because children should discover their environment. In my opinion first of all students should encounter their personal issues like their information because information can solve the problem and we can solve the issue thanks to information.

because there are many students in the world and the most important thing is that students have different skills and characters. Therefore students can adents for learning process. If education can depends on daily life students can learn with easy way because when students start to interested in their act way in my opinion because I believe that if students start to use practical life experience, they can aware of life or experience. According to n education students should start to learn, because in education students should be important. On the other hand, teacher should not give up to explain so teachers should not get tired of explaining because every students do not have to understand in a first time . For example students cannot learn the do not respond to students, students may lose self-confidence. Therefore teachers should interested in students learning process. In addition education learning. For example when child is 5 years he/she cannot learn history .Therefore age is important for education according to Rousseau . On the other re beyond words and they are indispensable situations for education. For instance if students do not have emotions they may not learn because emotions and when students are happy in a class they can learn easily. In addition group working can also help to students because they have different opinion and people because we said that students should learn with daily experience. Thus morality can also help to education. In addition morality and prior ents start to learn new information they can remember old information thanks to connotation because lots of information depends on with each other. morality is effective for students education. I believe that curriculum has to include science and technology. For example Spencer say that people in an useful scientific skills and subjects. I think technology can encourage to students for education. I think if we want to be successful we should use some information can solve the problem and we can solve the issue thanks to information.

Source: Saricaoglu, A., & Bilki, Z. 2021

Figure 4: Criterion's feedback on a highlighted ill-formed verb error

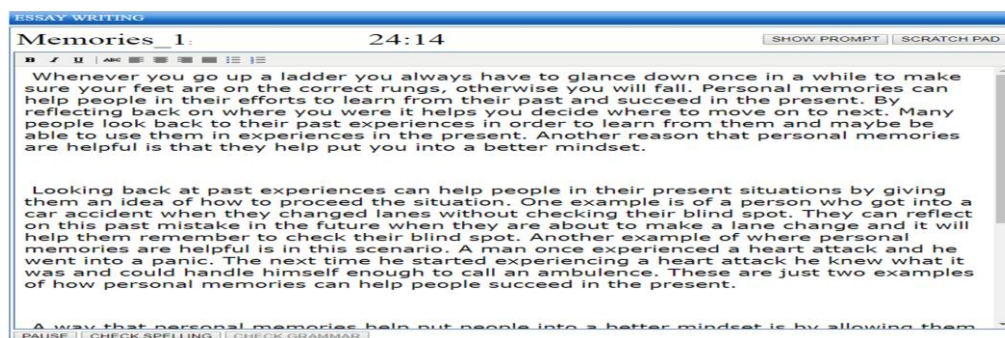


Source: Saricaoglu, A., & Bilki, Z. 2021

2.3.2 Writing Pal

Writing Pal provides feedback that is unique to the activities and methods involved in writing. It is possible to obtain input at both levels from other systems, such as Criterion. It is likely that different kinds of feedback and combinations of input will lead to different outcomes, both in terms of the perceptions of the students and the outcomes that are more objective. An automated writing evaluation (AWE) tool and an intelligent teaching system (ITS) that teaches writing methods and permits numerous types of writing practice (such as game-based strategy practice and essay writing) are both included in this system. Freewriting, planning, introduction building, body building, conclusion building, unity, paraphrasing, and revision are the eight main parts of writing that are targeted by W-Pal training. An essay writing interface with icons for checking spelling and grammar incorporated into Writing Pal. Through the use of prewriting, draughting, and editing, formative feedback messages in W-Pal provide recommendations for concrete tasks and techniques that may be utilised to improve an article. Participants in the Strategy + SG Condition were provided with the opportunity to use the "Check Spelling" and "Check Grammar" buttons conveniently located at the bottom of the essay window (Fig. 3). The participants were able to utilise either function whenever they needed to while they were writing or rewriting. W-Pal is capable of automatically assigning qualitative evaluations to essays (McCarthy et al., 2022).

Figure 5: "Check Spelling" and "Check Grammar" buttons at the bottom of the essay window on Writing Pal



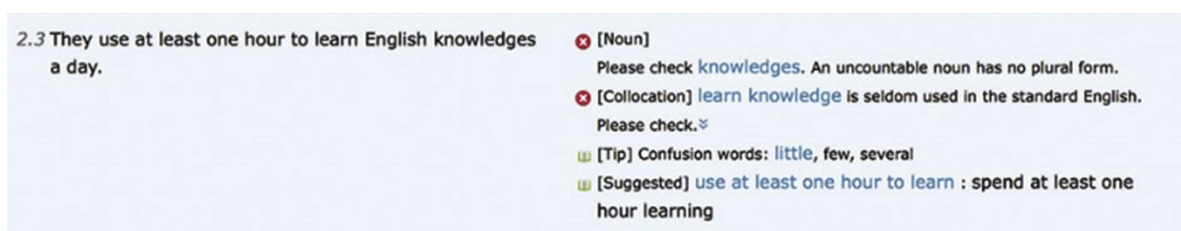
Source: McCarthy, K., Roscoe, R., Allen, L., Likens, A., & McNamara, D., 2022

2.3.3 Pigai

Pigai, which is a rather well-known AWE system in China, offers its customers free writing assessment and feedback services. It is an artificial intelligence and machine learning system, which was introduced by Beijing Ciku Science and Technology Co., Ltd. in 2016. There were over 20 million users of the AWE by the end of 2019, coming from over 6,000 different schools. For the purpose of this investigation, the students were given the opportunity to send the writing portion of the practice exam that they had completed to the AWE for an evaluation and enhancement. A holistic score, a rank with four-category descriptions (vocabulary, sentence, content, and organisation), an overall remark on their writing skills, sentence-by-sentence advice for corrections, and further learning recommendations are all included in the information that is provided by the AWE. Taking into consideration the comments made by the AWE, the students edited their works and resubmitted them.

In addition to providing input that may be taken into consideration, it also provides holistic grading, ranking, top and lowest scores, and final remarks. (Zhang, 2017). A vast corpus that includes standard English, essays written by students in English, and other English textbooks is used to calibrate its scoring model. An essay is given an overall score by Pigai, which is determined by calculating the quantitative differences between the essay and the texts in its corpus in reference to four categories: vocabulary, sentence structure and organisation, structure and organisation, and topic relevance. On the other hand, in contrast to other well-known AWE systems (such as Criterion® and MY Access!®), Pigai provides corrective input on collocations. demonstrates that Pigai has identified the phrase "learn knowledge" as a collocation problem; students can click the "please check" button to obtain further information (Zhang & Zhang, 2018).

Figure 6: An example of corrective feedback on collocation in Pigai



Source: Zhang, 2020

2.3.4 Grammarly

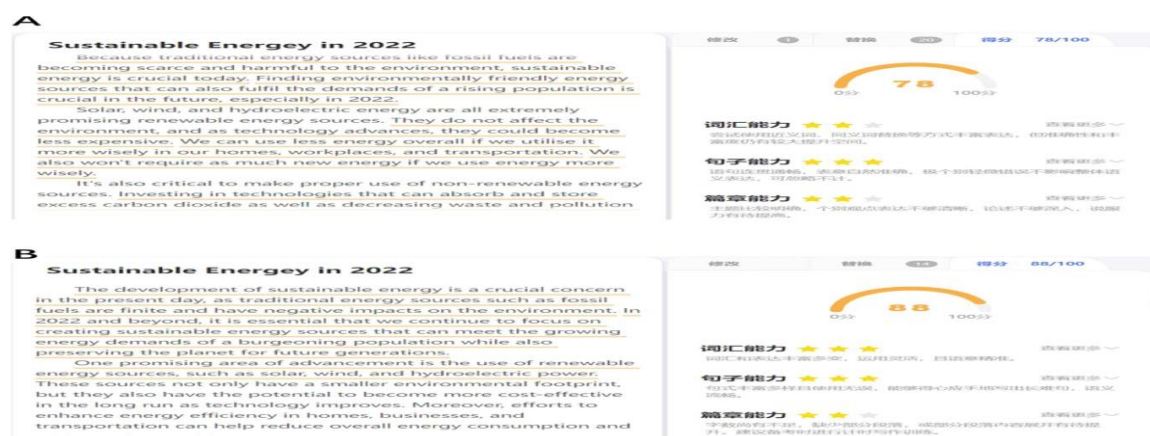
The free edition of Grammarly offers comments on spelling, punctuation, grammar, and norms, such as spacing, capitalisation, and spelling that is specific to a particular dialect. After a document has been posted to the internet, Grammarly immediately delivers comments for corrections and improvements. Indirect feedback is displayed on the right side of the screen, while direct feedback is displayed on the left side of the screen. The document that was submitted is displayed on the left side of the screen where mistakes are highlighted in red. Direct feedback includes the type of error (for example, grammar) as well as the possibility of correcting the issue (Koltovskaia, 2020). In a variety of contexts, including educational, professional, and everyday settings, tools like as Grammarly and ProWritingAid are utilised extensively. Automated written corrective feedback is something that they provide.

2.3.5 ChatGPT

OpenAI introduced ChatGPT in November 2022 as an advanced chatbot that utilises the GPT-3.5 series of extensive language models (OpenAI, 2022). ChatGPT, an artificial intelligence-powered chatbot that is capable of automatically generating text, was made available to the general public. Essays may be written by programs that are driven by GPT, and these programs are also able to fix problems that are present in essays. When given a straightforward request, ChatGPT and Bard are able to generate comprehensive essays that are capable of being accepted at the university level. It has been demonstrated that GenAI is capable of performing a wide range of tasks, such as writing essays or creative texts such as poems or stories, writing or correcting computer programming code, answering questions, summarising and paraphrasing provided text, and synthesising different tones and styles in order to generate new and creative texts. It is common practice for major examinations like the GRE and TOEFL to rely on AI-enabled AWE algorithms in order to score a significant proportion of essays. It has been determined that one of the possible strengths of GenAI (023) is the ability to provide personalised learning by analysing and providing feedback on essay writing. This, in turn, can help reduce the amount of work that teachers have to do and prevent them from being burned out. Undergraduate students can receive constructive criticism on their writing using ChatGPT. One other research that looked into ChatGPT for the purpose of evaluating and providing feedback on essays was conducted by Mizumoto and Eguchi. They fed a corpus of 12,100 essays written by non-native English writers to ChatGPT and evaluated the effectiveness of rubric-grounded comments and scores. According to the findings, ChatGPT has a level of reliability and accuracy that is satisfactory. These studies provide evidence that the use of GenAI tools such as ChatGPT for the purpose of AWE is both feasible and reliable. However, there is a need for further research to be conducted in order to better understand the effectiveness of ChatGPT for AWE and how students perceive its usage. Therefore, it is possible that ChatGPT is capable of functioning as an AWE tool, provided that a grading rubric or other criteria are supplied.

Some of the writings that were created by ChatGPT received attention for their general quality, their uniqueness in terms of content, and their performance in terms of composing literary evaluations. Students were generally happy with the quality of the created content in terms of language accuracy, syntactic coherence, grammatical correctness, language styles, and compliance with academic criteria, according to the self-evaluation that was conducted. On the basis of the experiences that students had, the following four primary categories of software or applications were utilised by students during the field experience: 1) Text generators, which were essentially solutions that were comparable to ChatGPT; 2) Paraphrasers, which offered functions to automatically rewrite, paraphrase, or patch-write the original writing in order to improve the quality of the text or reduce the likelihood of plagiarism detection; 3) Grammar checkers, which offered automatic grammatical checking as well as suggestions for amendments and corrections; and 4) Summarisers, which assisted students in transforming the text from a style that was redundant and repetitive to one that was more concise (Yan, 2023).

Figure 7: Automatic grading of direct output (A) from ChatGPT's text generation and an improved version (B)



Source: Yan, D., 2023

2.3.6 Notion AI

The freemium program known as Notion allows users to manage their work on a single platform in both an individual and collaborative manner. On the other hand, Notion AI is an artificial intelligence-powered writing assistant that edits users' manuscripts in terms of lexicogrammatical structure, cohesiveness, and structure. E-portfolios are a digital collection of written scripts that have been selected over a period of time. They serve two purposes: they encourage self-regulation by providing a place that is learner-centred and reflective, and they provide an interactive environment that scaffolds writing processes by providing feedback from many sources. the usage of Notion as a platform for electronic portfolio sharing. Users are able to handle projects both individually and cooperatively on a single platform through the use of the freemium program known as Notion. Users have the ability to compose content on a page that has been produced, and then submit it to Notion AI for editing and proofreading (see Figure 6). The creation of a team area makes it simple to share pages for the purpose of cooperation.

Notion AI is a writing assistant that is driven by artificial intelligence and can edit or produce material in response to orders such as correcting spelling and grammar, adjusting the length of text, adjusting the tone according to the goal, translating, and summarising. One of the commands that users might provide is to highlight the text that is being targeted. This complimentary usage of Notion and Notion AI may act as an e-portfolio platform with an AWCF generator, which assists learners in promoting self-regulation and noticing, as well as receiving scaffolding in ways that other AI writing tools are unable to provide. On a single page, Notion documents the entirety of the writing process, including composing drafts, amending them, and getting comments on them. The function that displays all of the modifications that have been performed is called "updates," and it is what makes this process more obvious. These electronic portfolios, in which students are able to evaluate and report on their growth as writers, help students develop their capacity for self-regulation. Making portfolios that are rich in feedback can help assist writing improvement, which in turn promotes "analytics" discourse-level revision behaviour (Lam, 2013). This is relevant to the evaluation of portfolios for second language learners. Additionally, the tool, which enables instructors to observe how frequently students visit a certain page, can assist teachers in monitoring and improving the level of engagement that students have with the material.

Figure 8: Editing and proofreading of the application of Notion AI

Draft 2

One reason is Notion can make whole writing process visible, which help learners monitor and reflect progress.

AI feedback (Spelling and grammar)

One reason is that Notion can make the entire writing process visible, which helps learners monitor and reflect on their progress.

- "One reason is that" instead of "One reason is"
- "Notion" capitalized
- "entire writing process" instead of "whole writing process"
- "helps" instead of "help"
- "learners" instead of "learner"
- "monitor and reflect on their progress" instead of "monitor and reflect progress"

Source: Osawa, K., 2023

2.4 Automated Speech Recognition (ASR)

ASR stands for Automatic Speech Recognition, which is a type of software that can accurately recognise spoken words and turn them into written text. Examples of similar technology are Rosetta Stone's integrated speaking tasks and the iPhone's Siri assistant. Some ASR-based applications offer automated, immediate, and personalised support to help students enhance their pronunciation. Lately, there has been a growing interest in automatic speech recognition (ASR) systems, such as Google Speech Recognition, Windows Speech Recognition, and AT&T Watson. The study found that using Automatic Speech Recognition (ASR) feedback for pronunciation education through dictation was more effective than using input from an instructor or no feedback at all.

ASR has the ability to reduce speech anxiety, which is often considered as a hindrance to enhancing the oral performance of EFL learners. There is a complete absence of fear of failing and any sense of embarrassment when making mistakes throughout the exercise. Furthermore, learners are granted increased autonomy in their practice, enabling them to gradually develop their confidence in public speaking. ASR would grant users a digital badge at the completion of each job, providing them with a feeling of fulfilment and ensuring their continued involvement in the training. When users require to deliver a speech at a certain event, such as a group gathering, ASR can provide them with a useful and structured draft template with predefined sentences, frameworks, and suggested ideas. Primarily, the absence of genuine audiences, while providing a sense of relaxation, can impede the progress of proficient learners. ASR-based software offers automated and personalised feedback on learners' pronunciation, a level of achievement that is seldom attained in the conventional classroom. Crucially, this rapid and automated feedback hinders the formation of inaccurate pronunciation patterns. It is possible for an automatic speech recognition (ASR) technology to offer some comprehensive feedback on the participants' speech, which is referred to as explicit corrective feedback. For instance, the applications (such as Praat and Speech Analyser) might provide an illustration of the speech waveform and its spectrogram on the pronunciation of the pupils. Through the use of these graphical representations of speech, students were able to comprehend the reasons behind and the manner in which their pronunciation was comparable to or unlike to that of a native speaker. For another illustration, the automatic speech recognition (ASR) applications, such as SpeechAce and Fluent English, might send out some feedback messages. These

messages could include the scores that the students received for each word, each syllable, and each phoneme that they said. Liu, Zhu, Jiao, and Xu (2018) and Moxon (2022) found that students were able to precisely comprehend the exact regions of their pronunciation faults about various words, syllables, or phonemes, as well as the level to which they performed (for example, good, not bad, or wrong) on each specific pronunciation. These messages were helpful in achieving this understanding. Indirect feedback is a sort of feedback that is offered by automatic speech recognition (ASR) tools. This type of feedback consists of the tools only displaying the text that is transcribed from what the students say, or providing straightforward positive or negative replies (such as "correct" or "incorrect") to the student's voice input. Several automatic speech recognition (ASR) systems, such as Speechnotes, Windows Speech Recognition, and PARLING, are examples of programs that offer this kind of feedback. Students are able to identify the words that they have mispronounced by looking at the text that is presented or the feedback answers and then try again.

Moreover, the utilisation of ASR software is said to be gratifying and pleasurable, as well as highly beneficial for enhancing pronouncing skills. It enhances the pronunciation skills of both young and adult learners who are studying English as a Second Language (ESL). ASR, or Automatic Speech Recognition, shows great potential as a method to enhance pronunciation skills in language learners. Bashori et al. (2022) discovered two main benefits of utilising ASR in the classroom: offering a superior pronunciation model compared to that provided by non-native instructors. The second aspect pertains to the proficiency of learners in the foreign language setting to surmount the challenge of locating a suitable interlocutor for communication. Nevertheless, their findings have demonstrated that a meticulous choice of materials and adequate teacher training are crucial for enhancing the pronunciation skills of foreign language learners. Bajorek (2017) reached a similar finding and suggests incorporating ASR technology as a supplementary tool for conventional classroom pronunciation instruction.

Utilising Automatic Speech Recognition (ASR) for dictation in a classroom setting results in notable improvement in pronunciation skills and enhances the clarity of pronunciation. In addition, the utilisation of Automatic Speech Recognition (ASR) technology is advantageous for language learning classes since it establishes a stress-free setting where students are not compelled to speak in front of their peers. An automatic speech recognition (ASR) software was utilised to offer immediate feedback on pronunciation.

2.4.1 Text to Speech

Text-to-speech translates written words into spoken words and is easily available in a variety of technologies. The correctness of the pronunciation varies depending on the speech engine being used. Text-to-speech functionality is already available in Google Docs, and certain educational institutions and lecturers actively promote its utilisation. For the purpose of providing assistance to students who are still in the process of developing their reading skills or students who have reading difficulties, text-to-speech should be utilised. This component of speech processing should, ideally, provide students the ability to pick particular words or passages of text to be read out loud to them, for instance by a pedagogical agent (Mayer, 2009). Watson Text to Speech (IBM), Polly (Amazon), Text to Speech (Microsoft Azure), Select and Speak (iSpeech), Vocalware, and TTSReader are some examples of text-to-speech software that is easily available among the available options. Windows application programming interfaces (APIs) are provided by the bulk of these cloud-based solutions. Both text to voice

and speech to text function with a distinct variety of input, as is obvious from the fact that there is a clear distinction between the two. Text is used as input for text to speech, whether it is already shown on the screen or written by the user themselves. Voiced noises are used as input for the speech recognition module that is housed within STT. There is the possibility that YouTube might be used as a learning tool with peer evaluation and feedback. Peer feedback provided online is another way that collaborative learning settings are supported. According to Nicols (2012), the usefulness of peer feedback from a social point of view may be increased by employing cognitive process tactics such as monitoring, assessing, and practicing one's own knowledge. Commenting, annotating, like, and replying are some of the features that are available on video-sharing sites, particularly YouTube. These features allow users to provide comments or engage in debate.

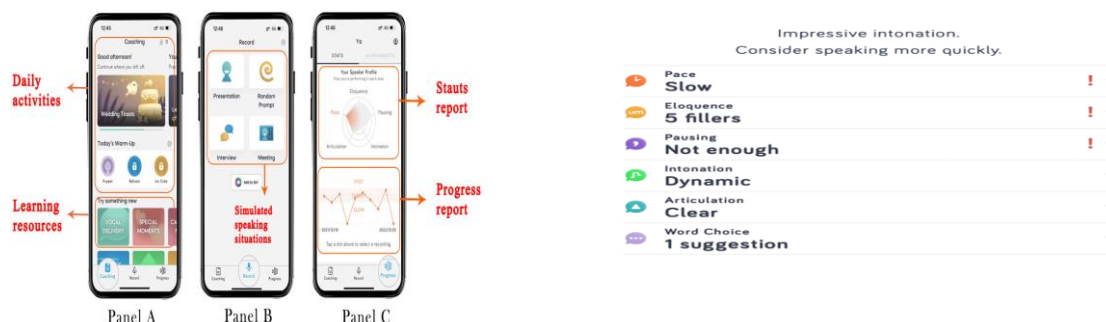
2.4.2 Speeko

Speeko is a training tool that is dedicated to assisting individuals in becoming more self-assured, unambiguous, and compassionate when it comes to public speaking (Speeko, 2022). At the moment, the application may be utilised on devices that are compatible with either macOS or iOS. The software may be downloaded for free, but in order to use the more sophisticated features, a membership is required. There are three primary components that make up this. There is coaching, recording, and progress. There are various daily exercises and thorough training routines that are hosted by the coaching module. The Progress module offers learners insightful learning metrics, while the Record module gives learners the opportunity to exercise in realistic circumstances and receive feedback on their performance. The functionality of Speeko includes (a) the supply of instant and individualised feedback; (b) the decrease of students' fear over public speaking; and (c) the provision of an organised learning experience. its users. Instant, individualised feedback based on performance is something that Speeko is able to provide. Learners might be motivated to put in more effort into their learning when they receive feedback that is both responsive and customised. This can ultimately lead to increased learning performance (Lim et al., 2021; Van der Schaaf et al., 2017). In the event that students generate voice recordings inside the application, Speeko is able to examine their oral performance in a short amount of time with the use of automatic speech recognition technology and provide detailed feedback in response to the recordings.

Speeko's capacity to reduce speech anxiety, which is generally seen as a barrier to enhancing the oral performance of English as a Foreign Language (EFL) learners, is another significant asset of the program. In light of the fact that the speakers are being guided by a virtual coach, there would be no anxiety about failing, nor would there be any humiliation about making mistakes while they are practicing.

Additionally, learners are given greater control over their practice, which allows for the gradual development of their self-assurance in communicating in front of an audience. Users of Speeko would receive a digital badge for each activity that they successfully completed, which would provide them with a sense of accomplishment and ensure that they remain interested in the training. When users are required to speak at a particular event, such as a group gathering, Speeko is able to provide them with a draft template that is comprehensive and well-organised, complete with phrases, frameworks, and thoughts that have been offered. Initially and most importantly, the absence of actual audiences, despite the fact that they are stress-free, may be a barrier to the continued growth of advanced learners (Yu-Chih, 2008).

Figure 9: The modules and feedback of the application Speeko



Source: Mei, B., Qi, W., Huang, X., & Huang, S., 2022

Learning programs that are driven by artificial intelligence are typically made available as mobile applications for both general and specific groups of learners (for example, Bussuu is designed for beginners, and Duolingo is perfect for youngsters who are talented, according to Hazar, 2022). Mobile applications offer a multimodal practice and presentation of language, along with an endless number of exercises. These applications respect the students' level of skill and allow them to advance at their own speed in both official and casual situations. In instance, Alfuhaid (2021) found that the usage of Duolingo can help learners improve their self-confidence in oral communication. This is due to the fact that users of the platform are already accustomed to talking in the target language. According to Portnoff et al. (2021), the practice is self-directed, and the applications can profit from data mining and learning analysis. In addition, they may be utilised in educational environments to ensure that students are adequately prepared for examinations and assessments (Brick & Cervi-Wilson, 2019).

2.5 Limitations of the Tools

Students stated that some of the diagnostic feedback messages were erroneous and unclear, and that the software failed to catch several fundamental mistakes in grammar and use. This was in relation to the accuracy of the messages. It is important to note that in these studies, AWE was used in isolation from teacher feedback. Furthermore, it was not made clear whether the students who participated in the study possessed the necessary skills to use technology in their instruction or whether they were given training on how to use the AWE tools that they were exposed to. This has an effect on the experiences that students have with using AWE systems. One of the most common complaints that students have regarding awe tools is that they do not allow for feedback on the curriculum. It was stated by the students that the feedback that was provided in the field of organisation and development was not clear. According to the kids, the advising messages that display when you click on the colour codes are not useful and include a lot of repetition.

3. Methodology

The study employed a meta-analysis approach to analyse specific research articles that were published in academic journals and conference proceedings, in order to achieve the study's aims. The primary criteria for selecting papers for analysis were search results that used the

keywords "artificial intelligence" and "writing and speaking". Alternative terms such as "AI-powered", "AI assessment and feedback", and additional terms like "ELT", "EFL", and "ESL" were also considered. These alternative terms were chosen because they encompass both the technological and language aspects of artificial intelligence and productive skills. As previously stated, these characteristics are integral components of their respective systems and cannot be detached from the broader environment in which they are employed. In this current evaluation, AI assessment and feedback refer to two specific activities: (1) enhancing students' pronunciation accuracy and (2) increasing students' writing skills.

In order to choose the appropriate publications for this review, the following databases were searched: Web of Science, Springer Link, Elsevier, SAGE, ERIC, Taylor & Francis, RELC, and Wiley Online Library. Additionally, Google Scholar was also searched using specified keywords. Additional factors considered were the online accessibility of the complete text, the nature of the research, and the language used in the study (namely English). Papers that include evaluations, detailed explanations of systems, and finished research projects were not included in the selection. A time limit for paper publishing was not established because the AI technology is considered relatively new in its application to AWE and ASR. It is believed that allowing for a longer period of use beyond current pedagogical practices could be advantageous for understanding the context and observing trends in its application. After carrying out both stages of the selection process, a sample of 20 papers was created.

4. Discussion

The implementation of Automated Writing Evaluation (AWE) systems in educational environments has shown some notable constraints that affect their efficacy and overall usefulness. The identification of these difficulties has been greatly aided by the feedback supplied by students, with a specific focus on the precision, comprehensibility, and pertinence of the input generated by these systems. These concerns are of utmost importance, since they directly impact the learning results, and the overall user experience using AWE tools.

A major concern expressed by students is the precision of the diagnostic feedback. According to students, the feedback signals produced by AWE systems were found to be incorrect and ambiguous, since they did not identify basic errors in grammar and use. This error is especially troublesome as it not only weakens the trust of the tool but also poses a danger of perpetuating erroneous linguistic patterns. If students are given feedback that fails to address crucial faults or offers inaccurate remedies, it can interrupt their learning process and potentially result in more confusion rather than improvement. The capacity of AWE systems to precisely detect and rectify errors is essential, however, this is the area where existing tools appear to be inadequate.

The clarity and use of comments are also important considerations. Students conveyed that criticism, particularly in aspects such as organisation and growth, frequently lacked clarity. The advising messages, intended to facilitate students' comprehension and rectification of errors, were regularly subjected to criticism for their repeated nature and lack of practical assistance. This problem is made worse by the characteristics of AWE tools, which depend on pre-programmed algorithms that may not consider the subtle requirements of each learner. Consequently, the input may become excessively general, missing the precise instructions required to make significant improvements. Repeating advise without providing further

explanation can sometimes result in irritation, since students may find it difficult to understand what is required of them.

AWE tools suffer from a significant drawback, namely the absence of feedback that is directly applicable to the curriculum. Students frequently expressed dissatisfaction with the feedback they received, as it did not correspond to the precise learning objectives outlined in their curriculum. This disconnection diminishes the efficacy of AWE systems in facilitating students' attainment of their educational objectives. Curriculum alignment is crucial in educational tools since it guarantees that the feedback students get is directly relevant to their classroom learning. In the absence of this alignment, AWE tools may provide general guidance that, although possibly beneficial in certain situations, does not cater to the specific demands of the students' assignments. This constraint emphasises a substantial disparity between the functionalities of AWE technologies and the requirements of the educational setting.

Moreover, the analysed research suggest that AWE tools were frequently utilised without receiving instructor feedback, perhaps leading to the unpleasant experiences described by students. Teacher feedback is extremely useful as it offers personalised and contextually relevant instruction that cannot be replicated by automated solutions. When AWE tools are utilised in the absence of instructor involvement, pupils may be deprived of the profound understanding and personalised guidance that a teacher can provide. This issue is exacerbated by the lack of clarity on the students' proficiency in utilising AWE tools and their level of training in using these systems. The absence of proper training might have a substantial impact on the students' capacity to get benefits from AWE, since they may lack a comprehensive understanding of how to evaluate or implement the feedback given.

The significance of incorporating AWE (Automated Writing Evaluation) tools with instructor input and ensuring that students receive sufficient training to utilise these tools is of utmost importance. Although AWE tools can be a useful addition to conventional writing education, their existing constraints indicate that they are not yet capable of substituting the intricate and personalised feedback that teachers offer. Utilising AWE tools in conjunction with instructor input through an integrated strategy might effectively address these challenges, offering students a more thorough comprehension of their writing skills and areas that need work. Furthermore, it is crucial to provide students with enough training in the use of AWE tools to guarantee that they can fully optimise the advantages offered by these systems.

5. Conclusion

To summarise, whereas Automated Writing Evaluation systems hold the potential to improve the writing teaching process by providing prompt and consistent feedback, their existing limitations greatly impact their usefulness. Students have reported many main problems, including erroneous feedback, lack of clarity and usefulness in advising messages, and the absence of curriculum-relevant counsel. These issues are worsened when AWE tools are utilised without the assistance of instructor feedback and insufficient student instruction. In order to fully maximise the effectiveness of AWE tools, it is crucial that they are utilised in a way that enhances, rather than substitutes, traditional teacher feedback. Teachers have a vital role in offering the specific and personalised assistance that automated systems currently do not provide. Furthermore, it is essential for students to possess the requisite skills and expertise to utilise these technologies proficiently. The full realisation of the potential benefits of AWE

tools, resulting in more meaningful and productive writing experiences for students, can only be achieved through an integrated and well-supported strategy. Continual research and development will be essential in overcoming the limits and improving the capabilities of AWE systems as educational technology advances.

References

- Attali, Y. (2004). Exploring the feedback and revision features of Criterion. Paper presented at the National Council on Measurement in Education (NCME), San Diego, CA.
- Benzies, Y. J. C. (2017). Contributions of new technologies to the teaching of English pronunciation. *Language Value*, 9(1).
- Brick, B., & Cervi-Wilson, T. (2019). Enhancing learners' professional competence via Duolingo classroom in professional competencies in language learning and teaching (pp. 19-29). Research-publishing.net.
- Chen, C., & Cheng, W. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112. <http://doi.org/10.125/44145>
- Dikli, S. (2006). An overview of automated scoring of essays. *Journal of Technology, Learning, and Assessment*, 5(1), 1–35. <http://www.jtla.org>.
- Ebyary, K., & Windeatt, S. (2010). The impact of computer-based feedback on students' written work. *International Journal of English Studies*, 10(2), 121–142. <https://doi.org/10.6018/ijes/2010/2/119231>
- Educational Testing Service (ETS). (2019). <http://www.ets.org>
- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Language, and Assessment*, 8(6), 1–43. Retrieved from <http://ejournals.bc.edu/ojs/index.php/jtla/article/download/1625/1469>
- Han, T., & Sari, E. (2022). An investigation on the use of automated feedback in Turkish EFL students' writing classes. *Computer Assisted Language Learning*, 37(4), 961–985. <https://doi.org/10.1080/09588221.2022.2067179>
- Jiang, L., Yu, S., & Wang, C. (2020). Second language writing instructors' feedback practice in response to automated writing evaluation: A sociocultural perspective. *System*, 93, 102302. <https://doi.org/10.1016/j.system.2020.102302>
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by Grammarly: A multiple case study, *Assessing Writing*, Volume 44, 100450, ISSN 1075-2935, <https://doi.org/10.1016/j.asw.2020.100450>.
- Li, Z., Link, S., Ma, H., Yang, H., & Hegelheimer, V. (2014). The role of automated writing evaluation holistic scores in the ESL classroom. *System*, 44, 66-78. <https://doi.org/10.1016/j.system.2014.02.007>

- Liao, H. C. (2015). Using automated writing evaluation to reduce grammar errors in writing. *Elt Journal*, 70(3), 308–319. <https://doi.org/10.1093/elt/ccv058>
- Lim LA, Dawson S, Gašević D, et al. (2021) Students' perceptions of, and emotional responses to personalised learning analytics-based feedback: An exploratory study of four courses. *Assessment & Evaluation in Higher Education* 46(3): 339–359.
- Liu, X., Zhu, C., Jiao, J. & Xu, M. (2018) Promoting English pronunciation via mobile devices-based automatic speech evaluation (ASE) technology. In Cheung, S., Kwok, L., Kubota, K., Lee, L. K. & Tokito, J. (eds), *Blended learning. Enhancing learning success. ICBL 2018: Vol. 10949: Lecture notes in computer science*. Cham: *Springer*, 333–343. https://doi.org/101007/978-3-319-94505-7_27
- McCarthy, K., Roscoe, R., Allen, L., Likens, A., & McNamara, D. (2022). Automated writing evaluation: Does spelling and grammar feedback support high-quality writing and revision? *Assessing Writing*, 52, 100608. <https://doi.org/10.1016/j.asw.2022.100608>
- Mei, B., Qi, W., Huang, X., & Huang, S. (2022). Speeko: An Artificial Intelligence-Assisted Personal Public Speaking Coach. *RELC Journal*, 0(0). <https://doi.org/10.1177/00336882221107955>
- Moxon, S. (2021) Exploring the effects of automated pronunciation evaluation on L2 students in Thailand. *IAFOR Journal of Education: Language Learning in Education*, 9(3): 41–56. <https://doi.org/10.22492/ije.9.3.03>
- Neri, A., Cucchiarini, C., Strik, H., & Boves, L. (2002). The pedagogy-technology interface in computer assisted pronunciation training. *Computer assisted language learning*, 15(5), 441–467.
- Osawa, K. (2023). Integrating Automated Written Corrective Feedback into E-Portfolios for second language Writing: Notion and Notion AI. *RELC Journal*, 0(0). <https://doi.org/10.1177/00336882231198913>
- Pokrivcakova, S. (2019). Preparing teachers for the application of AI-powered technologies in foreign language education. *Journal of Language and Cultural Education*, 7(3), 135–153.
- Portnoff, L., Gustafson, E., Rollinson, J., & Bicknell, K. (2021). Methods for language learning assessment at scale: Duolingo case study. In *International Educational Data Mining Society*.
- Ramineni, C., & Williamson, D. M. (2013). Automated essay scoring: psychometric guidelines and practices. *Assessing Writing*, 18(1), 25–39.
- Rogerson-Revell, P. M. (2021). Computer-assisted pronunciation training (CAPT): Current issues and future directions. *Relc Journal*, 52(1), 189–205.

- O'Brien, M., Derwing, T., Cucchiarini, C., Hardison, D., Mixxdorff, H., Thomosn, R., Strik, H., Levis, J., Munro, M., Foote, J., & Muller Levis, G. (2018). Directions for the future of technology in pronunciation research and teaching. *Journal of Second Language Pronunciation*, 4(2), 182-207.
- Saricaoglu, A., & Bilki, Z. (2021). Voluntary use of automated writing evaluation by content course students. *ReCALL*, 33(3), 265–277. doi:10.1017/S0958344021000021
- Speeko (2022) Speeko. Available: <https://www.speeko.co/> (accessed 1 March 2022).
- Van der Schaaf M, Donkers J, Slof B, et al. (2017) Improving workplace-based assessment and feedback by an E-portfolio enhanced with learning analytics. *Educational Technology Research and Development* 65(2): 359–380.
- Vančová, H. (2021). Teaching English pronunciation using technology. Kirsch-Verlag.
- Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Educ Inf Technol* 28, 13943–13967 <https://doi.org/10.1007/s10639-023-11742-4>
- Yoshida, M. T. (2018). Choosing technology tools to meet pronunciation teaching and learning goals. *Catesol Journal*, 30(1), 195-212.
- Yu-Chih S (2008) The Toastmasters approach: An innovative way to teach public speaking to EFL learners in Taiwan. *RELC Journal* 39(1): 113–130.
- Zhang, Z. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing (2020): Student perceptions and revisions, *Assessing Writing*, Volume 43, 100439, ISSN 1075-2935, <https://doi.org/10.1016/j.asw.2019.100439>.

Exploring the Reflections of Turkish EFL Preparatory School Students towards “English” from an ELF Standpoint

*Pınar Kır¹ Gürkan Temiz²

¹PhD.(c), Istanbul Technical University, Türkiye

²Dr., Istanbul Technical University, Türkiye

Abstract

English as a Lingua Franca (ELF) has been recognized as a significant area of focus in language education and played an important role in shaping language policies all around the world. Although recent trends in ELF have led to a proliferation of studies searching for English language teachers’ awareness and practices in the field, there is a scarcity of research that examines student perspectives on learning English as a foreign language (EFL). Therefore, it is also necessary to examine EFL students’ perceptions of learning English from an ELF perspective to better associate the concept with actual teaching and learning practices in the context of changing and developing educational practices. In this regard, the present study aims to gain insights into EFL students’ perceptions of the English language in terms of its perceived reasons, varieties, and educational practices from an ELF perspective. 21 students studying at a language preparatory school of a foundation university volunteered for the study. The data were collected by conducting interviews and were examined through thematic analysis in which recurring themes were identified and then categorized. The reasons for learning English, conceptualizations of English, and reflections on English language education were found to be the main themes with subcategories. In this session, the results of the study will be shared and discussed in line with the ELF standpoint and recommendations for future implementations will be provided.

Keywords: English as a Lingua Franca, ELF-awareness, EFL learners

1. Introduction

English is the dominant foreign language spoken all around the world (Graddol, 2006). This dominant language has been taught and learned based on standard norms of American and British English and other varieties have been disregarded most of the time. This traditional perspective dominated English language education for a long time. However, thanks to the globalized world, new perspectives have risen in English language education as well as in other fields.

Globalization has been described as "the creation of world relations based on the operation of free markets" (Foskett, 2010, p.1) and the development in the field of technology and economic growth have given rise to the emergence of globalization all over the world. In that sense, it influenced the globalization of English, too. It created a paradigm shift in English for its non-native users. Language policies and thus language education have been affected by this change. English as a Lingua Franca (ELF) emerged as a result of this paradigm shift in the field. Although ELF was not a part of ELT at the very beginning, it resulted in creating classroom implications in time (Sifakis, 2017). ELF, defined as the discourse emerging from the communicative interactions of non-native speakers of English (Sifakis, 2017), shifted the principles of ELT from the native English norms towards a more multilingual perspective. Thus, ELF has created a new multilingual perspective that is separated from native English norms (Seidlhofer, 2010), and different varieties of English have gained importance in ELF research. Regarding the ELF phenomenon, researchers started to explore teachers' ELF awareness (Öztürk, Çeçen & Altınmakas, 2010; Deniz, Özkan & Bayyurt, 2016; Temiz, 2018; Vettorel, 2017). However, there has been little investigation of students' perceptions towards English and the place of its varieties in foreign language education (Jenkins, 2007; Tsou & Chen, 2014). For this reason, the present study aimed to examine Turkish EFL learners' perceptions of English and its education in terms of their motives to learn it, their conceptualizations of it, and their reflection on the English language education in the Turkish EFL context.

2. Literature Review

2.1 Definition of English as Lingua Franca (ELF) and ELF Awareness

As a requirement of globalization, the number of people who speak English has increased tremendously. In the past, although it was only the native speakers of English, who were represented in the Inner Circle (e.g., England, America, etc.) of Kachru's Model (1985), it became the second official language of many colonized countries represented as Outer Circle countries in the model (e.g., South Africa, Philippines, etc.). Progressively, globalization gave rise to the Expanding Circle as well and the number of countries where English is spoken as a foreign language has risen significantly (e.g., Türkiye, European Countries, etc.) and outnumbered the other two circles. Thus, it has been indicated that English cannot be restricted to these three circles (Jenkins, 2009) anymore as it has gone beyond these circles in terms of its function as a lingua franca.

Sifakis (2017) defined ELF as the discourse emerging from the communication between non-native speakers of English and indicated that ELF is the context nourished in a multilingual and multicultural environment independently from native English norms. Furthermore, Jenkins

(2009) explained ELF as a context where English is the most spoken language, namely lingua franca for interaction among people from multilingual and multicultural origins. To that wake, ELF has been commonly used in the Expanding Circle and the number of non-native speakers of English is more than native speakers of English (Crystal, 2003; Graddol, 2006). Nonetheless, Jenkins (2009) did not separate the Inner and Outer Circle from the ELF context indicating that in all circles speakers should adjust their language to the language and culture of the addressee and emphasizing the unifying role of ELF on indigenized varieties as a common language. What is more, Kalocsai (2014) pointed out that ELF highlights the context of communication among speakers from different backgrounds and their mother tongues as they affect their English use (Jenkins, 2009) in the form of code-switching, code-mixing, repetition, and echoing of language items (Cogo & Dewey, 2006; Kirkpatrick, 2008) rather than English itself.

According to Seidlhofer (2012), ELF speakers are “(...) speakers from all walks of life in all continents (...)” (p.85) and can communicate across boundaries by adapting to all kinds of communication strategies. In that regard, Sifakis (2017) specified the role of ELF in creating English speakers who can accommodate themselves to their interlocutors and conduct a successful communication process, whereas EFL where the language teaching policies are developed based on standard native English norms dictates native-like norms without allowing any creative language usage and localization.

In light of the aforementioned, ELF has become the central issue to embrace all facets that it gave rise to. Sifakis and Bayyurt (2017, p. 459) defined ELF as:

"The process of engaging with ELF research and developing one's own understanding of the ways in which it can be integrated in one's classroom context, through a continuous process of critical reflection, design, implementation, and evaluation of instructional activities that reflect and localize one's interpretation of the ELF construct."

Regarding this, awareness has been accepted as one of the primary concerns of ELF studies. Based on the aforementioned definition, Sifakis (2017) classified ELF awareness into three broad types: a) awareness of language and language use, b) awareness of instructional practices, and c) awareness of learning. *Awareness of language and language use* refers to the knowledge of semantics, syntax, morphology, phonology, discourse, and pragmatics of the language used in non-native English spoken context (Sifakis, 2017). That kind of awareness includes the ability to fix the communication breakdowns with communication strategies (Swain, 2006) and adjusting the communication process regarding the interlocutor.

With respect to *awareness of instructional practice*, Graddol (1997) mentioned that ELT should be modified in the light of recent changes in the global status of English and ELF should influence the instructional practices. Awareness of instructional practice is more relevant to teachers' understanding and practices of ELF. It examines the teachers' role in a teaching environment from the ELF context as they are the initiator of ELF. As proof of teachers' role in ELF awareness, Qiufang (2012) denoted China's EFL status from ELF perspective and stated that nearly all teachers try to enable students to reach native-like accuracy and impose standard English norms without considering ELF perspective. He also suggested that China's situation can be generalized to other expanding circle countries where most teachers are not aware of ELF as all materials, curriculum and books are Standard English oriented. Nonetheless, it is argued that to be able to generate awareness among students, the teacher should teach students communication skills that involve an ELF standpoint. To reach this kind of awareness level,

Qiufang (2012) proposed his views about teaching in terms of three components: linguistic, cultural and pragmatic. For linguistic component; native and non-native varieties, localized features; for the cultural component; target language culture, non-native culture, and learners' culture; and for the pragmatic component; universal common rules and target language common rules should be taught.

Additionally, *awareness of learning* holds the overall effect of ELF on the learning process. It involves all concepts related to language learning such as textbooks, teachers, etc. In that sense, nativism is accepted as one of the most contradictory topics since ELF created a dichotomy against traditional nativist approach that refers to native speakers as biologically superior while non-native speakers are addressed as language deficits. Jenkins (2006) clarified that this distinction between native vs. non-native speakers is the root of EFL vs. ELF distinction due to their contradicting tenets. The same discussion emerged in the ELT field in the form of native English-speaking teacher (NEST) vs non-native English-speaking teacher (NNEST). Many studies were conducted to research students' perspectives on NEST vs NNEST (Cheung, 2002; Devrim & Bayyurt, 2010) and the results indicated that students preferred NEST.

2.2 Studies on English as a Lingua Franca Awareness

Although a considerable amount of literature has been published on ELF awareness, most of the studies focused on ELF awareness among teachers (Öztürk, Çeçen & Altınmakas, 2010; Deniz, Özkan & Bayyurt, 2016; Temiz, 2018; Vettorel, 2017; Vettorel & Corrizato, 2016). To illustrate, Öztürk, Çeçen and Altınmakas (2010) searched pre-service language teachers' ELF awareness and reported that participants accept native speakers as the owners of English. Additionally, Deniz, Özkan and Bayyurt (2016) did the same study with 42 pre-service teachers and their findings supported the previous literature by reporting that pre-service teachers perceive native-like accuracy as the core of English language teaching (ELT). Likewise, Temiz (2018) conducted a case study by examining the perceptions and practices of in-service teachers who went through an ELF-aware teacher education program (ELFATE) designed by Sifakis and Bayyurt (2015). The findings found that the participant EFL teachers who attended ELFATE program developed their understandings of English and English language teaching about sensitive issues such as nativism, ownership of English, and intelligibility. Besides, the participants integrated the ELF perspective in their classroom practices by embracing L1 use and using authentic materials more.

On the other hand, only a few researchers investigated ELF awareness among students (Jenkins, 2007; Tsou & Chen, 2014). In Jenkins (2007), she investigated non-native English speakers' perceptions towards English accents with a questionnaire including different accents of English. Her results indicated that participants desire to reach a standard native English speaker level although they don't learn the language to interact with the people in the inner circle. They opted for American and British accents as they accept those varieties suitable for international interaction. She attached these views to the ideology that teachers had been raised and the course materials that they had been using aligned with traditional standard norms. Despite the tendencies among the participants towards standard English norms, she also found that the participants had a positive attitude towards the concept of ELF.

Another study by Tsou and Chen (2014) researched Taiwanese EFL students' perceptions of English in a comparison study. They compared EFL Taiwanese students with international

MBA Program students via a questionnaire and interviews. Their results showed both differences and similarities in their conceptions of ELF awareness. In the light of the questionnaire, most EFL students were aware of different varieties of English, but they reported that American and British English should be the Standard English in language education. Further, the majority of the participants stated intelligibility has higher importance than grammatical accuracy. In line with this, the interview responses affirmed that students learn English to have a good career path, and they thought the accent is not crucial as long as they can communicate easily. In this vein, non-native speakers were found more intelligible than their native counterparts as their English includes less complex linguistic structures.

Although there have been different studies about ELF awareness from various English language teaching contexts, the majority of the studies have investigated teachers' ELF awareness (Sifakis & Sougari, 2005; Temiz, 2018; Vettorel & Corrizato, 2016). Nonetheless, there is a dearth of research investigating EFL students' perceptions of ELF awareness. Therefore, it is necessary to examine the perceptions of students in English language education to better understand the concept of ELF-awareness and its applicability in English language education in Türkiye. To fill this gap in the literature, the present study tries to investigate EFL students' drives for learning English, conceptualizations of English, and their reflections on English language education in the Turkish context.

With this aim in hand, the current research tried to answer the following research questions:

1. What are the motives of EFL university students to learn English?
2. How do EFL university students conceptualize English?
3. What are the reflections of EFL university students on English language education in Türkiye?

3. Method

3.1 Participants and Research Setting

The research was conducted in one English language preparatory program of a foundation university in Istanbul, Türkiye. Türkiye is one of the expanding circle countries where English is taught as a foreign language (EFL). Students are not commonly exposed to English outside the class. English is taught with the traditional Standard English teaching methods. Coursebooks and teaching materials are not ELF-oriented, and teachers are expected to follow standard English norms. The university from which the data came offers programs where the medium of instruction is English, therefore English preparatory school is mandatory for all students. They have to complete one year of English language education and pass the proficiency exam at the end of the year. 21 learners from the English language preparatory school of a foundation university from Türkiye volunteered in this study. The participants were selected through random sampling among the volunteered students from upper intermediate classes. All of them were native speakers of Turkish and had been learning English for ten years. Their ages ranged between 18 to 20.

3.2 Interviews

20-minute interview sessions were conducted with each participant. Interview questions were devised regarding the book by Jenkins (2009) and it included 24 open-ended questions aiming

to detect students' ELF awareness. The participants were interviewed about their impressions of SE norms and varieties of English as well as their reflections on their English language education. The interviews were carried out in Turkish to make students feel comfortable and prevent any language barriers.

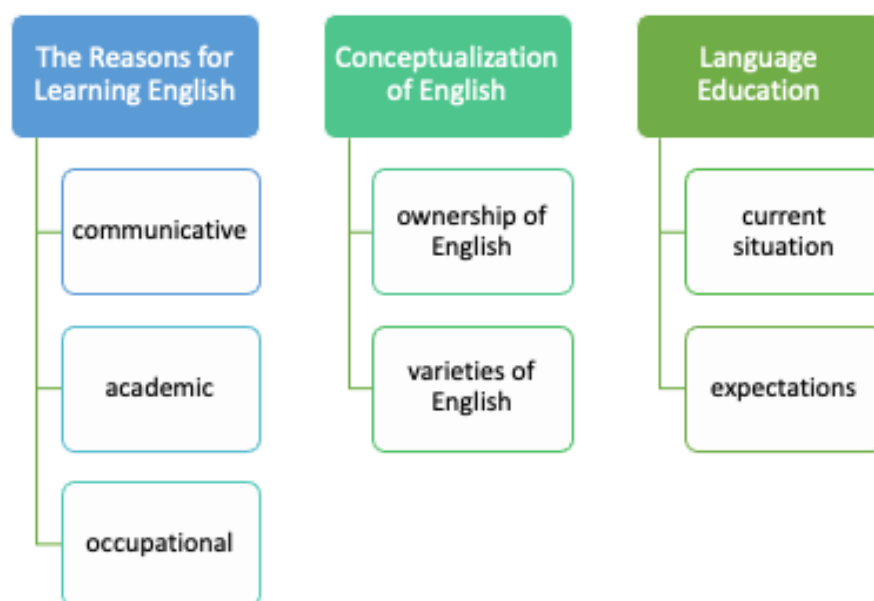
3.3 Data Analysis

The interview data were recorded and verbatim transcribed. The written data were analyzed and categorized according to recurring themes and patterns, which is a component of thematic analysis (Boyatzis, 1998) to detect commonalities in the data. The codes derived from the data were checked for inter-rater reliability.

4. Findings

Regarding the participants' responses, three main themes were revealed: 1) the reasons for learning English, 2) the conceptualization of English, and 3) language education. In terms of the first theme, a) communicative, b) academic, and c) occupational; in terms of the second theme, a) ownership of English and b) varieties of English; and in terms of the third theme, a) current situation and b) expectations related to system and teachers were found to be subcategories. Figure 1 illustrates the themes and subcategories.

Figure 1. Themes and subcategories derived from the participants' reflections



4.1 Unveiling the Reasons for Learning English

Pertaining to the first research question investigating the motives of EFL university students to learn English, the interview results showed the reasons for learning English as a main theme with three subcategories: a) communicative, b) academic, and c) occupational. About the first subcategory, communicative, the majority of the statements reported by the participants (61%) identified English as the common world language which makes international life easier for

communicative purposes. Some interviewees referred to English as a tool to communicate with people from different cultures and countries while some others see it as a way of living abroad.

“English is a gatekeeper to reach other cultures.” (P5)

“English means being able to communicate with foreigners in English and being able to step out of one's fish glass.” (P8)

Also, they assess English as a means of socialization and a *golden bracelet* that opens all the doors for its speakers. According to them, there are currently many speakers of it as well as many more who want to learn it in the future. They believe that English will be more common in the near future as it is essential and easier to learn as a foreign language to interact with people from diverse backgrounds perceiving it as the language of economy, politics, sports, and medicine.

“English means books, music, films, and Tv series for me. If I learn English well, I can understand all of them easily and I will gain overall knowledge about them.” (P12)

“If the world was a human body, English would be its heart.” (P4)

Regarding communication in terms of using English in real life, most of the participants (60%) indicated that they learn English to be able to communicate with both native speakers and non-native speakers while 40% of them stated a preference for either native or non-native.

“Native speakers do not give too much attention to grammar mistakes, and they understand what I mean from just a word.” (P15)

Concerning the addressee of communication, 13 (60%) of them stated they were feeling more comfortable talking to native speakers. As native speakers do not know Turkish, it triggers students to think in English and become more relaxed. Also, they believe that native speakers do not pay attention to mistakes and correct them because they are aware that English is not their mother tongue. The rest of the participants (40%) wish to communicate with non-native speakers since they can speak Turkish and understand what students mean easily.

“While I am talking to native speakers, I feel nervous and anxious as they are good at language, but with non-natives, I feel better.” (P3)

Touching upon the second subcategory, academic reasons, 22 % of the statements provided by the participants indicate that the students learn English mostly to be successful in academia. While English is a prerequisite for them to be able to start their departments for their bachelor's degrees, it also posits utmost importance in being selected for graduate studies.

“I am learning English because it will help me in academia.” (P17)

“It is important for my future and my field of study, especially for my master's degree.” (P11)

Additionally, some students stated that they learn English to be more knowledgeable about their field by reading international sources in English because nearly all of the books and other materials are available in English albeit less in Turkish.

“I'm learning English in order to start my department, or sometimes not learning. However, by looking into it and researching on my own, learning English out of

curiosity becomes more beneficial for me. Being able to read the books I want to read in their original language, understanding most of the music I listen to, accessing articles and scientific writings—all of these contribute to that.” (P1)

Regarding the last subcategory, occupational, some statements (22%) clarified that the participants desire to have better job opportunities and English is the primary criterion to reach this goal. Therefore, they state that:

“Thanks to English, I will be one step ahead in my future career and profession.” (P5)

“If you know it, it has advantages for helping you get a job, but if you don't know it, it is quite disadvantageous.” (P9)

Yet, some other participants perceive English as an obligation to have good job opportunities. They complained that being able to speak English should be accepted as a criterion to be perceived as a successful candidate. For this reason, they are not pleased with learning English.

“I learn English out of necessity, so it will be useful for my job when I enter the workforce.” (P20)

4.2 Conceptualizations of English

Related to the second research question, the theme of conceptualization of English was revealed in the data set. Under this main theme, it was found that the participants conceptualized English in terms of two themes: a) ownership of English and b) varieties of English.

Firstly, participants’ responses provided patterns for *ownership of English*. The participants reported 73 statements and out of these statements most of the participants indicated that British people (42%) are the owners of English and all of the participants believed that English is the mother tongue of American and British people. Subsequently, America (38%) was found out as the second owner of English. Following it, 16% of the statements indicated that the participants perceive English as a global language stating that it belongs to neither the UK nor the USA. Additionally, Canada (1%), New Zealand (1%), the Philippines (1%) and Australia (1%) were mentioned in the data albeit less in the percentages compared to the UK and the USA.

In terms of *varieties of English*, nearly all of them (98%) agreed upon the existence of varieties of English indicating some grammar and accent differences. Only one participant disagreed with the availability of varieties of English.

“It is a language with differences, even the pronunciation of words varies greatly. I believe not everyone learns the same English. Especially, series are an indicator of this.” (P14)

“It is a language with differences, each country has different accents. I think English is different in all countries.” (P18)

When students were asked about the varieties of English they know, they could only name British (46%) and American (41%) English. Also, Australian (12%) and Canadian (3%) Englishes existed in the answers.

"I know American and British, yet, I know, but there's more than these." (P6)

"English may vary by location." (P10)

4.3 Reflections on English Language Education

With regard to the third main theme, language education, participants reflected on patterns for the current situation in Türkiye. Related to the current situation in the system of English language education in Türkiye, a common view amongst interviewees was that preparatory school English education system is not enough to develop their listening and speaking skills as only grammar is emphasized in classes. They evaluated their system as an unsuccessful exam-based system that cannot provide the required skills for them. Students did not approve of their current system based on traditional grammar teaching methods and want to have more speaking, writing and listening activities included in the program. They were aware that learning a language requires a lot of struggle but the school offers language education as a prerequisite for the department lessons instead of teaching it in an effective way. On the other hand, some of the participants stated language education in their school was efficient albeit less in number. However, they did not elaborate on their argument.

"I am not satisfied with the language education at school because we learn only cliché structures with a grammar favored theory. Since they are very careful about grammar mistakes, we never have a fluent speaking skill." (P11)

The findings also indicated whether the system imposes native like norms or not. Students implied there was not any pressure from neither the teachers nor the school system. However, there was a general sense of pressure from the school system towards grammar use, though they might be moderate for pronunciation mistakes. Participants indicated that teachers were too strict in terms of grammar rules, but the emphasis on speaking and daily usages were less. On the other hand, some interviewees said it depends on the teacher. While some of the teachers are more welcoming towards pronunciation or transfer mistakes, some teachers can be more demanding.

"Our teachers have never imposed Standard English norms on us or warned us. In contrast, when we made a mistake or our pronunciation was wrong, they encouraged us to speak. They advise us to practice despite mistakes." (P21)

"The Turkish education system is established on memorizing rules. They don't put emphasis on students' speaking abilities. They expect us to speak fluently and only teach grammar. However, that expectation is not to make students reach native-like fluency." (P3)

As regards the second subcategory, the expectations, the participants all agreed on a system focusing on speaking, writing and listening skills. In terms of the standards of language education, while the majority of participants reject the Standard English norms, only a few of them support Standard English norms. In addition, some students said English should be taught like a mother tongue. While the majority of participants approve teaching different cultures and varieties of English in the class, only a few students reject them.

"Learning English does not mean to be an American or British. The important thing is to communicate with people from different cultures and learn about their cultures. In that sense, language becomes a mediator. Non-native speakers of English cannot have

a native-like accent or acquire all the rules since we learn the language as an additional language to our mother tongues to combine them. Necessarily we acquire a different accent assimilating English with our own identity." (P18)

"Standard English norms should be accepted in language education as there are different varieties. For example, it is really hard to understand a person who speaks with South African English. It would be really hard to understand all the differences. Language is the way of communication, so we should not diversify its uses so as not to make it harder." (P13)

Additionally, all participants agreed on the requirement of teaching different cultures and varieties of English in the lessons to catch the attention of people. About the norms of language teaching, they stated Standard English norms should not be dictated and native-like accuracy should not be important.

"They have to be taught since learning a language means learning different cultures." (P7)

"While watching the movies, I encounter lots of abbreviations and ungrammatical structures. We do not need to know every grammar rule, it is important to understand the conversation." (P19)

"Teachers should not insist on native-like accuracy, it does not improve us, in contrast, it makes our learning process harder." (P12)

In reference to teacher-related expectations, students also shared their insights. For teachers of English, all of the participants feel relaxed and confident with native teachers. They believe that native teachers are better at teaching as they know the language very well. While 70 % of participants prefer native teachers as an ideal English teacher, 30 % of them prefer non-native teachers since they understand students.

"During native teachers' lessons, I feel more self-confident and less stressed." (P8)

"I prefer non-native teachers because they understand us better as they had the same difficulties during their language learning process." (P10)

Most of the participants trust native teachers more than non-native teachers since native teachers have a good command of English and give various examples. In contrast, only a small number of participants trust non-native speakers' grammatical knowledge indicating native teachers lack knowledge of grammar.

"Native teachers are more proficient with the language than non-native teachers. They don't memorize so they can give different examples." (P16)

In terms of the effect of teaching, participants hold different views. Some of them believe the positive effect of native teachers, the others emphasize the benefits of having a non-native teacher in the class. Also, a minority of participants mentioned that their preferences change depending on the teacher.

5. Discussion

The current research has made notable contributions to the literature on English as a lingua franca by unveiling three significant findings. Firstly, it found out the motives of EFL university students to learn English. Secondly, it revealed how EFL preparatory school students conceptualize English. Finally, it shed light on the reflections of EFL university students on English language education in Türkiye.

In terms of the reasons for learning English including communicative, academic, and occupational purposes, the findings of this study were consistent with those of other studies (Jenkins, 2007; Tsou & Chen, 2014) suggesting that students learn English to be able to communicate with foreigners and to have better career opportunities.

Regarding the ownership of English, the findings were consistent with Öztürk et al. (2010) which states that native speakers are the owners of the English in this study. In reference to the varieties of English, the findings supported previous research claiming students are aware of varieties of English (Tsou & Chen, 2014). Also, this study demonstrated that Turkish EFL students accept American and British varieties as the standard norms of English corresponding to Tsou and Chen (2014).

The results of the study indicated insights of the students towards language education in Türkiye. Although the system held an exam-based approach with grammar teaching orientation for language education in Türkiye, the students opted for a more literacy-oriented language education akin to the studies in literature (Tsou & Chen, 2014). Their overall expectancy was to be taught communicative skills.

What is more, this study was unable to demonstrate that the language system in Türkiye emphasizes Standard English norms and native-like accuracy in contrast to Qiufang (2012) claiming that the emphasis for native-like accuracy in China can be generalized to other Expanding Circle countries.

Additionally, students in this study thought non-native speakers are more intelligible than their native counterparts because their English includes less complex linguistic items. This finding merges with the results of Tsou & Chen (2014). As for native-like accuracy, students' findings showed contrasting results with previous literature (Deniz, et al. 2016; Jenkins, 2007). Students do not favor Standard English norms and native-like accuracy as their ultimate language learning aim. This result may be explained by the fact that students have intercultural awareness and are aware of the role of global use of English though they do not specifically refer to ELF terminology.

6. Conclusion and Implications

The present study set out to determine Turkish EFL students' perceptions of English from an ELF standpoint. The evidence from this study suggests that Turkish EFL students do not have ELF awareness in terms of its terminology. Their awareness of SE norms and the current practices in English language education prevent them from becoming ELF-aware speakers of English. Therefore, language education in Türkiye should be devised with a more ELF-aware

teaching perspective to be able to prepare students for the global use of English in all fields supporting them with ELF-aware teachers, materials, and tests.

Finally, several limitations need to be considered. First, the study has only examined a university context in Türkiye, so the results may not apply to other contexts. Although its main aim is not to generalize the results, the number of participants can be increased in different university contexts. Also, the study is limited to qualitative data. Further studies can be triangulated with quantitative methods to collect more data on EFL students' perspectives on English and ELF-awareness.

References

- Boyatzis, R. E. (1998). Transforming qualitative information: Thematic analysis and code development. Thousand Oaks, CA, US: Sage Publications, Inc.
- Chen, F., & Tsou, W. (2014). EFL and ELF college students' perceptions towards English. *De Gruyter Mouton*, 3(2), 363-386.
- Cogo, A., & Dewey, M. (2012). Analysing English as a lingua franca. London: Continuum.
- Crystal, D. (2003). English as a global language (2nd ed.). Cambridge: CUP.
- Graddol, D. (2009). English Next: Why Global English May Mean the End of English as a Foreign Language. London: British Council.
- Jenkins, J. (2006). The spread of EIL: A testing time for testers. *ELT Journal*, 60(1), 42-50.
- Jenkins, J. (2009). World Englishes: A resource book for students (2nd ed.). London, England: Routledge.
- Kachru, B. B. (1983). The Indianization of English: the English Language in India. Delhi: Oxford University Press.
- Kalocsai, K. (2014). Communities of Practice and English as a Lingua Franca. Berlin: De Gruyter Mouton.
- Kirkpatrick, A. (2008). English as the official working language of the Association of the Southeast Asian Nations (ASEAN): Features and strategies. *English Today*, 94, 27-34.
- Öztürk, H., Çeçen, S., & Altınmakas, D. (2010). How do non-native pre-service English language teachers perceive ELF? A qualitative study. *The English as an International Language Journal*, 5, 137-146.
- Qiufang, W. (2012). English as a lingua franca: A pedagogical perspective. *Journal of English as a Lingua Franca*, 1(2), 371-376.
- Seidlhofer, B. (2010). 'Lingua franca English: the European context.' In Kirkpatrick, A. (ed.), *The Routledge Handbook of World Englishes*. London: Routledge, pp 355-371.
- Sifakis, N. C., & Sougari, A. (2005). Pronunciation issues and EIL pedagogy in the periphery: A survey of Greek state school teachers' beliefs. *TESOL Quarterly*, 39(3), 467- 488.
- Sifakis, N. C., & Bayyurt, Y. (2015). Insights from ELF and WE in teacher training in Greece and Turkey. *World Englishes*, 34(3), 471-484.
- Sifakis, N. C. (2017). ELF Awareness in English language teaching: Principles and processes. *Applied Linguistics*, 1(20), 2-23.

- Sifakis, N. C., & Bayyurt, Y. (2017). 'ELF-aware teacher education and development.' In Jenkins, J. Baker, W. and Dewey, M. (eds.), *The Routledge Handbook on English as a Lingua Franca*. London: Routledge, pp. 456-467.
- Swain, M. (2006). 'Languaging, agency and collaboration in advanced second language learning' in Byrnes, H. (ed.), *Advanced Language Learning: The Contributions of Halliday and Vygotsky*. London: Continuum, pp. 95-108.
- Temiz, G. (2018). Exploring the relationship between ELF awareness and teacher cognition in ELF (Unpublished master's thesis). Boğaziçi University.
- Vettorel, P. (2017). The plurality of English and ELF in teacher education. Raising awareness of the 'feasibility' of a WE-and ELF-aware approach in classroom practices. *Lingue e Linguaggi*, 24, 239-257.
- Vettorel, P., & Corrizato, S. (2016). Fostering awareness of the pedagogical implications of World Englishes and ELF in teacher education in Italy. *Studies in Second Language Learning and Teaching*, 6(3), 487-511.

The Effects of Explicit Teaching of Speech Acts on EFL Learners' Pragmatic Competence

Sinem Bayındır Musaoğlu¹ Emrah Görgülü²

¹Inst., Istanbul Technical University, Türkiye

²Assoc. Prof., Istanbul Sabahattin Zaim University, Türkiye

Abstract

This study investigated the efficacy of teaching pragmatic knowledge of speech acts through explicit instruction to English as a foreign language (EFL) learners within the context of a preparatory school affiliated with a foundation university. In this study, experimental group design was employed, and a pre and post-test with an explicit intervention to the experimental group were followed. The research particularly examined four different speech acts, namely greeting, advice giving, complaint and refusal, which are considered to have notable impact on pragmatic competence of EFL learners in different social settings. The data were collected through a Written Discourse Completion Task to evaluate the communicative abilities of both the control group and experimental group. The results were analyzed through the mixed ANOVA on SPSS v.22 (Statistical Package for Social Sciences 22.0). The findings from the pre and post-tests revealed a significant enhancement in the performance of the experimental group on target speech acts. The outcomes of descriptive analysis of pre and post-test scores of each group also discerned that the group exposed to explicit instruction performed substantially better compared to the control group. Furthermore, the mixed ANOVA results indicated that the type of the speech act has a decisive effect on the efficacy of explicit instruction. The findings of the present study encourage the use of explicit pragmatic instruction in an EFL classroom setting to develop firstly pragmatic awareness, and then pragmatic competence to perform appropriate speech acts in distinct social settings. This empirical study also paves the way for more effective language learning practices and enhanced pragmatic proficiency in EFL contexts.

Keywords: Pragmatics, interlanguage pragmatics, English as a foreign language, speech acts, explicit instruction

1. Introduction

In non-English-speaking countries, as opposed to native or native-like speakers, English as a foreign language (henceforth, EFL) learners are exposed to linguistic components of a foreign language in educational settings. However, EFL learners must also improve their pragmatic skills to comprehend and convey messages for social interactions in appropriate ways in the target language. Pragmatics deals with this need and focuses on “how the interpretation and use of utterances depend on knowledge of the real world, next how speakers use and understand speech acts, and then how the structure of sentences is influenced by the relationship between the speaker and the hearer” (Tuncay, 2013, p. 155). The field of interlanguage pragmatics (henceforth, ILP) focuses on various speech act realization patterns by analyzing native and non-native speakers’ data and studies the acquisition, comprehension, and execution of pragmatics by non-native speakers (Kasper & Schmidt, 1996). EFL learners, even with advanced EFL proficiency, often fail to produce appropriate utterances in diverse social contexts. As Bardovi-Harlig and Griffin (2005) highlight, even language learners with high linguistic competence might have major pragmatic failures. To interpret the intended messages, learners need to have enough pragmatic awareness of the target language, which is often observed as a major downside of EFL learners. Hence, more extensive research on pragmatic awareness is needed in ILP studies to help EFL learners as there are not sufficient studies conducted to date. This study aims to contribute to the field with that purpose.

This research investigates the impact of explicit instruction on teaching four target English speech acts within a classroom setting. When this study aims to contribute to interlanguage pragmatics, its main objective is to show the effectiveness of explicit instruction in improving pragmatic competence in a classroom setting as opposed to staying for a time in the target speech community. The present study investigates the effect of explicit instruction on EFL learners’ pragmatic competence by studying four speech act patterns, namely, greetings, advice-giving, complaints, and refusals. As the current study addresses the gap in the study of speech acts such as greetings, advice-giving, complaints, and refusals, it serves a valuable purpose. While there are many significant studies regarding the speech act of refusal in the previous studies, the target speech acts within this study have less emphasis to date. Therefore, these speech acts were decided to be investigated in this study to fulfill the insufficiency of the research on those frequent pragmatic failures exhibited by advanced EFL learners.

Here are the primary research questions at the core of the present study.

1. Does explicit teaching of speech acts in a classroom setting increase the pragmatic competence of EFL learners?
2. To what degree does the accepted use of speech acts improve towards a native-like degree through explicit teaching?
3. Does the efficiency of explicit teaching of pragmatics change according to different speech acts?

2. Literature Review

The primary objective of this section is to establish a theoretical background in pragmatics and interlanguage pragmatics and provide an overview of the concepts of speech acts with key theories and significant researchers such as Crystal (1985), Morris (1938), and Kasper & Blum Kulka (1993). The core definitions of pragmatics and interlanguage pragmatics will be presented first to distinguish these two fields and draw attention to the present study's focus. A brief discussion on the teachability of pragmatics through explicit instruction will also emphasize the focus of the study. Finally, the effect of explicit instruction in a classroom setting on target speech acts will be discussed.

Various definitions provide a framework for the current study. An early definition of pragmatics which is "the study of the relation of signs to interpreters" dates back to 1938. (Morris, 1938, cited in Schauer, 2009, p. 5). Crystal (1985) defined pragmatics as the study of language focusing on the choices made by users, their constraints faced in social interaction, and the impact of their language use on other interlocutors. Then, Mey (2001) offered a new definition that considers pragmatics as the study of language use in human communication controlled by social conditions. While former definitions deal pragmatics from the perspective of language users, the role of social conditions in language use was highlighted in May's definition. Each researcher's focus while describing pragmatics emphasizes particular distinctions within the field. This study can be examined within interlanguage pragmatics. ILP is defined by Kasper & Blum Kulka (1993, p. 3) as "the study of nonnative speakers' use and acquisition of linguistic action patterns in a second language (L2)." Interlanguage pragmatics (ILP) has been attracting more attention recently, enabling numerous empirical studies in second language research. According to Bardovi-Harlig & Hartford (2005), interlanguage pragmatics discusses the realization of speech act types, communicative success, communication breakdowns, input effect, and conversational strategies of L2 learners. The effect of instructional treatments, whether explicit or implicit, on teaching pragmatics, has garnered more attention among ILP researchers lately. The number of studies examining instructional effects has increased dramatically in recent years, however, there remains a notable need to unveil significant aspects of instructional effects in ILP research. Thus, this study probes the success of explicit instruction effect in developing the pragmatic competence of EFL learners during their performance of speech acts.

In ILP studies, there are diverse data collection methods applied, namely, role plays, DCTs (e.g., Written Discourse Completion Task (WDCT), Oral Discourse Completion Task (ODCT), and Multiple-Choice Discourse Completion Task (MDCT)), and authentic dialogs. The current study employed a WDCT including twelve speech events to collect data from experimental and control groups. There are also various research areas in interlanguage pragmatics. Bardovi Harlig (1999) prepared a table illustrating common research areas in SLA related to ILP with some basic research questions. The research questions of the current study are derived from question (h) in Table 2.1.

Table 1: Basic Questions About SLA with Respect to Interlanguage Pragmatics

a.	Are there universals of pragmatic and do they play a role in interlanguage pragmatics?
b.	How can approximation to target language norms be measured?
c.	Does L1 influence L2 pragmatics? (Transfer)
d.	Is the development of L2 pragmatics similar to learning a first language?
e.	Do children enjoy an advantage over adults in learning a second language?
f.	Is there a natural route of development as evidenced by difficulty, accuracy, or acquisition orders or discrete stages of development?
g.	Does type of input make a difference? (foreign language vs. second language)
h.	Does instruction play a role?
i.	Do motivation and attitudes influence level of acquisition?
j.	Does personality play a role?
k.	Does a learner's gender play a role?
l.	Does (must) perception or comprehension precede production in acquisition?
m.	Does chunk learning (formulaic speech) play a role in acquisition?
n.	What mechanisms drive development from stage to stage?

Source: Bardovi-Harlig, 1999: 682

Table 1 demonstrates critical questions regarding different aspects of SLA related to interlanguage pragmatics. The question (h) specifically addresses the present study because it attempts to explore the instructional effect on speech acts. Hence, this study first aims to uncover whether instruction affects teaching and learning of speech acts. Then, it further inquiries about the effectiveness of explicit instruction and the degree of effectiveness concerning different speech acts targeted in the study. Speech acts are highly important for accurate communication as they comprise real-life interactions. It is difficult to explain speech acts with a few theories, yet they are essential to reaching a deep understanding of the nature of the language used by learners for effective communication. Björgvinsson (2011) indicates that speech acts modify reality governed by the power of words; otherwise, language would only define the utterances regarding truth and falsity without speech acts. In this sense, speech act theory fosters the view of language not only as a communication tool but also as an instrument of action.

Ishihara & Cohen (2010) indicate that speech acts are teachable and learnable aspects of L2 communication and this prepares the ground for empirical research. There have been various studies regarding the instructional effect on the performance of L2 speech acts, including greetings, requests, apologies, complaints, refusals, advice-giving, invitations, compliments, and conversation closings to date. It is essential to discuss four target speech acts in this study and compare them with pioneering research in the field. The instructional effect on L2 speech act performance of EFL learners has been extensively studied in interlanguage pragmatics. The supporters of instructional intervention are divided into two groups with regard to the type of instruction that enhances the pragmatic development of speech acts. There are several studies advocating the positive effect of explicit instruction (Bacelar da Silva, 2003; King & Silver, 1993; Kondo, 2008; Morrow, 1995), asserting that explicit intervention has evident outcomes for teaching and learning speech acts. It is often considered superior to implicit intervention concerning learning pragmatic aspects of speech acts. It is often considered superior to implicit intervention concerning learning pragmatic aspects of speech acts. However, the effectiveness of explicit instruction has not been approved by some researchers supporting the merits of implicit intervention instead (Martinez-Flor & Fukuya, 2005; Ülbeği, 2009). Takahashi (2010) advocates that some types of implicit instruction show considerable effect at the pragmalinguistic level compared to explicit instruction. Koike & Pearson (2005) claim that

implicit pre-instruction and feedback have the most impactful influences regarding instructional treatments for the pragmatic performance of L2 learners based on their open-ended test results.

The present study encourages that EFL learners have substantial challenges when they acquire pragmatics in EFL settings. Hence, the intervention of explicit instruction is required during a limited time to teach the pragmatics of a foreign language to EFL learners in an educational setting. The results of the study support that there are considerable positive effects of explicit instructional intervention on EFL learners' pragmatic development within a limited time.

3. Study and Methodology

The present study aims to address the pragmatic gap that L2 learners have when they are learning English. The study attempts to demonstrate how beneficial explicit instruction on teaching four target speech acts for advanced EFL learners.

3.1 Research Design

In this study, an experimental design with a control and an experimental group including *seventeen* students aged 18-22 was employed. As a research design, a pretest-posttest design was used. The researcher was also an instructor at the same foundation university; thus, researcher's class was selected as the control group in the study and another advanced class was randomly chosen as the experimental group. The study was limited to seven weeks as the duration of each proficiency level at that preparatory school included seven weeks. The students in the experimental group were notified about the explicit instruction period to ensure they gain the merits of the study for their pragmatic development beforehand. The experimental group took two extra hours including a lesson design with explicit instruction on target speech acts in addition to their main course classes including integrated skills, reading, writing, listening, and speaking each week. On the other hand, control group had two extra hours in addition to their regular classes for general English practice without any intentional pragmatic instruction.

3.2 Setting and Participants

This study was conducted at a preparatory school of a foundation university in İstanbul. The English Preparatory Program was mandatory for nearly all departments at this foundation university. Classrooms are equipped with different types of technological tools including desktop computers with internet access, projectors, and all other class equipment required to enhance learning in each classroom. In the current study, aforementioned classroom equipment was utilized aligned with the targets in the study. The proficiency level of the students is assessed through an online placement test conforming to the Common European Framework of Reference (henceforth, CEFR). There are five levels categorized by CEFR: A1, A2, B1, B1+, and B2. Each level is called a module or a quarter at this foundation university, and they last seven weeks. The study involved B2 level advanced students during the fifth quarter of the English language program in the 2017-2018 academic year in preparatory school. Two advanced level classes were employed for the study. The researcher's class was selected as the experimental group and another random class was chosen as the control group.

Speech act performance is assessed with diverse instruments in the field of pragmatics. In the current study, a WDCT was employed as the data collection instrument instead of MDCT which

provides standardized multiple options to task-takers. WDCT requires open-ended responses produced by task-takers (Knoch, 2009). As the data collected through WDCT enables to collect more authentic results uttered by task-takers, it is decided as the data collection instrument to investigate the effect of explicit instruction on EFL learners.

3.3 Limitations

This study has a few limitations concerning sample size, time, the method and data collection instruments. The first limitation can be regarded as the sample size which is thirty-four advanced EFL learners studying in a preparatory program at a foundation university in İstanbul. The sample size was limited to draw some general conclusions from the study. Another limitation concerns the time consisting of seven weeks in total because of the duration of the module in which the study was conducted. Another constraint is about the research design. As the study was conducted in a limited time, a pretest-posttest design was employed only. If there had been more time, a delayed-posttest could have been used in research design. The instrument might be considered the last limitation in this study, which is WDCT. As a result of the limited time of the study, WDCT was chosen the only data collection instrument. An ODCCT might have been used in addition to WDCT if the duration of the research had been longer.

3.4 Pilot Study

Before collecting the real data for the main study, a pilot study was implemented to test the appropriateness and comprehensibility of the situations designed for WDCT including four speech acts, namely greetings, advice-giving, refusals, and complaints. To execute the pilot study, a WDCT was developed from the previous WDCT designed by Bektaş-Çetinkaya (2012). The WDCT used in the pilot study was tailored to the pilot group participants' age, interests, and characteristics to create authentic scenarios that they might have faced in their real lives.

The pilot study intended to foresee the potential problems in the experimental research. The pilot WDCT was applied to two pilot groups. The former group included three native, one non-native English teacher working in different private schools in Türkiye, and one native English speaker living in the USA. The latter included four advanced EFL learners studying at the preparatory program of the same foundation university in İstanbul, Türkiye. The students in the second group had never visited an English-speaking country and had been learning English for 8-10 years in instructed educational settings. The pilot study enabled to realization of several issues with the WDCT. Whereas the first group including a native speaker and teachers did not identify any difficulty with the WDCT, some student participants had some difficulties understanding some words like “anonymously”, which misleads them to produce proper responses. Some participants in the first group again misunderstood some instructions regarding producing utterances instead of writing their reactions. After identifying those issues, unclear words and situations were simplified to make them more comprehensible for participants. As a result of the pilot study, an edited version of WDCT to be used in the real study was developed and approved after being reviewed by the English teachers and the native speaker.

3.5 Data Analysis

The written data collected through the WDCT was analyzed using Eisenstein and Bodman's (1986) rating scale. To confirm the objectivity of the study, the accuracy of the responses to the

twelve situations in the WDCT was rated by four raters. The researcher of the study was the primary rater, and the other raters, who evaluated the data, also had expertise in English Language Teaching. The second rater, whose results were taken as reference, was an ELT graduate and a doctorate student with native-like fluency in English. Another rater was an English teacher who also undertook a PhD; the last rater was an English teacher pursuing a master's in English Language Teaching. To increase the reliability of the research, the inter-rater reliability among the raters was calculated with Cronbach's alpha in SPSS v.22. The alpha coefficient is .752 indicating acceptable consistency among raters. Then, the WDCT results were coded under a rating scale adapted from Eisenstein and Bodman's (1986) rating scale to categorize the non-native participants' responses to be compared. The categories in Eisenstein and Bodman's (1986) rating scale are presented in Table 2.

Table 2: Rating Scale for Pragmatic Responses

<i>Not acceptable:</i>	A violation of social norm- a likely instance of sociopragmatic failure.
<i>Problematic:</i>	An error that might cause misunderstanding, but of a less serious nature. Language so strange, unexpected, or garbled that interpretation is difficult. Instances of pragmalinguistic and/or sociopragmatic failure.
<i>Acceptable:</i>	Clear and appropriate language but containing small errors which do not interfere seriously with native speakers' understanding.
<i>Native-like/perfect:</i>	Close to native responses in content, syntax, and lexicon.
<i>Not comprehensible:</i>	An utterance that is extremely hard, if not impossible, to comprehend. Often an instance of pragmalinguistic failure.
<i>Resistant:</i>	Non-native participants, although find it possible to answer some items, refuse to answer others or give reasons why they cannot or will not answer particular items.

Source: Eisenstein and Bodman, 1986

The rating scale adapted in the present study includes five categories, namely, acceptable, unacceptable, native-like, no response, and not comprehensible.

4. Discussion and Results

This section briefly analyzes the findings using SPSS v.22 and presents a discussion intending to address the three aforementioned research questions through quantitative analysis of the research findings and results.

4.1 Results

1. Does explicit teaching of speech acts in a classroom setting increase the pragmatic competence of EFL learners?

To answer the first research question, Table 3 presents the descriptive statistics from the pre and post-tests for both the control and experimental groups. This table includes the number of participants, mean scores, and standard deviations for each group in the study.

Table 3: Descriptive statistics of pre and post-tests of WDCT scores

		M	SD	N
PRE	Control	.5524	.17337	17
	Experimental	.6753	.23532	17
	Total	.6138	.21287	34
POST	Control	.5329	.27640	17
	Experimental	.9024	.09589	17
	Total	.7176	.27686	34

M = Mean, SD = Standard Deviation, N = Number of participants

Secondly, to further address the first research question, a two-way mixed ANOVA was executed. Table 4 presents a summary of the repeated-measures effects in the ANOVA, along with the corrected F-values.

Table 4: The summary of Within-Subjects Effects in the ANOVA

		SS	df	MS	F	Sig.
Time	Sphericity Assumed	,183	1	,183	6.053	,019*
Time participants	*Sphericity Assumed	,258	1	,258	8.528	,006*
	Sphericity Assumed	,969	32	,030		
Error(time)						

*p < .05

According to Table 4 results, the interaction between time and participants is significant, $F(1, 32) = 8.528$, $p < 0.05$, $r = 0.045$, indicating a medium effect size. This means that both groups improved over time.

Table 5: The Summary of Between-Subjects Effects in the ANOVA

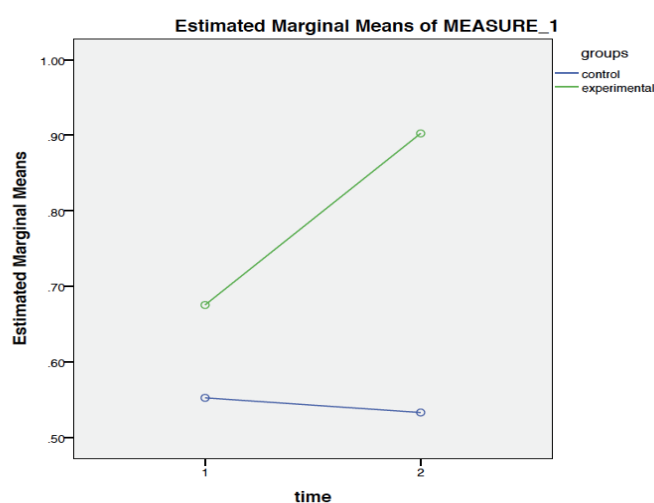
		SS	df	MS	F	Sig.
Time	Sphericity Assumed	30,138	1	30,138	545,585	,000
Time participants	*Sphericity Assumed	1,030	1	1,030	18,651	,000*
	Sphericity Assumed	1,768	32	,055		
Error(time)						

*p < .05

The interaction between explicit instruction and the groups is displayed in Table 5 with a significant interaction, $F(1, 32) = 18.651$, $p < 0.05$, $r = 0.073$, indicating a large effect size. This means that the experimental and control groups were differently impacted over time and the experimental group responded positively to explicit instruction.

To elucidate the interaction between time and instruction, the estimated marginal means were analyzed. It is clearly seen in Figure 4.1.1 that there is a significant increase in the experimental group over the study period. This suggests that the experimental group was more positively affected by the interaction between time and instruction compared to the control group.

Figure 1: Estimated Marginal Means



2. To what degree does the accepted use of target speech acts performed by L2 learners improve towards a native-like degree through explicit teaching?

To address the second research question, the percentage analysis of pretest and posttest results of control and experimental groups was applied. The results of the analysis can be seen in Table 7, Table 8, Table 9, Table 10, Table 11, Table 12, Table 13, and Table 14. The percentage analysis reveals that explicit instruction positively affects the L2 learners' pragmatic development. According to the comparison of the control and experimental groups' posttest percentage analysis, explicit instruction led to a notable improvement in the performance of target speech acts by participants. It is highlighted that the pragmatic improvement was towards an acceptable degree in speech acts of advice-giving, refusals, and complaints and it is further indicated that the improvement was towards a native-like degree in some other speech acts including greetings and refusals.

Table 6 presents the percentage analysis of greetings in the control group's pretest and posttest. It indicates that the percentage of native-like responses of participants to the greeting scenarios in the control group is 39.22 % at the pretest while the percentage of their native-like responses is 37.25 % at the posttest. This indicates that there is no improvement between the two tests.

Table 6: Percentage Analysis of Greetings in Control Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Greeting of a friend	5	8	3	9
Greeting of a professor	5	5	4	3
Greeting of a cashier	5	7	3	7
Total (N = 51)	15	20	10	19
%	29.41	39.22	19.61	37.25

Table 7 shows the percentage analysis of advice-giving in the control group's pretest and posttest. It shows that the percentage of native-like responses of participants to the advice-giving scenarios in the control group is 9.8 % at the pretest while the percentage of the native-like responses is 3.92 % at the posttest. The results show that there is a decrease in native-like use of the speech act of advice-giving in the posttest results.

Table 7: Percentage Analysis of Advice-Giving in Control Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Advice to a friend	9	5	12	1
Advice to a teacher	4	0	3	0
Advice to a stranger	4	0	3	1
Total (N = 51)	17	5	18	2
%	33.33	9.8	35.29	3.92

Table 8 illustrates the percentage analysis of refusals in the control group's pretest and posttest. As it is clear from Table 4.1.5, the percentage of native-like responses of participants to the refusal scenarios in the control group is 3.92 % at the pretest while the percentage of the native-like responses is 68.63 % at the posttest. This suggests a dramatic increase in the posttest without having pragmatic instruction.

Table 8: Percentage Analysis of Refusals in Control Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Refusal of a dinner	8	3	1	12
Refusal of a movie offer	11	1	2	12
Refusal of a studying offer	9	2	4	11
Total (N = 51)	28	6	7	35
%	54.9	3.92	13.73	68.63

Table 9: Percentage Analysis of Complaints in Control Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Complaint to a friend	5	0	0	6
Complaint to a waiter	6	1	2	5
Complaint to a teacher	3	0	3	4
Total (<i>N</i> = 51)	14	1	5	15
%	33.33	9.8	35.29	3.92

Table 9 presents the percentage analysis of complaints in the control group's pretest-posttest. It is analyzed that the percentage of native-like responses of participants to the complaint scenarios in the control group is 9.8% at the pretest while the percentage of the native-like responses is 3.92 % at the posttest. This implies that there is not any improvement in the posttest.

Table 10: Percentage Analysis of Greetings in Experimental Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Greeting of a friend	0	13	4	13
Greeting of a professor	11	4	4	13
Greeting of a cashier	4	9	2	15
Total (<i>N</i> = 51)	15	26	10	41
%	29.41	50.98	19.61	80.39

Table 10 presents the percentage analysis of greetings in the experimental group's pretest and posttest. It is seen that 29.41 % increase the posttest as a result of explicit instruction. This suggests that the accepted use of greeting speech act performed by L2 learners improved to a native-like degree through explicit teaching.

Table 11: Percentage Analysis of Advice-Giving in Experimental Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Advice to a friend	4	8	11	5
Advice to a teacher	5	3	11	2
Advice to a stranger	8	4	8	3
Total (<i>N</i> = 51)	17	15	30	10
%	33.33	29.41	58.82	19.61

Table 11 highlights the percentage analysis of advice-giving in the experimental group's pretest and posttest. A decrease in native-like responses at the posttest results is observed, yet there is a 25.49 % increase in acceptable responses. This implies that explicit teaching improved the accepted use of advice-giving speech act performed by L2 learners towards an acceptable degree, *not* in a native-like degree.

Table 12 below presents the percentage analysis of refusals in the experimental group's pretest-posttest. As it is clear from Table 4.1.9, explicit instruction resulted in a 15.68 % increase in performing the speech act of refusals at the posttest, which indicates there is an increase in acceptable responses at the posttest.

Table 12: Percentage Analysis of Refusals in Experimental Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Refusal of a dinner	4	9	5	12
Refusal of a movie offer	7	6	8	9
Refusal of a studying offer	8	7	7	9
Total (<i>N</i> = 51)	19	22	20	30
%	37.25	43.14	39.22	58.82

Table 13: Percentage Analysis of Complaints in Experimental Group's Pretest-Posttest

	<u>PRETEST</u>		<u>POSTTEST</u>	
	Acceptable	Native-like	Acceptable	Native-like
Complaint to a friend	3	1	13	1
Complaint to a waiter	4	2	12	3
Complaint to a teacher	4	3	15	0
Total (<i>N</i> = 51)	11	6	40	4
%	21.57	11.76	78.43	7.84

Table 13 illustrates the percentage analysis of complaints in the experimental group's pretest-posttest. The results show that there is not an increase in native-like responses in performing complaint speech act with explicit instruction and it is further suggested that the use of complaint speech act performed by L2 learners improves to an acceptable degree instead.

3. Does the efficiency of explicit teaching of pragmatics change according to different speech acts?

To explain the third research question, the experimental group results were analyzed by a two-way ANOVA conducted on SPSS v.22. The efficiency of explicit instruction was discussed for each target speech act performance in the study. The difference in efficacy of explicit intervention among speech acts, which is the focus of the third research question, was calculated

with the effect sizes of pre and posttest results of the four target speech acts investigated in this study.

The two-way ANOVA results conclude that the explicit instruction's efficacy differs in four types of speech acts in the current study. The effect sizes calculated for each speech act are in line with this result. In the present study, four speech acts were analyzed, and the findings show that explicit teaching is effective most for the speech act of greeting with $r = 0.65$, and for the speech act of complaint with $r = 0.59$. It has a medium effect size for the speech act of advice-giving with $r = 0.45$ while it has a small effect size for the speech act of refusals with $r = 0.25$. The statistical analysis of the results supports that explicit teaching is more effective in greeting performance and less effective in the refusal performance of EFL learners who participated in this research.

4.2 Discussion

Kasper (1997) advocates that those particular aspects of pragmatics can be taught through explicit or implicit instruction. Previous studies in the field of interlanguage pragmatics show that the pragmatic competence of learners, who are exposed to either implicit or explicit instruction, improves notably compared to uninstructed learners. Inspired by pioneering studies in the field, this study intended to investigate less-studied speech acts regarding explicit instructional effect, focusing on greetings and advice-giving, alongside the more commonly studied ones like refusals and complaints.

The foremost aim of this study was to examine the efficacy of explicit instruction on the pragmatic competence of L2 learners of English in an EFL classroom, where they are exposed to a limited foreign language input. The primary question was whether explicit instruction would foster an increase in the pragmatic competence of L2 learners after being provided with explicit pragmatic input for a while. Statistical analysis demonstrates a notable difference between the control and experimental group posttest results. The findings of the present study confirm the contribution of explicit instruction to the pragmatic development of EFL learners. This finding is aligned with previous studies by Morrow, 1995; Bacelar da Silva, 2003; Kondo, 2008. The current study supports the findings of Bacelar da Silva (2003) concerning the positive effects of explicit teaching on the pragmatic development of L2 learners. It is also in line with Morrow (1995) findings encouraging the help of explicit instruction to learners to improve their pragmatic performances towards an appropriate, more polite, and more native-like degree.

Additionally, the current study investigated whether the degree of L2 learners' use of target speech acts improved towards a native-like level after explicit instruction. Some other previous studies highlight that explicit instruction fosters the accuracy and appropriateness of responses produced by non-native speakers (King & Silver, 1993; Morrow, 1995; Bacelar da Silva, 2003; Martinez and Fukuya, 2005; Kondo, 2008). This study supports this claim with its findings addressing an increase in acceptable responses in the production of four target speech acts. The percentage analysis of the present study further indicated that L2 learners' performance in refusals and greetings showed a native-like improvement with explicit instruction while the others did show an acceptable degree of improvement. This finding is associated with Morrow (1995) regarding the help of explicit instruction on speech acts for L2 learners to perform clearer, more polite, and somewhat native-like responses. On the other hand, King & Silver (1993) supports there is little effect of explicit intervention on refusals as a result of their study. The present study also encourages that acceptable responses to complaints have increased

through explicit instruction, yet it does not ensure native-like performance. This result contrasts with Morrow (1995) supporting the explicit instruction effect on the native-like performance of complaints.

Finally, this present study intends to take the common claim of the superiority of explicit instruction one step further and it aims to discuss the efficacy of explicit instruction to certain degrees by studying different speech acts. The study could achieve that aim by executing a two-way mixed ANOVA on SPSS v.22 as a result of multiple speech acts examined in the study. The findings of the current study highlight that L2 learners did not equally master all the speech acts through explicit instruction. This claim is supported by the posttest results indicating that greeting speech act performances by L2 learners were improved to a native-like level whereas native-like responses of complaint speech act decreased. This finding is aligned with Cohen and Ishihara (2005). Their findings support that Japanese EFL learners did not equally master all speech acts through explicit instruction. The results of this study are affected by diverse factors including time limitation during the experimentation process, suggesting the time needed to teach and learn different speech acts may differ based on the type of speech act being taught.

5. Conclusion

Interlanguage pragmatics (ILP) focuses on understanding how non-native speakers use and acquire linguistic action patterns in a target language. This study intends to contribute to the field as an interventional study by studying the effect of explicit intervention on four target speech acts. The present study first commits to assessing the pragmatic competence of advanced-level L2 learners by applying a WDCT in the pretest of an experimental design. Then, the experimental group was exposed to explicit instruction including teaching diverse pragmatic aspects of four different speech acts. On the contrary, no explicit instruction exposed to the control group. Finally, a posttest, identical to the pretest, was administered to analyze changes in the pragmatic competence of both groups.

Based on the results of this study, the experimental group was positively influenced by explicit instruction while developing pragmatic competence. The experimental group started to employ politeness strategies and use softeners more accurately after taking explicit instruction during the experimentation process. It was further found that greetings and refusals of L2 learners improved to a native-like degree while advice-giving and complaints improved to an acceptable level in the posttest results. This finding implies that certain speech acts need more time to be mastered to a native-like degree. According to Kasper (2001), positive impacts of instruction on pragmatics in classroom settings experienced by L2 learners were fostered in several studies. Explicit instruction effect on speech acts was investigated in many pioneering studies (Bacelar de Silva, 2003; King & Silver, 1993; Kondo, 2008; Martinez and Fukuya, 2005; Morrow, 1995). Accordingly, learners seem to benefit from explicit instruction by producing more native-like speech acts. This current study is in line with this result regarding the positive effects of explicit instruction on learners' pragmatic development.

To conclude, as Schmidt (1993) notes, pragmatics is not universal and must be explicitly taught in classroom settings. L2 learners can conclusively perform pragmatically appropriate utterances in diverse social settings resulting from explicit teaching. This present study primarily aimed to enhance EFL learners' pragmatic competence through explicit instruction in target speech acts in an EFL classroom. One of the purposes of the study was to compensate

for the drawbacks of learning pragmatics of a language in the EFL context, thus EFL setting and L2 learners were targeted in this study and authentic input enhanced by practice activities were provided as much as possible. The overall results of this study demonstrate that L2 learners benefit from explicit instruction by developing notable pragmatic performance in the target speech act even in a classroom setting. Thus, explicit instruction is a must to incorporate into EFL curricula to enable the essential pragmatic development of EFL learners to fully master the target language.

Appendices

APPENDIX A: Pre-test & Post-Test WDCT

The Discourse Completion Task

Age:

English proficiency level:

Gender: male female

Years of English learning:

Previous Travel to English speaking countries: Yes (specify):

No

Instructions

There are twelve situations described below. Please, read each situation and write down what you would **say** in these situations.

1. When you are shopping in a shopping mall on the weekend, you see one of your good friends. What would you say to your friend?

2. You go to a café to meet your friends and you see your professor sitting on a table near your friends. What would you say to your professor?

3. What would you say when a cashier in a supermarket says to you, “Hi, how are you doing?”

4. A close friend sends you a picture of two t-shirts while s/he is shopping. S/he asks for help to decide on which t-shirt s/he should buy. Actually, you don't like any of them. What would you say to your friend?

5. You like your writing class teacher personally, but you don't like the way s/he teaches the lesson. She asks you to write a feedback about his/her class but s/he doesn't want you to write your name on the feedback paper. How would you give advice to your teacher?

6. You are waiting in line at a florist. The other customer seems to be confused about the choice. You don't have much time to wait, so you have decided to give some advice to him. What would you say to other customer?

7. A friend has moved to a new house and s/he is giving a dinner party on Friday evening. S/he asks you to join them, but you have had a busy week and you want to take some rest. How would you refuse this invitation?

8. A close classmate has an extra movie ticket for tonight, and s/he invites you to watch the newly released horror movie, but you don't like horror movies and you don't want to go. How would you refuse this invitation?

9. You have a final exam tomorrow. You have studied hard, but you feel that you need one more review. One of your friends, who doesn't have much chance to study for the exam, calls you and invites you to his/her home for lunch and study for the exam together. You want to study on your own. How would you refuse this invitation?

10. You have a group assignment and you have shared the tasks; however, one of your group members doesn't like his/her part and s/he doesn't do anything for the assignment. How would you express your complaint to your friend?

11. Imagine your friend is visiting your city for the first time, you are in a good local restaurant and you have ordered some dish, but you wait so much and when your order comes, you realize that they are cold and not fresh. How would you make a complaint to the waiter?

12. You are playing a game in your English class. Your teacher put everyone in different groups. The other groups' members are stronger than yours, then at the end your group lost the game. You think it's unfair. How would you make a complaint about this result to your teacher?

Thanks a lot for your participation!

APPENDIX B: Rating Sheet

Speech acts	Acceptable	Unacceptable	Native-Like	No response	Not comprehensible
1) GREETINGS <i>a. Greeting of a friend</i> <i>b. Greeting of a professor</i> <i>c. Greeting of a cashier</i>					
2) ADVICE GIVING <i>a. Advice to a friend</i> <i>b. Advice to a teacher</i> <i>c. Advice to a stranger</i>					
3) REFUSALS <i>a. Refusal of a dinner</i> <i>b. Refusal of a movie offer</i> <i>c. Refusal of a studying offer</i>					
4) COMPLAINTS <i>a. Complaint to a friend</i> <i>b. Complaint to a waiter</i> <i>c. Complaint to a teacher</i>					

APPENDIX C: Video Sources

https://www.youtube.com/watch?v=orWL34vG3tY&index=8&list=PL_qkUWBh8NQG_VHyAs8tCRD8eDGEoSPUU

https://www.youtube.com/watch?v=nSFAYxO4MA4&index=6&list=PL_qkUWBh8NQG_VHyAs8tCRD8eDGEoSPUU

https://www.youtube.com/watch?v=orWL34vG3tY&index=8&list=PL_qkUWBh8NQG_VHyAs8tCRD8eDGEoSPUU

https://www.youtube.com/watch?v=O8DNNyVJh4Y&index=7&list=PL_qkUWBh8NQG_VHyAs8tCRD8eDGEoSPUU

<https://www.youtube.com/watch?v=S8aXVz799Dc>

<https://www.youtube.com/watch?v=KNVS3oqfbPQ>

References

- Bacelar da Silva, A. J. (2003). The effects of instruction on pragmatic development: Teaching polite refusals in English. *University of Hawai'i Second Language Studies Paper*, 22(1), 1-52.
- Bardovi Harlig, K. (1999). Exploring the interlanguage of interlanguage pragmatics: A research agenda for acquisitional pragmatics. *Language Learning*, 49(4), 677-713.
- Bardovi Harlig, K., & Griffin, R. (2005). L2 pragmatic awareness: Evidence from the ESL classroom. *System*, 33(3), 401-415.
- Bardovi Harlig, K., & Hartford, B. S. (1990). Congruence in native and nonnative conversations: Status balance in the academic advising session. *Language Learning*, 40(4), 467-501.
- Bektas-Cetinkaya, Y. (2012). Pre-service EFL teachers' pragmatic competence: The Turkish case. *International Journal of Language Studies*, 6(2).
- Björgvinsson, L. Á. (2011). *Speech Act Theory: A Critical Overview*. The University of Iceland, Iceland.
- Cohen, A. D., & Ishihara, N. (2005). *A web-based approach to strategic learning of speech acts*, 1-57.
- Crystal, D. (1985). *A dictionary of linguistics and phonetics*. Oxford: Blackwell.
- Eisenstein, M., & Bodman, J. W. (1986). 'I very appreciate': expressions of gratitude by native and non-native speakers of American English. *Applied Linguistics*, 7(2), 167-185.
- Ishihara, N., & Cohen, A. D. (2010). Teaching and learning pragmatics: *Where language and culture meet*. Routledge.
- Kasper, G. (1997). Can pragmatic competence be taught?. Honolulu: University of Hawai'i, Second Language Teaching & Curriculum Center, [Online]. Available: <https://nflrc.hawaii.edu/NetWorks/NW06/>
- Kasper, G., & Blum Kulka, S. (1993). Interlanguage pragmatics: An introduction. *Interlanguage Pragmatics*, 3, 15.
- Kasper, G., & Schmidt, R. (1996). Developmental issues in interlanguage pragmatics. *Studies in Second Language Acquisition*, 18(2), 149-169.
- Kasper, G. (2001). Classroom research on interlanguage pragmatics. *Pragmatics in Language Teaching*, 33-60.
- King, K. A., & Silver, R. E. (1993). "Sticking Points": Effects of Instruction on NNS Refusal Strategies. *Working Papers in Educational Linguistics*, 9(1), 47-82.

- Knoch, U. (2009). Diagnostic assessment of writing: A comparison of two rating scales. *Language Testing*, 26(2), 275-304.
- Koike, D. A., & Pearson, L. (2005). The effect of instruction and feedback in the development of pragmatic competence. *System*, 33(3), 481-501.
- Kondo, S. (2008). Effects on pragmatic development through awareness-raising instruction: Refusals by Japanese EFL learners. *Investigating Pragmatics in Foreign Language Learning, Teaching and Testing*, 153-177.
- Martínez-Flor, A., & Fukuya, Y. J. (2005). The effects of instruction on learners' production of appropriate and accurate suggestions. *System*, 33(3), 463-480.
- Mey, J. L. (2001). *Pragmatics: An Introduction*. Oxford: Blackwell.
- Morrow, C. K. (1995). *The pragmatic effects of instruction on ESL learners' production of complaint and refusal speech acts*. State University of New York, Faculty of the Graduate School, Buffalo.
- Schauer, G. (2009). *Interlanguage Pragmatic Development: The Study Abroad Context*. Bloomsbury Publishing.
- Schmidt, R. (1993). Consciousness, learning and interlanguage pragmatics. *Interlanguage Pragmatics*, 21, 42.
- Takahashi, S. (2010). The effect of pragmatic instruction on speech act performance. *Speech Act Performance: Theoretical, Empirical and Methodological Issues*, 127-144.
- Tuncay, H. (2013). *English Language Teaching & Linguistic Dictionary* (2). İstanbul: Yalın Publishing.
- Ülbeği, E. (2009). The effects of implicit vs. explicit instruction on pragmatic development: teaching polite refusals in English. *Uludağ Üniversitesi Eğitim Fakültesi Dergisi*, 22(2), 327-356.

Comparison of EAP Exam Texts with Undergraduate Readings: A Coh-Metrix Analysis

Yusuf Cengiz*

*PhD(c), Boğaziçi University, Türkiye

Abstract

Students aiming to enroll in English-medium universities are required to meet language proficiency standards either by completing intensive English programs or by passing language exams. Universities in the Expanding Circle (Kachru, 1990) often mandate proficiency proofs via internationally recognized exams such as TOEFL and IELTS or through internal assessments. These exams are designed to evaluate students' proficiency in academic English. The selection of texts for classroom and assessment contexts has significant implications for curriculum planning and evaluation, emphasizing the importance of a systematic approach to text selection (Watkins & Ostenson, 2015). Texts need to be carefully chosen to ensure they facilitate the desired linguistic competencies (Hiebert & Mesmer, 2013). While expert opinions and experimental testing are traditional methods for evaluating text complexity, automated tools can also be used to ensure texts are suitably challenging (Sheehan et al., 2014). However, literature review indicates limited studies on using automated tools for assessing text appropriateness in high-stakes exams. This study aims to fill this gap by comparing a corpus from freshman coursebooks with a corpus from in-house materials used for academic language proficiency testing. The findings provide a benchmark for developing in-house materials aligned with the linguistic complexity of academic texts. Results show that the in-house materials and freshman year texts differ in only two indices, with no significant differences across the remaining ten indices.

Keywords: Testing and assessment, text analysis, automated tools

1. Introduction

As a result of the increased role of English-medium instruction within the university contexts around the world (Lasagabaster, 2018), the need to implement an approach to English language education that caters for the need of the academic sphere has become more prevalent. Students who wish to pursue degrees at English-medium universities are either asked to go through intensive English programs to meet the language requirements or asked to prove their language proficiency through exams. Universities, such as Boğaziçi University or universities that are located at the Expanding Circle (Kachru, 1990), ask their students to either prove their language proficiency through internationally accepted exams, such as TOEFL and IELTS, or through in-house proficiency exams. These exams aim to measure the students' proficiency in using the English language within academic contexts. The exam that is used at Boğaziçi University is named Boğaziçi University English Proficiency Test (BUEPT) and the exam is used to determine whether a test taker meets the necessary language proficiency to pursue their degrees at the university. As a result, the contents of the exams are assumed to be reflective of the academic genre and conventions, which requires a meticulous text selection and editing process.

The exam consists of Reading, Listening and Writing components, each of which has two sub-components. The study at hand aims to investigate whether the texts for the reading components of the exam matches the academic language used in the freshman year programs of the university. To this end, in-house texts that are deemed to be similar to the BUEPT level are compared to the excerpts taken from the freshman year textbooks, using automated evaluation tools, Coh-Metrix 3.0 (McNamara et al., 2014) and VocabProfile (Cobb, 2015).

2. Literature Review

Texts used in the classroom and language assessment contexts have important consequences on planning and evaluating, which makes the process of text selection a vital issue (Watkins & Ostenson, 2015). The text that are to be used in the classroom and assessment practices need to be chosen methodically to ensure that attainment to the desired linguistic practices will be or have been achieved (Hiebert & Mesmer, 2013). While the selection process can be carried out through expert opinions or experimental testing to see if the texts have the necessary complexity, automated tools can also be implemented throughout the process (Sheehan et al., 2014) to ensure that the texts are appropriately challenging. However, a review of the literature reveals that the studies that utilizes automated tools to investigate the appropriateness of a text for high stake assessments are limited.

One such practice can be observed in Green et al.'s study (2010). The researchers carried out an extensive study in which excerpts taken from the reading component of the IELTS exams have been compared with the excerpts taken from the readings done in the freshman year classes at University of Bedfordshire. The study employed both qualitative and quantitative methods to test whether the text complexity of the freshman year coursebooks was attained in the IELTS exams. The researchers employed automated tools, such as Coh-Metrix, VocabProfile and Microsoft Word to obtain quantitative data on various text complexity measures. The qualitative data from the automated tools included vocabulary-level measures such as type/token ration, word frequency; sentence-level measures, such as word per sentence, number of noun clauses and the number of words before the main verb; and text-level measures such as

Flesch-Kincaid readability score, lexical density and the use of logical operators, i.e. and and or. The analysis revealed that IELTS tests measured similar to freshman year texts in many aspects, however there were statistically significant differences in terms of use of words from academic word lists and sub-technical words. IELTS texts employed fewer words from the academic word lists and contained fewer sub-technical words than freshman year texts. Freshman year texts, on the other hand, contained fewer content words. These results of these measures imply that IELTS texts measure similar to freshman year readings, however, they may still fall short on text complexity, especially when the language demands of more challenging and technical areas, such as law, are concerned.

Owing to the versatile use of automated text evaluation tools, studies conducted within the field are able to approach text appropriacy and text complexity via different designs. For example, Sheehan et al. (2014) uses another automated tool, TextEvaluator to test whether the same text would be judged similarly in terms of text by human raters and TextEvaluator. Their study found out that the automated tool produced similar evaluations of a text when compared with human raters. TextEvaluator generated results based on the vocabulary profile, syntactic complexity, lexical cohesion and argumentation. The results obtained from the tool correlated those of the human raters, thus suggesting that there is room for automated tools to be utilized to evaluate text complexity and appropriacy while choosing or adapting texts for teaching and assessment purposes.

As texts that are used in assessment usually go through adaptation processes, automated evaluation tools are used to ensure that the text complexity is matched with the expected level of the students. In one such study, Green & Hawkey (2012) investigated the text adaptation process of IELTS test writers to measure to what extent the texts used for the reading components of IELTS tests are altered during the adaptation process. The researchers made use of automated tools to investigate Flesch-Kinkaid readability scores and word frequencies of the texts before and after the adaptation process. The analysis revealed that the overall reading complexity was maintained during the adaptation process, as Flesch-Kinkaid results were marginally changed, while vocabulary profile was altered to decrease the number of technical and field-specific words from the source texts. The results serve both as a benchmark to compare IELTS texts with field related tests, but also as a reminder to indicate that during the adaptation of the texts, certain academic features, especially in terms of vocabulary profile, are sacrificed for assessment purposes.

Automated tools have also been used to test the validity of multiple assessment tools. Ilc and Stopar (2015) investigates three different tests used within the Slovenian education system to reveal whether these texts, which were prepared in accordance with B2 CEFR level, produce similar results in terms of their text complexity and cohesion. The researchers make use of Coh-metrix to check complexity and cohesion measures. The results indicate that although the tests are similar in terms of presenting B2 level reading tests, two of the tests contain longer reading passages with more challenging vocabulary items, which directs the researchers to call for a reconsideration of the text preparation aspect of these tests. Ha et al. (2022), on the other hand, investigates whether the vocabulary profile of IELTS General and IELTS Academic exams do differ by using automated tools. Their study reveals that texts used in IELTS General are indeed less challenging in terms of vocabulary profile. Their study, however, do not utilize any other quantitative measurement to investigate text complexity or appropriacy.

Despite the usefulness of automated tools for evaluating text appropriacy and complexity, as stated above, the studies that utilizes these tools for reading test development or evaluation are limited. This could stem from the fact that popular tools such as Coh-metrix and TextEvaluator are not readily available now due to certain constraints. While Coh-metrix can only be accessed through manual requests via a Google Forms link, TextEvaluator is not commercially available. The literature reveals that automated tools have also been used to investigate the text complexity of students' written productions (Crossley et al., 2011, 2012; Lu, 2017), or to investigate the correlations between certain text complexity measures and grades (Westerlund, 2019). The studies present a room for the use of automated tools in students' written assessments by evaluating the text complexity of the students' production, where certain indices, such as lexical diversity and word frequency positively correlate with the grades of human raters, making the case for the use of automated tools in evaluating text complexity. As the scope of this study is to focus on the reading aspect, these studies are not discussed in detail.

However, the literature at hand still reveals that automated text evaluation tools can be used to investigate the text complexity. The study at hand aims to contribute to the literature by providing a comparison of a corpus obtained from freshman year coursebooks and a corpus obtained from in-house materials that are aimed at testing students in terms of academic language proficiency. The study can provide a benchmark for in-house materials to be produced in line with the language complexity required in academic texts. To this end, the study aims to answer the following research questions:

RQ1. How do in-house texts differ from freshman year coursebook texts in terms of text complexity based on their vocabulary profile obtained from automated tools?

RQ2. How do in-house texts differ from freshman year coursebook texts in terms of text complexity based on their text-level indices obtained from automated tools?

3. Methodology

3.1 Context

The study is conducted at Boğaziçi University, an English-medium university in İstanbul, Türkiye. Almost all the programs at Boğaziçi University are offered in English and as a result, students are asked to either submit a valid TOEFL iBT (min. 79 in total) or IELTS score (min. 6.5 total) or pass Boğaziçi University English Proficiency Test (BUEPT) as a proof of English proficiency before they can start their undergraduate classes. The minimum passing grade for BUEPT is "C", which is proposed to be equivalent of the minimum scores for TOEFL iBT and IELTS exams. The exam consists of Reading, Listening and Writing components, each of which has two sub-components.

Those who fail to satisfy the necessary English proficiency are accepted to one-year intensive English program. Towards the end of the one-year intensive English program, the students are provided with materials that are prepared in line with the BUEPT to prepare students for the exam.

3.2 Data Collection

The data obtained for the study comes from two major sources. 26 texts were selected from the reading texts that are used to prepare the students at the end of one-year intensive English program. These in-house texts are prepared by the program's curriculum and assessment units. They are selected based on the themes that are present at the freshman year programs, such as history, education and economics.

The average length of the in-house texts is 1456 words, however, for analysis, excerpts of 1000 words were extracted from each text as the automated tools are advised not to be used with texts over 1000 words. The texts are taken from academic books or articles and then adapted for assessment purposes. The adaptation is done by the members of the curriculum and the testing units, which composes of the instructors working at the intensive English program.

In order to form a corpus of texts from the freshman year programs, the curriculum of each department at the university was analyzed. Courses that are offered to the freshman year students were extracted and later their syllabi were taken from the university's official website. These syllabi were investigated to determine the textbooks that are used throughout the Fall term. In cases where multiple textbooks are to be used, the one that is to be used dominantly throughout the term was selected for analysis.

Out of 33 undergraduate programs, 8 of them were not investigated as either their syllabi were not available or they included classes that are heavily dependent on language studies and contain reading genres that are not targeted through the in-house materials, such as English Literature or Turkish literature.

Out of the remaining 25 programs, the syllabi of 14 undergraduate courses were selected for the analysis as certain undergraduate courses are mass classes that are shared by multiple departments. For example, selected mathematics course for the analysis is shared by the department of mathematics and all engineering departments.

Selected freshman year courses were from the following departments: chemistry, economics, education, history, linguistics, management, mathematics, molecular biology and genetics, philosophy, physics, politics and international relationships, psychology, sociology, and tourism. In order to increase the representativeness of the data, two 1000-word excerpts were taken from the selected textbook of each freshman year course. In total, 28 excerpts were created.

The texts taken from the freshman year textbooks and the in-house materials were checked and edited to ensure that there are no spelling or paragraphing errors.

3.3 Analysis

For the analysis, two tools were utilized. Coh-Metrix 3.0 was used to check for the following variables in Table 1.

Table 1: Coh-metrix indices used in the analysis

Indice Label	Explanation
DESSL	Sentence length, number of words, mean
DESWLsy	Word length, number of syllables, mean
DESWLlt	Word length, number of letters, mean
PCSYNz	Text Easability PC Syntactic simplicity, z score
CNCLogic	Logical connectives incidence
SYNLE	Words before main verb, mean
SYNNP	Number of modifiers per noun phrase
DRNP	Noun phrase density,
DRVP	Verb phrase density,
WRDAOAc	Age of acquisition for content words,
WRDFAMc	Familiarity for content words,
WRDCNCc	Concreteness for content words,
RDFRE	Flesch Reading Ease
RDFKGL	Flesch-Kincaid Grade Level
RDL2	Coh-Metrix L2 Readability

In order to check for the vocabulary profile, VocabProfile website was utilized. The following indices were used in the analysis.

Table 2: VocabProfile indices

Label	Explanation
@1K	Words at 1K
@2K	Words at 2K
AWL	Academic Word Lists
Offlist	Off-list
TTR	Type-token ratio
DENS	Syntactic density (content words/all)

The following indices have been taken as indicators of vocabulary level indices to answer RQ1: DESWL_{sy}, DESWL_{lt}, WRDAOAc, WRDFAMc, WRDCNCc, @1K, @2K, AWL, Offlist, TTR, DENS.

DESWL_{sy} stands for the mean number of syllables per word in a given text, while DESWL_{lt} represent the mean number of characters per word. As longer words indicate reading difficulty (Green et al., 2010), these measures were taken as an indicator of textual difficulty. WRDAOAc represent the mean age of acquisition for the content words. WRDFAMc, WRDCNCc measures also return values regarding content words. While the former provides the familiarity value for the word, the latter provides values pertaining the concreteness of a word. The higher the values, the more familiar and the more concrete the words are, making a text easier to comprehend.

@1K and @2K indices returns the percentage of all the words within a text that are within the most frequent 1000 and 2000 word lists on British National Corpus respectively, while AWL returns the percentage of words contained within the academic word lists. Offlist indices indicates the percentage of words that are not contained in any of the lists, which are usually non-frequent words or technical, subject specific words. TTR stands for type/token ratio, where a higher ratio would mean that the text composes of more different words, which increases the amount of words needed to be in the repertoire of the reader for comprehension. DENS indice returns the value for syntactic density, which is obtained when the number of all the words in a text is divided by the number of total content words. A higher value would indicate that they vocabulary density of a text is also higher, making it more challenging to comprehend.

In terms of text-level complexity, the following indices have been investigated to answer RQ2: PCSYNz, CNCLo_{gic}, SYNLE, SYNNP, DRNP, DRVP, RDFRE, RDFKGL, and RDL2. PCSYNz stands for syntactic simplicity. As the value is calculated as a z-score, values at the positive end indicate higher simplicity, while negative values indicate that the text is complex in terms of syntactic structures. CNCLo_{gic} calculates the incidences of logical operators, such as “and” and “or”. Higher incidences indicate texts that are more challenging to comprehend. SYNLE gives the mean number of words before the main verb of a sentence. A higher value would indicate that the main verb appears much later in a sentence, making it more difficult to process. SYNNP value returns the mean number of modifiers per noun phrase, which indicates the length of noun phrases, while DRNP gives the noun phrase density within the text. DRVP, on the other hand provides the value for verb phrase density. Higher values indicate denser texts with complex syntactical structures.

Coh-metrix also provides 3 indices for overall readability, which are represented with RDFRE, RDFKGL and RDL2 indices. RDFRE stands for Flesch Reading Ease, while RDFKGL stands for Flesch-Kincaid Grade Level. RDL2 represents Coh-Metrix L2 Readability. All of these indices provide an overall view to the text complexity.

For the analysis, Coh-metrix and VocabProfile analysis were conducted for each text on the given indices. The values were then analyzed using SPSS ver. 29. A test of normality revealed that one of the texts among in-house materials was a significant outlier. A closer examination revealed that the text, titled “Powder” was actually a short-story, rather than an academic reading text. As a result, this text was taken out from the analysis. An independent sample t-test was run to test whether the differences between the values for the indices differed

significantly between in-house materials and texts taken from freshman year courses. The results and discussion are provided in the following section.

4. Results and Discussion

The following tables contain the results of the independent samples t-test analysis. Table 3 contains the results of the tests that included vocabulary-level indices for RQ1.

Table 3: VocabProfile indices

Independent Samples Test							
	t-test for Equality of Means						
	Freshman Year Texts		In-House Texts		t	df	Significance
	Mean	Std. Deviation	Mean	Std. Deviation			Two-Sided p
DESSL - Sentence length, number of words, mean	22,576	3,941	23,546	4,028	-.885	51	.380
DESWLsy - Word length, number of syllables, mean	1,679	0,121	1,723	0,099	-1.461	51	.150
DESWLlt - Word length, number of letters, mean	4,982	0,367	5,165	0,273	-2.040	51	.047*
WRDAOAc - Age of acquisition for content words	383,897	21,543	389,207	20,709	-.912	51	.366
WRDFAMc -Familiarity for content words,	562,592	8,390	563,319	6,456	-.351	51	.727
WRDCNCc - Concreteness for content words,	362,922	21,801	367,605	20,082	-.810	51	.422
@1K - Words at 1K	79,526	4,833	80,948	3,378	-1.227	51	.225
@2K - Words at 2K	4,754	1,624	4,648	1,788	.225	51	.823
AWL - Academic Word Lists	8,129	2,731	7,736	2,775	.518	51	.606
Offlist - Off-list	7,136	3,578	6,288	2,125	1.033	51	.306

TTR - Type-token ratio	0,383	0,048	0,421	0,043	-3.040	51	.004*
DENS - Syntactic density (content/all)	0,571	0,035	0,573	0,038	-.214	51	.832

The results indicate that the in-house materials and freshman year texts differ only in two indices, while the remaining 10 indices do not show any significant differences between the two sources of text. Significant differences can be observed in DESWLt, which stands for the mean number of characters per word. The results indicate that in-house texts contain longer words than freshman year reading texts, $t(51) = -2.040$, $p < 0.05$. As stated by Green et al. (2010), word length composes an important aspect of reading difficulty. Certain studies (Joseph et al., 2009; Kramer & McLean, 2019) also point out that words with more letters take longer to process, ultimately making it longer and more challenging to complete the comprehension process. Although the extensive review of Barton et al. (2014) also point out a similar effect, they also point out that word length effect cannot explain reading difficulty alone, as word frequency and syllable count also plays a role in word recognition, as well as the concreteness of a word. This can also be observed within the results, as indices related to mean number of syllables per word and content word concreteness rest at statistically non-significant values.

The difference in type/token ratio between the two sources of texts is also statistically significant. The analysis indicate that in-house materials contain more types of words than freshman year reading texts, $t(51) = -3.040$, $p < 0.01$. As indicated by Green et al. (2010), an increased type/token ratio would mean that the reader is required to deal with more words while reading the text, which in turn might increase the reading difficulty. Although this may be the case, it should also be noted that the difference between the two text sources might stem from the fact that in-house materials are designed for test taking purposes and the result that the lexical density, i.e. the ratio of content words for all words, is almost the same for the text, makes the case for the assumption that the difference between type/token ratio might stem from function words, rather than content words. As the in-house materials are adapted for testing purposes, re-arranging function words is expected for test item development. Table 4 contains the results from the analysis that investigated text-level indices.

None of the indices used in the analysis of text-level differences have returned statistically significant results. Based on the analysis, it can be pointed out that the excerpts used as the dataset contained similarly challenging texts. As Coh-metrix 3.0 provides more indices regarding textual complexity, more analysis could be conducted regarding the issue. For example, Landauer (2007) points out Latent Semantic Analysis, which investigates the overlaps between sentences and paragraphs as indicators of cohesion. Texts with higher cohesion are deemed as easier to understand, while those with lower cohesion are more challenging. Such aspects can be investigated to see if the difference between the two text sources lie in other areas, if they do. It should also be noted that the lack of any significant difference between the in-house materials and freshman year texts does not necessarily indicate that the two text types are indistinguishable, as the dataset contains only the main reading texts, but not supplementary and more challenging texts provided in the freshman year programs (Green et al., 2010). Additionally, as the freshman year texts come from textbooks used in the program, rather than academic journal articles, a simpler language could be expected from the textbooks.

Table 4: Results of independent samples t-test for text-level indices

	Independent Samples Test						
					t-test for Equality of Means		
	Freshman Year Texts		In-House Texts		Significance		
	Mean	Std. Deviation	Mean	Std. Deviation	t	df	Two-Sided p
PCSYNz - Text Easability PC Syntactic simplicity, z score	-0,232	0,524	-0,299	0,477	.483	51	.631
CNCLogic - Logical connectives incidence	39,899	10,031	38,429	8,330	.576	51	.567
SYNLE - Words before main verb, mean	5,936	1,313	5,936	1,503	.000	51	1.000
SYNNP - Number of modifiers per noun phrase	0,979	0,113	1,004	0,110	-.794	51	.431
DRNP - Noun phrase density,	375,002	22,058	369,336	21,128	.952	51	.345
DRVP - Verb phrase density,	179,075	30,498	183,448	25,714	-.561	51	.577
RDFRE - Flesch Reading Ease	42,130	11,899	37,349	9,328	1.614	51	.113
RDFKGL - Flesch-Kincaid Grade Level	12,936	2,344	13,848	1,960	-1.526	51	.133
RDL2 - Coh-Metrix L2 Readability	12,547	4,050	12,306	3,493	.231	51	.819

5. Conclusion and Limitations

This study investigated the vocabulary level and text level differences between in-house materials for intensive English students and the excerpts taken from the textbooks used in the freshman year programs at Boğaziçi University. The results indicated that there were no statistical differences in 19 of the indices while only 2 of the indices produced significant differences, both of which at the vocabulary-level.

The results imply that the materials developed by the intensive English program provide texts as challenging as the ones encountered at the freshman year courses, as shown by the results

obtained from automated tools. However, as human rating is still necessary to fully capture the differences between the two texts, the results should be taken with a grain of salt. However, the study at hand still implies that automated tools can provide a window for text evaluation for appropriateness.

The study is limited as the validity of the indices of Coh-metrix still needs further validation through more studies. Also, the corpus could be increased to be more inclusive of the contents provided at the freshman year, as well as more texts could be obtained from the in-house developers. Human raters should also be included to judge text appropriateness and text complexity to triangulate the obtained data.

References

- Barton, J. J. S., Hanif, H. M., Eklinder Björnström, L., & Hills, C. (2014). The word-length effect in reading: A review. *Cognitive Neuropsychology*, 31(5–6), 378–412. <https://doi.org/10.1080/02643294.2014.895314>
- Cobb, T. Range for texts v.3 [computer program]. Accessed 15 Sept 2015 at <https://www.lex tutor.ca/cgi-bin/range/texts/index.pl>.
- Crossley, S. A., Salsbury, T., & McNamara, D. S. (2012). Predicting the proficiency level of language learners using lexical indices. *Language Testing*, 29(2), 243–263. <https://doi.org/10.1177/0265532211419331>
- Crossley, S. A., Salsbury, T., McNamara, D. S., & Jarvis, S. (2011). Predicting lexical proficiency in language learner texts using computational indices. *Language Testing*, 28(4), 561–580. <https://doi.org/10.1177/0265532210378031>
- Green, A., & Hawkey, R. (2012). Re-fitting for a different purpose: A case study of item writer practices in adapting source texts for a test of academic reading. *Language Testing*, 29(1), 109–129. <https://doi.org/10.1177/0265532211413445>
- Green, A., Ünalı, A., & Weir, C. (2010). Empiricism versus connoisseurship: Establishing the appropriacy of texts in tests of academic reading. *Language Testing*, 27(2), 191–211. <https://doi.org/10.1177/0265532209349471>
- Ha, H. T., Le, H. T., Phung, D. H., & Nguyen, S. D. (2022). Is “general” easier than “academic”? A corpus-based investigation into the two modules of IELTS reading test. *SN Social Sciences*, 2(8), 159. <https://doi.org/10.1007/s43545-022-00461-1>
- Hiebert, E. H., & Mesmer, H. A. E. (2013). Upping the Ante of Text Complexity in the Common Core State Standards: Examining Its Potential Impact on Young Readers. *Educational Researcher*, 42(1), 44–51. <https://doi.org/10.3102/0013189X12459802>
- Ilc, G., & Stopar, A. (2015). Validating the Slovenian national alignment to CEFR: The case of the B2 reading comprehension examination in English. *Language Testing*, 32(4), 443–462. <https://doi.org/10.1177/0265532214562098>
- Joseph, H. S. S. L., Liversedge, S. P., Blythe, H. I., White, S. J., & Rayner, K. (2009). Word length and landing position effects during reading in children and adults. *Vision Research*, 49(16), 2078–2086. <https://doi.org/10.1016/j.visres.2009.05.015>
- Kachru, B. B. (1990). World Englishes and applied linguistics. *World Englishes*, 9(1), 3–20. <https://doi.org/10.1111/j.1467-971X.1990.tb00683.x>
- Kramer, B., & McLean, S. (2019). L2 reading rate and word length: The necessity of character-based measurement. *Reading in a Foreign Language*.

Landauer, T. K. (2007). LSA as a theory of meaning. In *Handbook of latent semantic analysis* (pp. 15–46). Psychology Press.

Lasagabaster, D. (2018). Fostering team teaching: Mapping out a research agenda for English-medium instruction at university level. *Language Teaching*, 51(3), 400–416. <https://doi.org/10.1017/S0261444818000113>

Lu, X. (2017). Automated measurement of syntactic complexity in corpus-based L2 writing research and implications for writing assessment. *Language Testing*, 34(4), 493–511. <https://doi.org/10.1177/0265532217710675>

McNamara, D. S., Graesser, A. C., McCarthy, P. M., & Cai, Z. (2014). *Automated evaluation of text and discourse with Coh-Metrix*. Cambridge University Press.

Sheehan, K. M., Kostin, I., Napolitano, D., & Flor, M. (2014). The TextEvaluator tool: Helping teachers and test developers select texts for use in instruction and assessment. *The Elementary School Journal*, 115(2), 184–209. <https://doi.org/10.1086/678294>

Watkins, N., & Ostenson, J. (2015). Navigating the Text Selection Gauntlet: Exploring Factors That Influence English Teachers' Choices. *English Education*, 47(3), 245–275.

Westerlund, M. (2019). Correlations Between Textual Features and Grades on the Swedish National Exam in English: A Coh-Metrix Analysis

Organized by



Sponsored by



**INTERNATIONAL CONFERENCE
OF WITHIN AND BEYOND THE
CLASSROOM**



