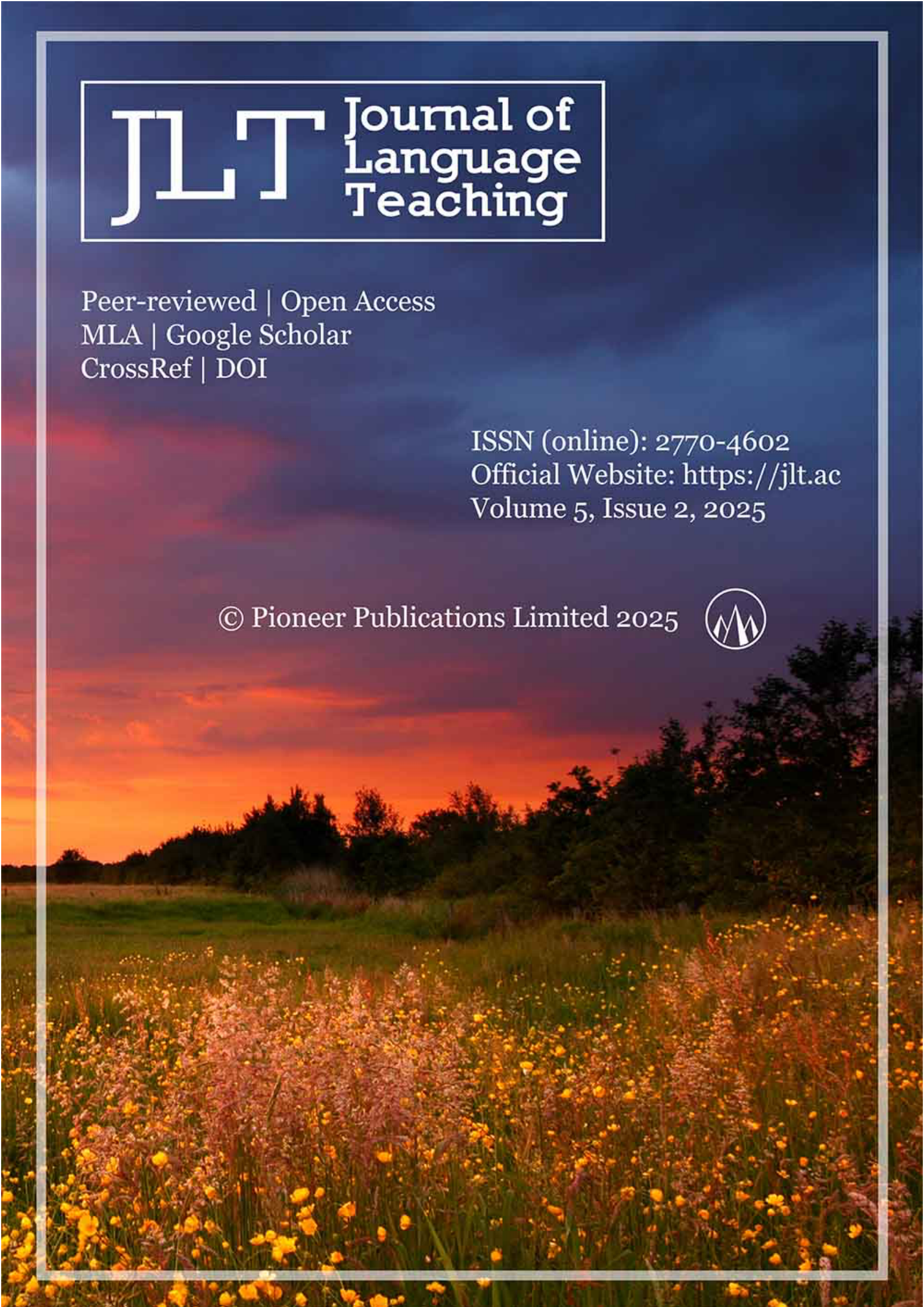
JLT Journal of Language Teaching

Peer-reviewed | Open Access
MLA | Google Scholar
CrossRef | DOI

ISSN (online): 2770-4602
Official Website: https://jlt.ac
Volume 5, Issue 2, 2025

© Pioneer Publications Limited 2025



Journal of Language Teaching

ISSN (online): 2770-4602
Vol. 5, Issue 2, 2025

CONTENTS

Please visit our website to read more articles for free: https://jlt.ac/
Article Submission: https://jlt.ac/home/about/submissions
© Pioneer Publications LTD 2021-2025

Journal of Language Teaching, *5*(2), 1-16, 2025
https://doi.org/10.54475/jlt.2025.006

Check for updates

ORIGINAL RESEARCH OPEN ACCESS

Migrant Students' Emotions and Perceptions of collaborative writing: The role of educational setting, educational level and (non-)migrant background

Anastasia Paspali, Despina Papadopoulou

LingLab, Department of Linguistics, School of Philology, Faculty of Philosophy, Aristotle University of Thessaloniki, Thessaloniki, Central Macedonia, Greece

Received: April 15, 2025 / Accepted: June 1, 2025 / Published Online: June 6, 2025
© Pioneer Publications LTD 2025

Abstract

Within a socio-cultural framework of language teaching and a mixed methods design, the present study explores the emotional impact of collaborative writing activities on (non-) migrant students in formal education as well as their perceptions towards collaboration. The study examined whether emotions and perceptions about collaboration are differentially affected by educational level (primary vs. junior high school), educational setting (reception vs. mainstream class) as well as students' background (migrant, non-migrant). The data were obtained by means of questionnaires and were analyzed both quantitatively by means of linear mixed-effects models and qualitatively by means of thematic analysis. The results revealed that students overall exhibited more positive than negative emotions. Even more, primary school students in the mainstream class exhibited weaker negative emotions compared to students in the reception class, and marginally stronger positive emotions and weaker negative emotions compared to junior high school students in the mainstream class. Primary school students in the mainstream class were the only group who did not exhibit negative emotions towards collaboration. Furthermore, junior high school students exhibited more negative emotions and reported more collaboration difficulties compared to primary school students, which could be associated with different conceptualization of group work among different ages as well as with different social group dynamics between the mainstream class of primary school and the mainstream class of junior high school along with different orientation of the school curricula.

Keywords migrant education, collaborative writing, dictogloss, language learning in migrant contexts, school setting, school level, migrant background, grammar teaching intervention, group work in formal education settings

1. Introduction

As the child population from migrant backgrounds grows, so too does the need for language teaching interventions to promote their development in the language of the host country and the language of schooling, which is critical for their integration. Crucially, the interventions need to be appealing to students since the emotional impact of a teaching intervention on students is associated with their learning process (see Franck & Papadopoulou, 2024 for discussion). Furthermore, recent studies embrace interventions and methods promoting cooperative learning in migrant students since it has been argued to facilitate language learning, socialization, conflict management, sense of belonging, and thus, emotional development and well-being (Ferguson-Patrick, 2020). Consequently, knowing how migrant students feel during the learning process when certain teaching intervention methods are employed, as well as how they collaborate with each other in various learning environments, is of utmost importance.

To date, studies on the emotional impact of language teaching interventions on migrant students as well as on their perceptions towards collaborative tasks in formal education are limited (Busse et al., 2020, 2021). This area of research is highly critical given that migrant students form a highly vulnerable population, often experiencing trauma and anxiety (Ferguson-Patrick, 2020). In addition, migrant students do not form a uniform population, and their educational needs may differ and/or dynamically change across the various learning environments. For this reason, it is highly important to study which factors may differentially affect their emotions and perceptions towards the learning process. This can, in turn, reveal new insights into good teaching practices and more tailored educational materials that promote migrant students' educational growth as well as their well-being.

The present study is the first one, to our knowledge, which explores (a) the emotional impact of a collaborative language teaching intervention with migrant students, (b) students' perceptions towards collaboration in classroom

while also exploring (c) the role of educational level (primary vs. junior high school), educational setting (reception vs. mainstream class), and student background (migrant vs. non-migrant).

2. Language teaching in migrant students: focusing on students' emotions and collaboration in formal education

Migrant students have been reported to exhibit lower school satisfaction and social belonging, along with lower school performance compared to non-migrant students (Göbel & Frankemölle, 2020; Henschel et al., 2019). Crucially, their school performance has been found to be higher when they experience more happiness at school and a high sense of belonging (OECD, 2015). Thus, designing appealing teaching interventions which promote (all) students' well-being, and stimulating positive emotions and attitudes is of paramount importance for integration and for lowering the risk of school failure and school dropout in the migrant population.

However, research on measuring the emotional outcomes of language teaching on the migrant child population is limited. To our knowledge, the existing studies focus on EFL classes in formal school settings (Busse et al., 2020; 2021). Given the limited studies on the topic, more attention needs to be paid to this line of research (Busse et al., 2020; Philp and Duchesne 2016; Swain, 2013), and specifically to language teaching interventions which target the language of the host country (both in reception as well as in mainstream classes). Busse et al. (2020; 2021) applied two language teaching intervention studies targeting vocabulary in EFL primary school learners of English from diverse backgrounds in Germany. Overall, they found that learners had significantly higher post-test performance in vocabulary as well as higher positive emotions when the intervention included either plurilingual practices and/or other affective-experiential activities, and less negative emotions when the intervention included stimulated appreciation of plurilingualism and positive language attitudes. Frank & Papadopoulou (2024) studied adult L2 migrant learners learning the language of their host country, i.e., Greek and French. They found a higher degree of positive and a lower degree of negative emotions, as well as more positive attitudes towards a multilingual language teaching intervention targeting derivational morphology compared to a traditional intervention on the same topic.

Meanwhile, despite the extensive line of research on cooperative learning, limited research has been carried out in school settings. When it is done, it mostly focuses on the FL classroom and highlights that learners have positive attitudes towards collaboration (Calzada & García Mayo, 2020). Studies in school settings have found that group work is of great help for children from diverse backgrounds, promoting their relationships, social inclusion, and team incorporation (Baines et al., 2017; Borůvková & Emanovský, 2016). However, recent meta-analyses reveal (a) that collaborative interactions between students from different backgrounds were highly effective among younger children and less effective in adolescents (Ülger et al., 2018), and (b) that collaborative learning activities targeting migrant, low-income, and Roma children in eight European countries accounted for only 6% of all interventions conducted (Aguiar et al., 2019).

3. Collaborative writing

Drawing from the notion of collaborative learning and the socio-cultural theory of learning (Vygotsky, 1978), collaborative writing is a communicative activity between two or more students who compose an entire written text together from the beginning until the end (Storch, 2018; Howard, 2001, p. 54). All students of the group participate and collaborate throughout all stages, and they are all responsible for making decisions and producing the text (Storch, 2005; Storch & Wigglesworth, 2007; Wigglesworth & Storch, 2009). In this way, students develop their critical thinking, argumentation, and negotiation since they analyze and produce their own texts while learning from each other and expanding their perspectives (Storch & Wigglesworth, 2009).

Moreover, many studies found that collaborative writing helps learners in their language acquisition as well as writing skills (Donato, 1994; Swain & Lapkin, 2001). Specifically, learners seem to attend to the form more cautiously, which enhances more accurate language production (Philp, Adams, & Iwashita, 2013). Students also improve their vocabulary and text coherence (Talib & Cheung, 2017) and establish clarity in writing (Fong, 2012). It has also been found that pairs produced more accurate texts compared to individuals in FL classrooms (Basterrechea & García Mayo, 2013; Dobao & Blum, 2013; Nassaji & Tian, 2010; Wigglesworth & Storch, 2009 but see also Kuikken & Vedder, 2012).

Additionally, collaborative writing has been argued to work as scaffolding (Wood, Bruner, and Ross, 1976) among peers (Cazden, 1988). Hence, it can provide a social context for students to offer and receive support and feedback (Berkenkotter, 1984; Donato, 1994). Studies on the topic (i.e., Cho et al., 2006 Daiute & Dalton, 1993; Stanley, 1992; Storch, 2005; Zhu, 2001) revealed several types of peer scaffolds (i.e., collaborating, seeking information, using repetition, providing compliments or praises etc.).

Furthermore, the role of social group dynamics as well as of affect (i.e., emotions that are generated in collaborative learning), can greatly affect the successful outcome of the activity (Swain and Miccoli, 1994). Group activity towards a common goal has been argued to promote acceptance of differences (Johnson & Johnson, 1999), enjoyment, self-esteem, and confidence in speaking (Shehadeh, 2011). It has also been argued that it develops opportunities for peer support even among students from diverse/different backgrounds (Kagan, 1992). Crucially, heterogenous groups seem to show more benefits than homogenous groups (Dishon & O'Leary, 1984). This is attributed to the fact that individuals from different backgrounds tend to focus on different information in the discourse and exhibit different perspectives (Gardner,

1999). Thus, they can benefit from each other, embrace diversity, and reflect positive interdependence (Slavin, 1995). However, collaborative work may also entail conflict and disagreement among group members and particular patterns of behavior and/or relations which are not always conducive to learning (Storch, 2002) such as slacking or free-riding (Brooks & Ammons, 2003; Pieterse &Thompson, 2010), especially when students lack the necessary collaboration skills to contribute to the team (Oakley et al., 2007).

Importantly, from a socio-cultural perspective, interaction between peers during group work and emotions are tightly linked. Many scholars argue that learners are likely to be more successful in language learning when they are socially engaged, i.e., listening and providing feedback to each other, and drawing from one another's expertise and ideas (Moranski & Toth, 2016; Sato & Ballinger, 2016; Philp & Duchesne, 2016).

Another important factor that can influence group dynamics is language proficiency, which can, in turn, influence individuals' willingness to participate (Storch, 1998). Low proficiency learners may benefit more when paired with higher-level proficiency partners (Kim and McDonough, 2008; Leeser, 2004 but see Storch, 1998 for opposite results). On the other hand, other studies have found that proficiency is not always the determining factor in participants' post-test performance, and that other factors may play a crucial role, such as pair interactions (Watanabe & Swain, 2007; Storch, 2002).

4. Dictogloss

Dictogloss (Swain et al., 1985) can be considered a form of collaborative writing intervention. It draws from the Output hypothesis (Swain, 2000), which argues that grammar is learnt via comprehensible output and the systematic effort of students to produce it promotes acquisition and the automation of the grammatical forms. Dictogloss includes a sequence of main stages (Kuiken & Vedder, 2002; Wajnryb, 1990): students read or listen to a text while they also keep notes, then they reconstruct the text individually or in pairs/groups, and finally they compare their version with the original one. Dictogloss aims to provide students with opportunities to practice all language skills (Qin, 2008) and to teach grammar (Jacobs & Small, 2003; Nurdianingsih & Rahmawati, 2018; Olioumtsevits et al., 2023; Yolanda, 2019). The reconstruction stage of a dictogloss encourages negotiation and thus languaging (Swain, 2006), i.e., opportunity for noticing and resolving language problems that the learners may not have been able to resolve on their own (Pica, 1994). Significant grammar gains have been observed in both comprehension and production, with the gains remaining for long after instruction (e.g., Gorman & Ellis, 2019; Qin, 2008). In addition, as a group task, dictogloss can be influenced by group dynamics. Thus, students with strong personalities may prevent other students from participating, especially during the reconstruction stage, while more reserved students may be reluctant to discuss or correct the text with team members (Deveci & Ayish, 2018, p. 7).

Importantly, students' attitudes towards dictogloss as well as their interactions during text reconstruction and collaboration have also been explored. Gallego (2014) and Steward et al. (2014) found that adult foreign learners, who were university students, exhibited positive attitudes towards dictogloss. Deveci and Ayish (2018) also found that adult EFL students found dictogloss appealing, motivating, and empowering when employed as a group activity. Crucially, the main challenge that students exhibited was conflict and disagreement along with time management, which in turn caused more tension. Other challenges were recalling the details and anxiety about writing the correct form. Limited engagement in writing and/or lack of involvement were also reported as a source of dissatisfaction. Kanazawa (2017) and Ahmadian et al. (2015) also found that dictogloss increased adult EFL students' motivation and reduced anxiety, respectively. Importantly, EFL children and adolescents also exhibited a positive attitude towards dictogloss (Calzada & García Mayo, 2020; Shak, 2006) despite their (initial) unfamiliarity with the activity (Shak, 2006).

5. The present study

5.1. Research questions

In the present study, we explore learners' emotions and attitudes towards dictogloss, and more particularly, its collaborative aspect by focusing on the role of migrant background, educational level, and school setting.

RQ1: To what extent does dictogloss influence migrant students' positive and negative emotions? And is this influence modulated by the students' educational level and educational setting?

We expect higher positive and lower negative emotions if the intervention is appealing.

RQ2: To what extent does dictogloss influence students' positive and negative emotions within the mainstream classroom? And is this influence modulated by the students' educational level and background?

We expect higher positive and lower negative emotions if the intervention is appealing.

RQ3: How did students experience their collaboration in class? Is this experience mediated by the students' educational level, educational setting, and background?

We expect that if students have a positive collaboration experience, they will exhibit more positive emotions towards collaborating with each other and fewer collaboration difficulties.

Given the lack of prior research in these populations, no hypothesis was formed regarding the role of educational level, educational setting, and students' background.

5.2. Participants

One hundred and sixteen students participated in the present study (Table 1). The data were collected from 6 primary and junior high schools in the western part of Thessaloniki (Greece), including the outskirts. To explore the role of educational level, educational setting, and (non-)migrant background, there were six groups of

students presented in Table 1 along with their age range, school grade, and mean years of stay in Greece.

Table 1. Number and groups of participants by students' Background, Educational level, and Educational setting, along with age range, school grade, and mean years of stay in Greece (standard deviation (SD) in parentheses).

N of participants	Background	Educational level	Educational setting	Age range	School grade	Mean years of stay in Greece
18	migrant	primary school	mainstream class	11-12	6th grade	9.3 (SD=2.9)
19	non-migrant	primary school	mainstream class	11-12	6th grade	-
17	migrant	primary school	reception class	10-12	5th & 6th grade	5.1 (SD=0.8)
27	migrant	junior high school	mainstream class	13-17	2nd & 3rd grade	9.3 (SD=5)
17	non-migrant	junior high school	mainstream class	13-15	2nd & 3rd grade	-
18	migrant	junior high school	reception class	12-18	1st-3rd grade	5.1 (SD=2.8)

Migrant students' languages were: Albanian, Arabic, Armenian, Chinese, Georgian, English, Polish, Russian, Ukrainian, and Kurdish. Five migrant students in junior high school were older than the typical age for these grades. Students in the reception classes had approximately 5 years of mean stay in Greece and had a proficiency level between advanced A2 and intermediate B1. This information was reported by their teachers based on their placement tests at the beginning of the school year. In the mainstream classes, 85% of the students had a proficiency level between advanced B1 and C1 level, and approximately 9 years of mean stay. All non-migrant students were native speakers of Greek (L1) and were all born and raised in Greece.

5.3. Methods

5.3.1. Dictogloss: teaching intervention protocols and procedure.

Two teaching protocols were implemented. In the first one, the (non-)migrant groups in the mainstream classes of junior high school conducted two dictogloss activities about the life and action of Nelson Mandela. The first text targeted vocabulary and the second one indirect speech and the formation of indirect questions. The second protocol was conducted by the rest of the groups and included two dictogloss activities about two friends exchanging voice messages about a theater performance and a funny day in the park. These texts targeted the formation and meaning of verbal aspect in Greek, which is a vulnerable phenomenon in L2 Greek (Karpava et al., 2012; Tsimpli & Papadopoulou, 2009). Both teaching protocols were applied after discussion with the teachers regarding the needs of their class. The students first listened to the pre-recorded text presented at a natural pace and were asked to listen for comprehension. Then, they listened to the text two more times at a slower pace and were asked to note down key words/notes that would later help them reconstruct the missing parts of the text. Then, they had to complete these parts in groups of two to four members. Finally, they had to compare their version with the original text and make amendments, using a pen with a different colour[1].

5.3.2. Emotions questionnaire

To explore RQ1 and RQ2, students conducted an emotion questionnaire. The questionnaire was an adapted version of the emotion questionnaire in Franck and Papadopoulou (2024). Students were asked to rate the strength of the emotions they had experienced during the intervention. Thirteen emotions were tested, including both positive and negative ones. Furthermore, epistemic emotions (confusion, curiosity, excitement, frustration, and surprise) and a subgroup of achievement emotions (anger, enjoyment, despair, hope, shame, and pride) from the Achievement Emotions Questionnaire AEQ-S (Frenzel et al., 2009) were employed. Emotions that are considered both epistemic and achievement emotions (boredom and anxiety) were also included. Answers were given on a 5-point Likert scale (1 = "Hardly at all"; 5 = "Very strong"). The meaning of each emotion was orally explained by the researcher, and examples were given to make sure that the students understood each one of them. Detailed instructions were also given by the researcher for the completion of the questionnaire. Students chose between the Greek version of the questionnaire and the version in their own first language.

5.3.3. Open-ended questionnaire

To explore RQ3, an open-ended questionnaire was conducted, including three questions (Q1-3). Q1 asked students how they felt while collaborating with their team members. Q2 asked students what difficulties they faced as a team (at group level) during collaborating with each other, and what would have mitigated these difficulties. Q3 asked students what difficulties they faced personally (as individuals) during collaborating with each other, and what would have mitigated these difficulties.

5.4. Data analysis

For RQs 1-2, linear mixed effects models were conducted in R (Version 4.4.1; R Core Team, 2023) on the emotional ratings of students predicted by the fixed effects of Emotion type, Educational level, and Educational setting as well as their interaction (RQ1) and Emotion type, Educational level, and Background as well as their interaction (RQ2). The random part included a random effect for students and a random slope for Emotion type.

In RQ3, responses were analyzed qualitatively by

[1] Since we do not focus on the language learning gains of the teaching protocols but on their emotional impact on students and the collaboration among them, we do not present the protocols into more detail in terms of their language structures in the present study.

means of thematic analysis (Braun & Clarke, 2006). In this way, recurring themes and insights related to the students' emotions and perceptions towards collaboration in the dictogloss task were identified. Both authors collaboratively analyzed the responses on several occasions. Similar responses were grouped and manually identified to generate common themes. The research team refined the themes, and after reaching a consensus on their relevance, the main themes were identified:

(a) emotions towards collaboration (Q1): positive, negative, and neuter responses,

(b) group level (Q2) and individual level difficulties (Q3) towards collaboration: internal difficulties (i.e., due to language proficiency, due to memory demands, due to note taking), collaboration difficulties (i.e., disagreeing, not sharing the notes, covering the text with their body, not participating in the discussion), no difficulties, and difficulties related to the materials.

6. Results

In terms of RQ1, the results are illustrated in Figure 1. Model results (Table 2) revealed a main effect of Emotion type, indicating that students assigned significantly higher ratings to positive emotions compared to negative ones. The three-way interaction was also significant, reflecting that the significant effect of Emotion type depends on both educational level and educational setting.

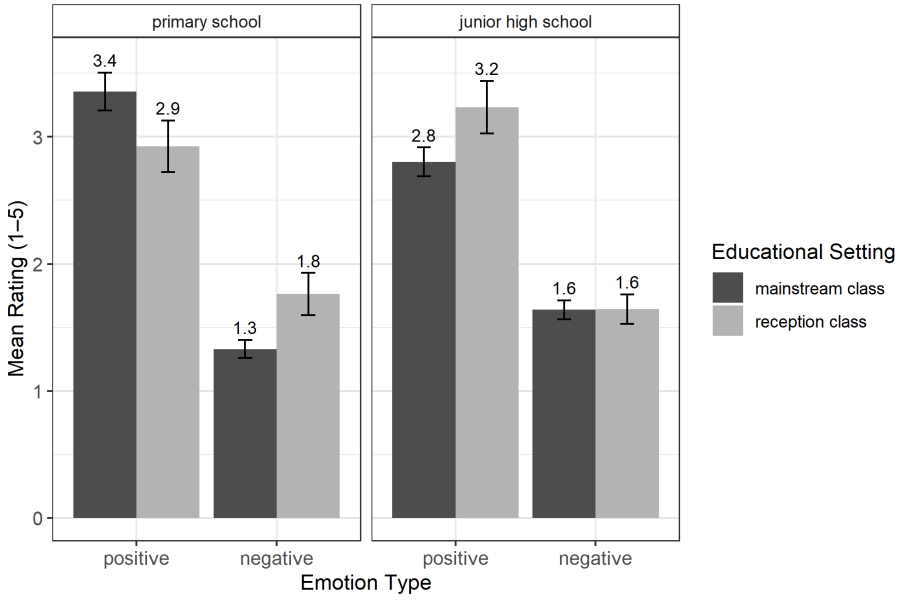


Figure 1. Emotional ratings by school level and educational setting.

Table 2. Model results for emotional ratings predicted by Emotion type, Educational level, and Educational setting

	b	se	t	p
Emotion type	1.501	0.152	9.86	**< .001**
Educational level	0.025	0.154	0.16	.872
Educational setting	-0.103	0.144	-0.72	.476
Emotion type * Educational level	0.247	0.305	0.81	.421
Emotion type * Educational setting	0.238	0.285	0.84	.406
Educational level * Educational setting	0.230	0.289	0.80	.427
Emotion type * Educational level * Educational setting	1.326	0.570	2.33	**.002**

Between-group pairwise comparisons (via emmeans package (Lenth, 2000) and adjusted p-values with Tukey correction) revealed that all groups assigned higher ratings to positive compared to negative emotions (all *p*-values< .05). Furthermore, when contrasting Educational setting (mainstream vs. reception class), there was a significant difference between the mainstream and reception class in primary school for the negative emotions (b= -0.439; se= 0.214; t= 2.05; *p*= .044), indicating that migrant students in the mainstream class of primary school assigned lower ratings to the negative emotions compared to migrant students in the reception class of primary school. The rest of the three comparisons were not significant (positive emotions of primary school students in the mainstream class vs. positive emotions of primary school students in the reception class: b= 0.462; se= 0.356; t= 1.30; *p*= .197, positive emotions of junior high school students in the mainstream class vs. positive emotions of junior high school students in the reception class: b= -0.431; se= 0.351; t= -1.23; *p*= .225, negative emotions of junior high school students in the mainstream class vs. negative emotions of junior high school students in the reception class: b= -0.006; se= 0.190; t= -0.030; *p*= .976).

When contrasting Educational level (primary vs.

junior high school), there were two marginally significant differences: migrant students in the mainstream class of primary school assigned marginally higher ratings for positive emotions compared to the migrant students in the mainstream class of junior high school (b= 0.595; se= 0.304; t= 1.95; p= .055) and marginally lower ratings for negative emotions compared to the migrant students in the mainstream class of junior high school (b= -0.315; se= 0.167; t= -1.89; p= .064). The rest of the two comparisons were not significant (positive emotions of primary school students in the reception class vs. positive emotions of junior high school students in the reception class: b= -0.298; se= 0.422; t= -0.707; p= .482, negative emotions of primary school students in the reception class vs. negative emotions of junior high school students in the reception class: b= 0.118; se= 0.236; t= 0.502; p= .618).

Furthermore, the role of Educational setting and Educational level was further explored for each emotion separately, following Franck & Papadopoulou (2024). The means are presented in Figure 2. The analysis revealed that there was an interaction between the two factors for the emotion of curiosity (b= -0.386; se= 0.165; t= -2.34; p= .023), reflecting that primary school students in reception classes were less curious than students in the mainstream class and the students at junior high school (p-values < .05). There was also a main effect of Educational setting for the emotion of confusion (b= -0.559; se= 0.171; t= -3.27; p= .002), reflecting that migrant students in reception classes felt more confused compared to migrant students in mainstream classes (p< .05). There was also a main effect of Educational level (b= 0.231; se= 0.110; t= 2.10; p= .004) such that junior high school students felt more shame than primary school students (p< .05).

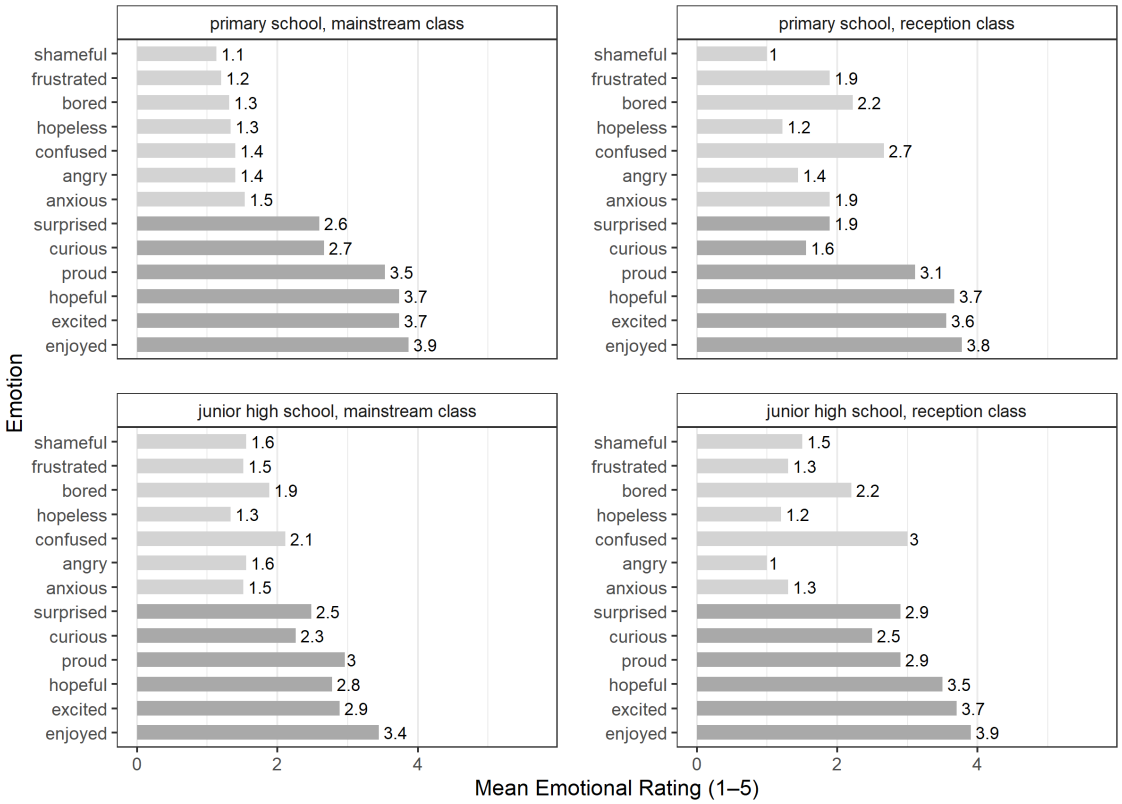


Figure 2. Mean emotional ratings per emotion by educational setting. Positive emotions are visualized with dark grey colour and negative emotions with light grey colour.

In terms of RQ2, the effects of Background, Educational level, and Emotion type were analyzed. The results are illustrated in Figure 3. The model results (Table 3) revealed a main effect of Emotion type, indicating that students assigned significantly higher ratings to positive emotions compared to negative emotions. The three-way interaction was also significant, reflecting that the significant effect of Emotion type depends on both Educational level and Background.

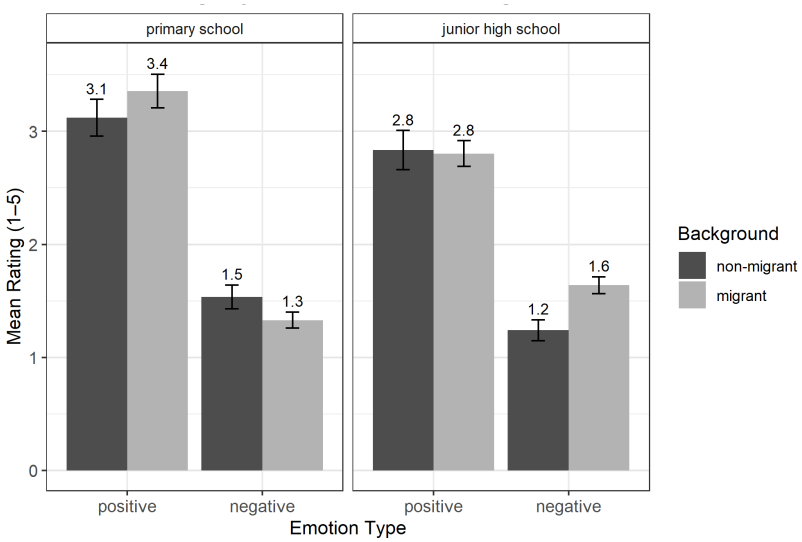


Figure 3. Mean emotional ratings by school level and background.

Table 3. Model results for emotional ratings predicted by Emotion type, students' Background, and Educational level.

	b	**se**	**t**	**p**
Emotion type	-0.795	0.044	-17.9	**< .001**
Educational level	-0.118	0.065	-1.82	.074
Background	-0.045	0.065	-0.69	.491
Emotion type * Educational level	0.107	0.044	2.407	.016
Emotion type * Background	0.002	0.044	0.06	.957
Educational level * Background	-0.040	0.066	-0.60	.549
Emotion type * Educational level * Background	-0.108	0.044	-2.44	**.015**

Pairwise comparisons revealed that all groups assigned higher ratings to positive emotions compared to negative emotions (all *p*-values< .05). Furthermore, when contrasting Background (migrant vs. non-migrant students), no significant differences were found (all p-values > .05), indicating that migrant students and non-migrant students within the same educational level (i.e., migrant vs. non-migrant students in primary school, migrant vs. non-migrant students in junior high school) assigned similar emotional ratings.

However, when contrasting Educational level (primary vs. junior high school), it was found that migrant students in primary school exhibited significantly stronger positive emotions compared to migrant students in junior high school (positive emotions of migrant students in primary school vs. positive emotions of migrant students in junior high school: b= 0.587; se= 0.187; t= 3.14; *p*= .002), while non-migrant students did not (positive emotions of non-migrant students in primary school vs. positive emotions of non-migrant students in junior high school: b= 0.313; se= 0.266; t= 1.18; *p*= .241). The rest of the comparisons were not significant (negative emotions of migrant students in primary school vs. negative emotions of migrant students in junior high school: b= -0.272; se= 0.178; t= -1.53; *p*= .128; negative emotions of non-migrant students in primary school vs. negative emotions of non-migrant students in junior high school: b= 0.318; se= 0.251; t= 1.27; *p*= .207).

Furthermore, the role of Background and Educational level in the mainstream class was further explored for each emotion separately. The means are presented in Figure 4. The analysis revealed that there was a main effect of Educational level for the emotion of hope (b= 0.596; se= 0.206; t= 2.89; *p*= .005), reflecting that primary school students felt more hopeful than junior high school students. There was also a significant interaction between Background and Educational level (b= 0.350; se= 0.158; t= 2.22; *p*= .031) reflecting that non-migrant students in primary school felt more confused compared to non-migrant students in junior high school and compared to migrant students in primary and junior high school (all p-values < .05).

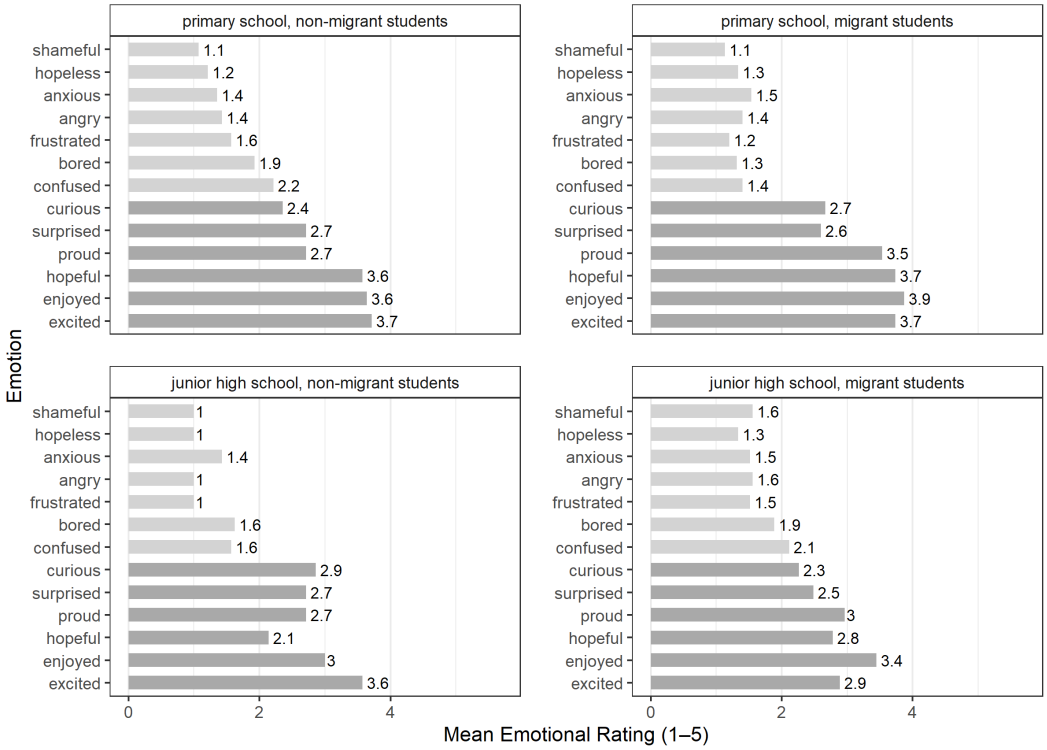


Figure 4. Mean emotional rating per emotion by Educational level and students' Background. Positive emotions are visualized with dark grey colour and negative emotions with light grey colour.

In terms of RQ3, the students' questionnaire responses (Q1-Q3) are presented. Below, we report students' responses by educational level, educational setting, and educational background. We first report the percentages of students' responses in each Question (Table 4) based on the thematic analysis. Then, we report the most representative responses of each group.

Table 4 reflects that there is a clear distinction between primary and junior high school students, with primary school students experiencing more positive emotions compared to junior high school students (primary school: migrant students in mainstream class 100%, non-migrant students in mainstream class 86%, migrant students in reception classes 88%; junior high school: migrant students in mainstream class 70%, non-migrant students in mainstream class 60%, migrant students in reception classes 65%) as well as more limited collaboration difficulties at group level (primary school: migrant students in mainstream class 5%, non-migrant students in mainstream class 27%, migrant students in reception classes 23%; junior high school: migrant students in mainstream class 38%, non-migrant students in mainstream class 55%, migrant students in reception classes 50%) and individually (primary school: migrant students in mainstream class 5%, non-migrant students in mainstream class 13%, migrant students in reception classes 0%; junior high school: migrant students in mainstream class 37%, non-migrant students in mainstream class 0%, migrant students in reception classes 10%).

Furthermore, primary school students experienced more internal difficulties compared to junior high school students at group level (primary school: migrant students in mainstream class 50%, non-migrant students in mainstream class 32%, migrant students in reception classes 30%; junior high school: migrant students in mainstream class 22%, non-migrant students in mainstream class 0%, migrant students in reception classes 20%) and individually (primary school: migrant students in mainstream class 45%, non-migrant students in mainstream class 60%, migrant students in reception classes 60%; junior high school: migrant students in mainstream class 22%, non-migrant students in mainstream class 0%, migrant students in reception classes 20%).

Thus, the present findings reflect that the increased positive emotions of primary school students are associated with having better collaboration experiences compared to junior high school students, even though the former faced more internal difficulties (i.e., difficulties in note taking, in remembering parts of the text, in time management etc.). The above finding is further reflected in the group of migrant students in the mainstream class, given that all students experienced positive emotions (100%) for their collaboration and exhibited scarce collaboration difficulties (5%). On the other hand, non-migrant students in mainstream class were the group who experienced the weakest positive emotions (60%) and the most frequent collaboration difficulties (55%). Difficulties with materials were reported less than 15% across groups.

Table 4. Percentage of responses per question and theme by students' Background, Educational level and Educational setting

	Primary school			Junior high school		
	Migrant students in mainstream class	Non-migrant students in mainstream class	Migrant students in reception class	Migrant students in mainstream class	Non-migrant students in mainstream class	Migrant students in reception class
Q1 Emotions towards collaboration						
Positive	100	86	88	70	60	65
Neutral	0	14	12	10	20	15
Negative	0	0	0	20	20	20
Q2 Difficulties as a group						
Collaboration	5	27	23	38	55	50
Internal	50	32	30	22	0	20
Materials	0	14	0	3	0	0
No difficulties	45	27	47	37	45	30
Q3 Individual difficulties						
Collaboration	5	13	0	37	0	10
Internal	45	60	60	22	0	20
Materials	0	0	13	3	0	20
No difficulties	55	27	27	38	100	50

All migrant students in the primary classroom reported positive emotions (see (1) and (2) below) for collaborating with their classmates and they mainly expressed joy and happiness (Q1). Crucially, no student reported negative emotions. Most of the difficulties they faced with their teams (Q2) were attributed to internal reasons, as noted in (3) (i.e., spelling errors/orthography, taking notes, and remembering words when filling out the gaps). However, there were two students who reported collaboration difficulties (4). At an individual level (Q3), half of the students reported that they experienced difficulties related to internal reasons see for example (5). The students reported that (a) paying attention to the pictures of the story, and (b) guessing the missing word/phrase based on the preceding/following part of the text were helpful strategies.

(1) "I felt happy because we did the activity together with my friend."

(2) "I liked it a lot. We worked in teams in nice and smart ways."

(3) "Our team needed a bit more time for the notes."

(4) "My classmate wrote fast, and I could not see where she was when writing. Thus, I could not do much."

(5) "I had some difficulties when I had to listen and write at the same time."

Non-migrant students in the mainstream classroom of primary school mainly reported positive emotions (Q1) for collaborating with their classmates (6) except for three students who gave neutral answers (7). Nobody gave a negative response. In Q2, the students reported some internal difficulties (8), difficulties due to collaboration (9), and no difficulties with their team (10). At an individual level (Q3), half of the students reported internal difficulties (11). There were few difficulties due to collaboration (12) and materials (13). In terms of their recommendations (i.e., what would have mitigated their individual or group difficulties), most students reported that they would like to learn how to collaborate more effectively, and to practice their skills in relation to dictogloss (note taking while listening, guessing what is missing based on the pictures etc.).

(6) "I felt very nice because I had a very helpful partner, and we collaborated nicely."

(7) "It was ok."

(8) "We had difficulties in finding/remembering the missing word because we did not take notes, but the pictures were very helpful."

(9) "My classmate wasn't listening to what I was saying at all. I wish he was!"

(10) "Nothing was difficult for us because we were really good and worked together."

(11) "The main difficulty I had was to fill out the gaps fast and accurately. Practicing this further would help me."

(12) "The partner I had."

(13) "The first text was difficult. The second one was easier."

Most of the migrant students in the reception class of primary school reported positive emotions (Q1) for collaborating with their classmates (14) and they mainly expressed joy and happiness. Three students gave more neutral responses (15). Nobody gave a negative response. Overall, half of the students reported no collaboration difficulties (Q2). Most of the difficulties were internal (16). However, there were also some difficulties due to collaboration (17). At an individual level (Q3), half of the students reported that they experienced some internal difficulties, which were mostly resolved via collaboration (18). Two students reported difficulties with materials (19).

(14) "Perfect! I really enjoyed that I was in the same

team together with my friends and I want to do that more often."

(15) "It was ok, but I wanted to do this by myself too."

(16) "We did not have any. When I did not remember something, my friend helped me, and I also helped him."

(17) "My partner wanted to complete some gaps herself, but she couldn't. I helped her and from that point all went well."

(18) "I forgot some words, but my team knew them and helped me a lot.

(19) "It was difficult, the text was fast."

Most of the migrant students in the mainstream class of junior high school reported positive emotions, feeling joy and closeness due to their collaboration (20). However, neutral and negative responses about collaboration were also present (21-23). The negative emotions were anger, dissatisfaction, and anxiety. In Q2, approximately half of the students reported that they had difficulties in collaboration (24, 25). Six students also reported internal difficulties and difficulties with the materials (26). At an individual level (Q3), half of the students reported that they experienced difficulties because of inefficient/insufficient collaboration (28-29). There were also some internal difficulties (30). In terms of recommendations, practicing their collaboration skills was reported by one student. For internal difficulties, collaboration with friends, translation, and listening to the text one more time were recommended.

(20) "I felt that I came closer to my classmates since we worked together, and we made it."

(21) "Not good, not bad. If I was in another team, it would have been better."

(22) "I got angry because of the other members of the team. They believe that I was the problem. I did not enjoy it at all."

(23) "In the beginning I was feeling comfortable and nice full of hope for the win. Then, everything changed, and this is why we did not win the first time."

(24) "We were shouting at each other and disagreed. Nothing helped. I did everything myself."

(25) "We could not collaborate well because two members of the team did not know Greek well."

(26) "The listening was fast; it should have been slower."

(27) "Personally, I had difficulties working with these members."

(28) "I could not think and concentrate because of my team. It would have been better if I was in another team."

(29) "I got stressed because of the members of my team."

(30) "I did not know some words. I needed translation from the phone or my friends."

Most of the non-migrant students in the mainstream class of junior high school reported positive emotions (31), feeling joy and satisfaction due to their collaboration (Q1). However, neutral and negative responses (32, 33) were also present. The negative emotion was dissatisfaction. Half of the students reported that the difficulties they experienced with their team (Q2) were attributed to collaboration (34). No individual difficulties were reported (Q3). With respect to their recommendations, they would like to have collaborated more effectively.

(31) "It was nice that we worked in teams, much better than if we had to work individually."

(32) "It was ok, but I prefer to work alone."

(33) "The team was not very responsive and active, and this is why did not performed well."

(34) "Difficulties in understanding each other. It would have been better if we had collaborated more effectively."

Most of the migrant students in the reception class of junior high school reported positive emotions (Q1), feeling joy and satisfaction due to their collaboration (35). However, neutral and negative responses about collaboration were also present (36). The negative emotions were disappointment and dissatisfaction. Half of the students reported difficulties due to collaboration (Q2) (37-38). Some limited internal difficulties were also reported (39). At an individual level (Q3), half of the students reported no difficulties. When difficulties were reported, they were attributed to internal reasons (40) or to materials (41). In terms of recommendations, one student reported that more time would have been beneficial due to communication issues, and another student reported that working individually would have been a solution. Finally, only one student reported that practicing their collaboration skills would be important.

(35) "Great! I really enjoyed it! My partner was very helpful."

(36) "Not so good. My partner could not help, and he was very slow."

(37) "My partner did everything herself so I could not do much."

(38) "I did not trust my partner and I did everything by myself."

(39) "My team needed more time to fill out the gaps."

(40) "I had to think a bit more about putting the verb in the correct form".

(41) "It was difficult to listen and write notes".

7. Discussion

The present study explored students' emotions and collaboration experience when performing collaborative writing by means of dictogloss activities. The study tested the role of educational level, educational setting, and student background on the emotional impact of dictogloss as well as students' perceptions about their collaboration.

Three RQs were addressed. RQ1 and RQ2 focused on the emotional impact of dictogloss on students. RQ1 asked whether educational level (primary school vs. junior high school), educational setting (mainstream class vs. reception class), and type of emotion (positive vs. negative emotions) influenced students' emotional ratings during dictogloss. RQ2 asked whether educational level (primary school vs. junior high school), type of emotion (positive vs. negative), and background (non-migrant vs. migrant

students) influenced students' emotional ratings in the mainstream class. RQ3 explored students' perceptions about their collaboration during the task.

The results for RQ1 showed that migrant students experienced more positive than negative emotions during dictogloss, and that educational level and educational setting differentially affected their emotions. Specifically, migrant students in primary school experienced weaker negative emotions in the mainstream class compared to the reception class, while migrant students in junior high school exhibited similar strength of negative emotions in the mainstream and the reception class. Furthermore, migrant students in the mainstream class of primary school experienced marginally stronger positive emotions and weaker negative emotions compared to the migrant students in the mainstream class of junior high school.

The fact that all students overall exhibited higher positive and lower negative emotions, even though they were all unfamiliar with the task, is in line with the findings of RQ3 (see below the discussion for Q1). The positive emotional impact of dictogloss has also been found in previous studies with FL learners (Ahmadian et al., 2015; Deveci & Ayish, 2018; Gallego, 2014; Kanazawa, 2017).

In addition, the weaker negative emotions found for the primary school migrant students in the mainstream class compared to the reception class could be attributed to the high sense of belonging and happiness that the former group has developed within the mainstream class. This is also enhanced by the fact that this was the only group who reported only positive feelings for their collaboration and had almost no collaboration difficulties (see below, RQ3). The present findings are in line with findings demonstrating that migrant students are better supported long-term in the mainstream class where the appropriate scaffolding promotes better integration and peer interaction (Hunt, 2024).

However, this group of students also exhibited (marginally) higher positive and lower negative emotions compared to the junior high school migrant students in the mainstream class. This marginal difference is present even though junior high school students also attended schools where inclusivity, interculturalism and peer collaboration were highly valued and embraced. This finding reflects that the mainstream class may not have the same positive impact on migrant adolescents and migrant children. Migrant adolescents in the mainstream class have been reported to feel anxiety and discomfort when they cannot express themselves (Horgan et al., 2022). Furthermore, adolescents face greater academic and social challenges in the mainstream class related to language acquisition, curriculum complexity and social integration, in addition to the commonly observed academic gaps in their formal education (Ahad & Benton, 2018; Horgan et al., 2022; Wong and Schweitzer, 2017). On the other hand, the mainstream class in primary schools tend to be more inclusive, and students perform group activities more often, which was also confirmed by the teachers. Thus, migrant children are more familiar and feel more at ease with group work in the mainstream class compared to the adolescent migrant learners.

When looking at each individual emotion, we found that primary school students in the reception class exhibited weaker curiosity compared to the rest of the groups. Curiosity is an epistemic emotion and for this reason, it may be less pronounced in migrant children with lower levels of proficiency. Perhaps due to the internal difficulties they faced (see discussion in RQ3), they were less likely to feel as curious as the other groups. Furthermore, students in the reception classes exhibited more confusion compared to students in the mainstream classes. This is not surprising given that dictogloss requires intensive cognitive processing and active listening skills, especially among L2 learners (Deveci & Ayish, 2018). Importantly, based on students' responses to the open-ended questions (see RQ3 below), this confusion seems to have stemmed from their unfamiliarity with the task and the cognitive demands of dictogloss rather than from not understanding the instructions. Even more, junior high school students felt more shame compared to primary school students, which relates to the fact that adolescents tend to exhibit more shyness when they are asked to do group work (see Bowker et al., 2023 for discussion on shyness in adolescence).

The results for RQ2 showed that both migrant and non-migrant students in the mainstream class experienced more positive than negative emotions during dictogloss, reflecting that the activity has positive emotional impact on all students of the mainstream class. However, migrant students experienced stronger positive emotions in primary school compared to migrant students in junior high school, as also found in RQ1 (see the discussion above). Furthermore, non-migrant students felt more confused than migrant students in primary school, and less confused than migrant students in junior high school. This interaction could be attributed to the fact that migrant students in primary schools, due to their flexibility to adapt into new contexts, follow new instructions and tasks more easily. On the other hand, in junior high school, migrant students in the mainstream class may find the content of the activity academically more challenging compared to the non-migrant students of the mainstream class. In addition, primary school students felt more hopeful compared to junior high school students, which also aligns with the characteristics of adolescents (i.e., they tend to be less optimistic) (Bowers & Powers, 2023; Long et al., 2024).

The results for RQ3 showed that students exhibited mainly positive emotions across the different groups, although not to the same extent. This is in line with the emotional ratings of the emotional questionnaire (RQ1-2). Furthermore, the group which exhibited only positive emotions when asked about their feelings towards collaboration (Q1) was the group of migrant students in the mainstream class of primary school. This is in line with the students' emotional ratings in the emotional questionnaire (see discussion in RQ1). Furthermore, primary school students overall did not report negative emotions as opposed to junior high school students, whose negative emotions mostly attributed to collaboration difficulties.

In Q2 of the open-ended questionnaire, junior high school students reported primarily difficulties in collaboration, while primary school students reported mostly internal difficulties. This also aligns with the responses in Q1. Thus, primary school students in the mainstream class who reported only positive emotions,

also reported the lowest extent of collaboration difficulties (5%). Furthermore, migrant students in mainstream classes seem to report better collaboration compared to migrant students in the reception classes and compared to non-migrant students in the mainstream classes. However, they mostly experience internal difficulties. Additionally, issues with materials were scarce, reflecting that they found them appropriate.

In Q3, migrant students in the mainstream class of junior high school exhibited an increased percentage of collaboration difficulties when reporting their difficulties at an individual level. Importantly, this is the group in which conflicts and intense disagreement occurred during collaboration.

With respect to the students' suggestions of mitigating the difficulties they experienced (either at group level or individually) throughout their collaboration, primary school students' responses reflected their wish and hope for better collaboration in the future and willingness to further practice this skill. Furthermore, in primary school, a series of internal difficulties were resolved via good collaboration since students reported that they helped each other when group and/or individual internal difficulties arose, reflecting solidarity, respect, and accountability. In junior high school, students mostly reported suggestions that are not constructive (i.e., change teams, stop collaborating, working individually). There was a limited number of students which acknowledged that further practice and knowledge is required to collaborate more efficiently.

The different collaboration patterns between the two educational levels can be explained from various perspectives. Primary school students may be more familiar with group work and practice their collaboration skills more often than junior high school students. On the other hand, the curriculum in Greek junior high schools is more oriented to individual written evaluation and less time, in turn, is devoted to group work, collaboration, negotiation. This is reflected in junior high school students' responses which were more focused on performance and successful completion of the dictogloss compared to primary school students. Furthermore, adolescents and children have been found to exhibit different perceptions about group work accompanied by different cognitive and social strategies (Bereiter & Scardamalia, 1987) employed to perform the group activity (Leman, 2015). Specifically, children have been found to view group work as a source of information, while adolescents as an opportunity to construct knowledge together. This could explain why adolescents emphasized on collaboration (difficulties) given the more "socialized" view of group collaboration (Leman et al., 2015, p. 819).

In terms of language proficiency and years of stay in the host-country, we observe no across-the-board difference between migrant and non-migrant students or between migrant students in mainstream classes and in reception classes. However, we found that migrant students in the reception classes experienced more confusion compared to migrant students in the mainstream class, which we attribute to their lower proficiency and years of stay in the host country along with the cognitive demands of the task.

7.1. Implications for teaching practice

The present study highlights the overall positive emotional impact of dictogloss on different groups of (non-)migrant students in formal educational settings and enhances its application within the mainstream and reception classroom.

However, the study shows that students also exhibited certain difficulties, mainly due to the cognitive and language demands of the activity as well as collaboration difficulties that arose mainly in junior high school. This is crucial for teaching practice given that these students were unfamiliar with this activity. Thus, the present study indicates that further practice is needed so that learners maximally benefit from dictogloss. The study also suggests that teachers should not assume that students already know how to work together efficiently (Johnson & Johnson, 1999). Specifically, teachers may need to explicitly explain the value of collaborative learning (Nunan, 1988) to their students to increase their awareness about group work. Group work is not always efficient and productive, and teachers should help learners further develop their decision making, trust-building and conflict management skills (Deveci & Ayish, 2018; Johnson & Johnson, 1999). Finally, teachers need to consider age differences accompanied by different social and cognitive competencies and their impact on group work.

Funding and ethics

This project has received funding from the European Union's AMIF program under the grant agreement No 101141078. The students' parents gave their written informed consent, and students gave their oral informed consent according to the Declaration of Helsinki. This study was approved by the Research Ethics Committee of the Aristotle University of Thessaloniki (91298/2025].

Dr. Anastasia Paspali (Corresponding Author) is a post-doctoral researcher at the Aristotle University of Thessaloniki. She is a linguist focusing on Psycholinguistics, Language teaching and language acquisition, bilingualism, language processing, first and second language teaching as well as language disorders. She works on the EU-funded project ACTIN (Act and Connect for Integration: Language Learning and Cultural Awareness). Her PhD (2019, Summa cum Laude) from the Humboldt University of Berlin is on the acquisition and processing of gender agreement and agreement attraction in Greek monolinguals and heritage bilingual speakers. Her post-doctoral project *"Acquisition and Processing of Voice: is non-active morphology a clinical marker?"* received funding from the German research foundation (DFG 2021-2024, awarded to particularly qualified early-career postdoctoral researchers). She has worked as a lecturer and post-doctoral researcher at the Humboldt University of Berlin, the University of Konstanz, University of Crete, Aristotle University of Thessaloniki, and New-York College Thessaloniki. She has published articles in international peer-reviewed journals (e.g., Language Acquisition, Applied Psycholinguistics, Frontiers in Psychology, Glossa etc.).
Email: paspalia@lit.auth.gr

Despina Papadopoulou is Professor of Linguistics at the Aristotle University of Thessaloniki and the Director of the School of Modern Greek Language. She is the principal investigator of the EU-funded project ACTIN (Act and Connect for Integration: Language Learning and Cultural Awareness). Her research interests lie in the areas of bilingual/multilingual development, second/foreign language acquisition and processing, and second/foreign language teaching. Over the last ten years, a considerable amount of her research and teaching has been devoted to migrant and refugee education, with particular emphasis on L2 learning and teaching. She participated in several research programmes on bilingual acquisition, language processing and second language acquisition, while recently she has been involved in MultiMind, a large-scale EU funded project on multilingualism that investigates among others second language acquisition and teaching in migrant and refugee children. She has designed and implemented a series of training programmes for educators who work with migrant and refugee populations in Greece, which are funded by UNICEF (https://www.teach4integration.gr).
E-mail: depapa@lit.auth.gr

References

Aguiar, C., Silva, C. S., Guerra, R., Rodrigues, R. B., Ribeiro, L. A., ... Pastori, G. (2019). Early interventions tackling inequalities experienced by immigrant, low-income, and Roma children in 8 European countries: A critical overview. *European Early Childhood Education Research Journal*, *28*(1), 58–76. https://doi.org/10.1080/1350293X.2020.1707363

Ahad, A. & Benton, M. (2018). *Mainstreaming 2.0: How Europe's education systems can boost migrant inclusion*. Migration Policy Institute Europe.

Ahmadian, M., Amerian, M., & Lavasani, E. (2015). The effect of the dicto-gloss as a cooperative learning technique on EFL learners' self-efficacy in writing. *Journal of Language Teaching and Research*, *6*(6), 1357. https://doi.org/10.17507/jltr.0606.25

Baines, E., Blatchford, P., & Kutnick, P. (2008). *Promoting effective group work in the primary classroom: A handbook for teachers and practitioners*. http://ci.nii.ac.jp/ncid/BA9124816X

Basterrechea M., García Mayo M.P. (2013). Language-related episodes (LREs) during collaborative tasks: a comparison of CLIL and EFL learners. In McDonough K., Mackey A. (Eds.), *Second language interaction in diverse educational contexts* (pp. 25–43). John Benjamins.

Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Lawrence Erlbaum Associates.

Berkenkotter, C. (1984). Student writers and their sense of authority over texts. *College Composition and Communication*, *35*(3), 312. https://doi.org/10.2307/357459

Borůvková, R., & Emanovský, P. (2016). Small group learning methods and their effect on learners' relationships. *Problems of Education in the 21st Century*, *70*(1), 45–58. https://doi.org/10.33225/pec/16.70.45

Bowker, J. C., Richard, C. L., Stotsky, M. V., Weingarten, J. P., & Shafik, M. I. (2023). Understanding shyness and psychosocial difficulties during early adolescence: the role of friend shyness and self-silencing. *Personality and Individual Differences, 209,* 1–7. https://doi.org/10.1016/j.paid.2023.112209

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

Brooks, C. M., & Ammons, J. L. (2003). Free riding in group projects and the effects of timing, frequency, and specificity of criteria in peer assessments. *Journal of Education for Business*, *78*(5), 268–272. https://doi.org/10.1080/08832320309598613

Busse, V., McLaren, L. M., & Dahm, A. (2021). Responding to migration-related diversity in the classroom: a comparison of diversity-sensitive approaches to stimulate word acquisition in early FL teaching. *Journal of Multilingual and Multicultural Development*, *45*(5), 1596–1615. https://doi.org/10.1080/01434632.2021.2005611

Busse, V., Cenoz, J., Dalmann, N., & Rogge, F. (2020). Addressing linguistic diversity in the language classroom in a resource-oriented way: an intervention study with primary school children. Language Learning, 70(2), 382–419 https://doi.org/10.1111/lang.12382

Calzada, A., & García Mayo, M. del P. (2020). Child learners' reflections about EFL grammar in a collaborative writing task: When form is not at odds with communication. *Language Awareness*, *30*(1), 1–16. https://doi.org/10.1080/09658416.2020.1751178

Cazden, C. B. (1988). *Classroom discourse: The language of teaching and learning*. http://ci.nii.ac.jp/ncid/BA12142720

Cho, K., Schunn, C. D., & Chamey, D. (2006). Commenting on writing typology and perceived helpfulness of comments from novice peer reviewers and subject matter experts. *Written Communication, 23*(3), 260–294. https://doi.org/10.1177/0741088306289261

Daiute, C., & Dalton, B. (1993). Collaboration between children learning to write: Can novice be masters? *Cognition and Instruction, 10*(4), 281–333.

Deveci, T. & Ayish, N. (2018). Effects of dictogloss as a collaborative writing activity on students' perceptions and language development. *The Journal of Limitless Education and Research*, 4–19. https://www.researchgate.net/publication/328900786_Effects_of_dictogloss_as_a_collaborative_writing_activity_on_students'_perceptions_and_language_development

Dishon, D. and O'Leary, P. W. (1984). *A guidebook for cooperative learning: A technique for creating more effective schools*. Learning Publications.

Donato, R. (1994). Collective scaffolding in second language learning. In J. P. Lantolf, & G. Appel (Eds.), *Vygostkian approaches to second language research*

(pp. 33–56). Ablex.
Dobao, A. F. and Blum, A. (2013). CW in pairs and small groups: Learners' attitudes and perceptions. *System*, *41*(2), 365–378.
Ferguson-Patrick, K. (2020). Cooperative learning in Swedish classrooms: Engagement and relationships as a focus for culturally diverse students. *Education Sciences*, *10*(11), 312. https://doi.org/10.3390/educsci10110312
Fong, L. S. (2012). Benefits of collaborative writing for ESL advanced diploma students in the production of reports. *US-China Education Review, 4*, 396–407.
Franck, J., & Papadopoulou, D. (2024). Pedagogical translanguaging in L2 teaching for adult migrants: Assessing feasibility and emotional impact. *Education Sciences*, *14*(12), 1308. https://doi.org/10.3390/educsci14121308
Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. E. (2009). Emotional transmission in the classroom: exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology, 101*(3), 705–716. https://doi.org/10.1037/a0014695
Gallego, M. (2014). Second language learners' reflections on the effectiveness of dictogloss: A multi-sectional, multi-level analysis. *Studies in Second Language Learning and Teaching, 4*(1), 33–50.
Gardner, H. (1999). *Intelligence reframed: Multiple intelligences for the 21st century*. Basic Books.
Göbel, K., and Frankemölle, B. (2020). Interkulturalität und Wohlbefinden im Schulkontext [Interculturalism and well-Being in the school context]. In T. Ringeisen, P. Genkova, and F. T. L. Leong (Eds.) *Handbuch Stress und Kultur: Interkulturelle und kulturvergleichende Perspektiven* (pp. 1–17). Wiesbaden: Springer. https://doi.org/10.1007/978-3-658-27825-0_30-1.
Gorman, M., & Ellis, R. (2019). The relative effects of metalinguistic explanation and direct written corrective feedback on children's grammatical accuracy in new writing. *Language Teaching for Young Learners*, *1,* 57–81. https://doi.org/10.1075/ltyl.00005.gor
Grief, S. (2007). *Collaborative writing*. The National Research and Development Centre for Adult Literacy and Numeracy. https://unesdoc.unesco.org/ark:/48223/pf0000204476
Henschel, S., Heppt, B., Weirich, S., Edele, A., Schipolowski, S., & Stanat, P. (2019). Zuwanderungsbezogene Disparitäten. In P. Stanar, S. Schipolowski, N. Mahler, S. Weirich, &S. Henschel (Eds.), *IQB-Bildungstrend* (pp. 295–336). Waxmann.
Horgan, D., Martin, S., O'Riordan, J., & Maier, R. (2022). Supporting languages: the socio-educational integration of migrant and refugee children and young people. *Children & Society*, *36*, 369–385. https://doi.org/10.1111/chso.12525
Howard, R. M. (2001). Collaborative pedagogy. In G. Tate, A. Rupiper, & K. Schick (Eds.), *A guide to composition pedagogies* (pp. 54–70). Oxford University Press.
Hunt, L. (2024). Young refugees' participation in post-compulsory education: mapping policies, challenges and ways forward in mainland Greece. In K. Sobczak-Szelc, M. Pachocka & J. Szałańska (Eds.), *The Integration of refugees in the education and labour markets: between inclusion and exclusion practices* (pp. 48–61). Routledge.
Jacobs, G.M., & Small, J.J. (2003). Combining dictogloss and cooperative learning to promote language learning. *The Reading Matrix: an International Online Journal, 3*.
Johnson, D.W. and Johnson, R. T. (1999). What makes cooperative learning work. In Kluge, D. McGuire, S., Johnson, D. and Johnson, R. (Eds.), *Cooperative Learning* (pp. 23–36). JALT Applied Materials.
Kagan, D.M. (1992) Implications of research on teacher belief. *Educational Psychologist*, *27,* 65–90. http://dx.doi.org/10.1207/s15326985ep2701_6
Kanazawa, N. (2017). *A study of the effectiveness of dictogloss on English grammar and motivation for Japanese junior college students in English*. Retrieved from https://komajo.repo.nii.ac.jp/?action=repository_uri&item_id=1242&file_id=22&file_no=1
Karpava, S., Grohmann, K. K., & Fokianos, K. (2012). Aspect in the L2 and L3 acquisition of Greek. In *Second language learning and teaching* (pp. 41–62). https://doi.org/10.1007/978-3-642-29557-7_3
Kim, Y., & McDonough, K. (2008). The effect of interlocutor proficiency on the collaborative dialogue between Korean as a second language learners. *Language Teaching Research, 12(2),* 211–234. https://doi.org/10.1177/1362168807086288
Kuiken, F., & Vedder, I. (2012). Speaking and writing tasks and their effects on second language performance. In S.M. Gass, & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 364–377). Routledge. https://doi.org/10.4324/9780203808184
Leeser, M. J. (2004). Learner proficiency and focus on form during collaborative dialogue. *Language Teaching Research*, *8*(1), 55–81. https://doi.org/10.1191/1362168804lr134oa
Leman P. J. (2015). How do groups work? Age differences in performance and the social outcomes of peer collaboration. *Cognitive science*, *39*(4), 804–820. https://doi.org/10.1111/cogs.12172
Lenth, R.V. (2020). *Emmeans: estimated marginal means. Aka least-squares means R Package Version 1.5.3*.
Long, K. N., Wilkinson, R., Cowden, R. G., Chen, Y., & VanderWeele, T. J. (2024). Hope in adolescence and subsequent health and well-being in adulthood: An outcome-wide longitudinal study. *Social Science & Medicine*, *347*, 116704. https://doi.org/10.1016/j.socscimed.2024.116704
Moranski, K. & Kim, F. (2016). Flipping lessons in a multi-section Spanish course: implications for assigning explicit grammar instruction outside of the classroom. *The Modern Language Journal*, *100*,

830–852. https://doi.org/10.1111/modl.12366
Nassaji, H., & Tian, N. J. (2010). Collaborative and individual output tasks and their effects on learning English phrasal verbs. *Language Teaching Research*, *14*(4), 397–419. https://doi.org/10.1177/1362168810375364
Nunan, D. (1988). *The learner-centred Curriculum: A study in second language teaching*. Cambridge University Press. https://doi.org/10.1017/cbo9781139524506
Nurdianingsih, F., & Rahmawati, O.I. (2018). Running dictation as an effective technique on the teaching writing skill. *ELLiC Proceedings*, *2*, 127–131.
Oakley, B. A., Hanna, D. M., Kuzmyn, Z., & Felder, R. M. (2007). Best practices involving teamwork in the classroom: Results from a survey of 6435 engineering student respondents. *IEEE Transactions on Education, 50*(3), 266–272.
OECD. (2015). *Immigrant students at school: easing the journey towards integration (OECD Reviews of Migrant Education)*. OECD Publishing. https://doi.org/10.1787/9789264249509-en
Olioumtsevits, K., Papadopoulou, D., & Marinis, T. (2023). Second language grammar learning in refugee children. *Pedagogical Linguistics, 4*(1), 50–76. https://doi.org/10.1075/pl.21012.oli.
Pieterse, V., & Thompson, L. (2010). Academic alignment to reduce the presence of social loafers and 'diligent isolates' in student teams. *Teaching in Higher Education, 15*(4), 355–367. http://dx.doi.org/10.1080/13562517.2010.493346
Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics, 36,* 50–72. https://doi.org/10.1017/s0267190515000094
Philp, J., Adams, R., & Iwashita, N. (2013). *Peer interaction and second language learning* (1st ed.). Routledge. https://doi.org/10.4324/9780203551349
Pica, T. (1994). Research on negotiation: what does it reveal about second-language learning conditions, processes, and outcomes? *Language Learning*, *44*(3), 493–527. https://doi.org/10.1111/j.1467-1770.1994.tb01115.x
Qin, J. (2008). The effect of processing instruction and dictogloss tasks on acquisition of the English passive voice. *Language Teaching Research, 12,* 61–82. http://dx.doi.org/10.1177/1362168807084494
Sato, M., & Ballinger, S. (2016). *Peer interaction and second language learning: Pedagogical potential and research agenda*. John Benjamins Publishing Company. https://doi.org/10.1075/lllt.45
Shak, J. (2006). Children using dictogloss to focus on form. *Reflections on English Language Teaching*, *5*(2), 47–62.
Shehadeh, A. (2011). Effects and student perceptions of collaborative writing in L2. *Journal of Second Language Writing*, *20*(4), 286–305. https://doi.org/10.1016/j.jslw.2011.05.010
Slavin, R. (1995). *Cooperative learning* (2nd ed.). Allyn & Bacon.
Stanley, J. (1992). Coaching student writers to be effective peer evaluators. *Journal of Second Language Writing*, *1*(3), 217–233. https://doi.org/10.1016/1060-3743(92)90004-9
Steward, B. L., Silva, L. H. R., & Gonzalez, J. A. T. (2014). Integrating language skills through a dictogloss procedure. *English Teaching Forum, 4*(2), 12–19.
Storch, N. (2018). Collaborative writing. *Language Teaching*, *52*(1), 40–59. https://doi.org/10.1017/s0261444818000320
Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing, 14,* 153–173.
Storch, N. (2002). Patterns of interaction in ESL pair work. *Language Learning*, *52*(1), 119–158. https://doi.org/10.1111/1467-9922.00179
Storch, N., & Wigglesworth, G. (2007). *Writing tasks: comparing individual and collaborative writing*. In M. del Pilar Garcia-Mayo (Ed.), *Investigating tasks in formal language learning* (pp. 157–177). Multilingual Matters.
Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching*, *46*(2), 195–207. https://doi.org/10.1017/S0261444811000486
Swain, M. (2006). Languaging, agency and collaboration in advanced second language proficiency. In: Byrnes, H., (Ed.), *Advanced language learning: The contribution of Halliday and Vygotsky* (pp. 95–108). Continuum.
Swain, M. (2000). The output hypothesis and beyond: mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 97–114). Oxford University Press.
Swain, M., & Lapkin, S. (2001). Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 99–118). Pearson Education.
Swain, M. & Miccolli, L.S. (1994): Learning in a content-based, collaboratively structured course: the experience of an adult ESL learner. *TESL Canada Journal, 12*, 15–28.
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass, & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
Talib, T., & Cheung, Y. L. (2017). Collaborative writing in classroom instruction: A synthesis of recent research. *The English Teacher, 46*(2), 43–57.
Tsimpli I.-M., Papadopoulou, D. (2009). Aspect and the Interpretation of Motion Verbs in L2 Greek. In Snape N., Leung Y K-I & M. Sharwood Smith (Eds.), *Representational Deficits in SLA: Studies in honor of Roger Hawkins* (pp.187-227). John Benjamins.
Ülger, Z., Dette-Hagenmeyer, D. E., Reichle, B., & Gaertner, S. L. (2017). Improving outgroup attitudes in schools: A meta-analytic review. *Journal of School Psychology*, *67*, 88–103. https://doi.org/10.1016/j.jsp.2017.10.002
Vygotsky, L. S. (1978). *Mind in society: Development of*

higher psychological processes. IHarvard University Press. https://doi.org/10.2307/j.ctvjf9vz4
Wajnryb, R. (1990). *Grammar dictation*. Oxford University Press.
Watanabe, Y., & Swain, M. (2007). Effects of proficiency differences and patterns of pair interaction on second language learning: collaborative dialogue between adult ESL learners. *Language Teaching Research*, *11*(2), 121–142. https://doi.org/10.1177/1362168806074599
Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, *26*(3), 445–466. https://doi.org/10.1177/0265532209104670
Wong, C. W. S., & Schweitzer, R. D. (2017). Individual, premigration and postsettlement factors, and academic achievement in adolescents from refugee backgrounds: A systematic review and model. *Transcultural Psychiatry*, *54*(5–6), 756–782. https://doi.org/10.1177/1363461517737015
Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry*, *17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x
Yolanda, D. (2019). The effect of running dictation method on students' writing ability in procedure text (a study at the xith grade students of sma Negeri 6 padangsidimpuan). *jurnal.ipts.ac.id*. https://doi.org/10.37081/liner.v2i3.1208
Zhu, W. (2001). Interaction and feedback in mixed peer response groups. *Journal of Second Language Writing*, *10*(4), 251–276. https://doi.org/10.1016/s1060-3743(01)00043-1

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2025 Paspali and Papadopoulou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

JLT Journal of Language Teaching

Peer-reviewed | Open Access | Google Scholar | MLA | Crossref | DOI

Call for Papers

Submit via https://jlt.ac/

Areas of Interest:

Language teaching intervention and experiments; Curriculum development; Language teacher education; Bilingual education; New technologies in language teaching; Testing, assessment, and evaluation; Educational psychology, and more.

We accept the following types of submission:

1. Research article: (6,000 to 8,000 words)
2. Review: (3,000 to 8,000 words)
3. Book review: (up to 3,000 words)
4. Features: (3,000 to 8,000 words)

Scan to submit your articles* & read more articles for free.

*Article Processing Charges Apply.

Contact: editor@jlt.ac

Pioneer Publications

ISSN (Online) 2770-4602

Journal of Language Teaching, *5*(2), 1-8, 2025
https://doi.org/10.54475/jlt.2025.005

Check for updates

Review OPEN ACCESS

Navigating intercultural competence with ChatGPT: Implications and recommendations for foreign language education

Yeon Mi Lee

Department of World Languages and Cultures, Samford University, Homewood, Alabama, United States

Received: May 10, 2025 / Accepted: June 2, 2025 / Published Online: June 4, 2025
© Pioneer Publications LTD 2025

Abstract

Despite the increasing corpus of research concerning the application of ChatGPT in adaptive learning, grammar, and lexical acquisition in foreign language education (FLE), its effectiveness on the cultivation of intercultural competence (IC) as an essential learning outcome in FLE remains comparatively underexplored. This review article introduces the theoretical background and definitions of intercultural competence in second language acquisition (SLA) before analyzing empirical studies that examine ChatGPT's cross-cultural performance across multiple languages, cultures, and methodologies. It then discusses the implications of utilizing ChatGPT for intercultural engagement in FLE. The article concludes by underscoring the imperative of critical artificial intelligence literacy (CAIL) among language learners and offers recommendations for best practices in FLE.

Keywords intercultural competence, foreign language education, second language acquisition, ChatGPT, critical artificial intelligence literacy

1. Introduction

> *Language is not merely a reproducing instrument for voicing ideas, but is itself a shaper of ideas. …We see and hear and otherwise experience very largely as we do because language habits of our community predispose certain choices of interpretation.*
>
> –Benjamin Lee Whorf, Language, Thought, and Reality (1959, p. 212)

While the integration of the intercultural learning outcomes into the mainstream L2 curricula and courses did not begin until the 1990s, driven by the new demands of globalization and international affairs, its theoretical underpinnings in SLA can be traced back to the 1920s. Benjamin Lee Whorf's ideas on the intrinsic link between language and culture, as illustrated in the preceding epigraph, derived from the rising notions in the field of cultural anthropology and linguistics of the 1920s and 1930s, primarily from the work of his mentor and linguist Edward Sapir (Subbiondo, 2005). Sapir and Whorf's ideas on the cultural determinants of human language laid the groundwork for the sociocultural theoretical framework in the following decades (Thorne & Tasker, 2011). In psychology, Sapir's ideas met the theoretical construct of Vygotsky's Sociocultural Theory (SCT), whose emphasis on the significance of contextual knowledge in learning later brought a great impact to the field of second language acquisition; for example, notions such as "interaction and negotiation" in speech analysis and development was introduced by Michael Long (1980), building awareness about the constructivist process wherein speech is co-constructed, navigated, and negotiated between speakers. Thus, sociocultural awareness shifted the traditional paradigms of SLA by taking language acquisition beyond the confines of lexical and grammatical structures and integrating cultural competence as an essential learning outcome. Wiseman et al. (1989) assert that "cultural knowledge is an important determinant of one's ability to minimize misunderstandings with someone from another culture. Cultural knowledge has a positive effect on other [cross-cultural competence] attributes and maximizes intercultural competency" (p. 351).

In recent years, the swift adoption of OpenAI's ChatGPT as a digital learning tool in FLE has driven significant scholarly interest, presenting its impact on adaptive learning, personalized feedback, grammar, and lexical acquisition (Anjum et al., 2024). While the advantages of integrating ChatGPT into L2 pedagogy continue to be widely explored, discussion concerning its impact and implications on the development of intercultural competence among L2 learners has garnered less attention. The current article introduces the theoretical frameworks and definitions that have shaped the discourse around intercultural competence in FLE. It then provides an analysis of the empirical studies that focus primarily on assessing ChatGPT's performance in engaging cultural knowledge and interaction across multiple languages and cultures. Drawing upon this review, the current article offers insights on potential implications and recommendations for FLE. Finally, the article

concludes by highlighting the significance of cultivating critical artificial intelligence literacy (CAIL) to educate L2 learners and instructors about the potential risks and benefits of utilizing ChatGPT in cross-cultural contexts.

2. Intercultural competence in FLE: Theoretical background and definitions

The expansion of globalization and international trade during the late 20th century created a high demand for professionals with optimal skills capable of navigating the new multicultural landscape (Garrett, 2025). Intercultural competence emerged as an essential outcome for L2 or foreign language programs, as language associations and institutions provided official guidelines on its implementation; in the United States, the World-Readiness Standards was established in 1996 by the American Council on the Teaching of Foreign Languages (ACTFL), offering specific guidelines for achieving cultural learning outcomes, known as the 5 C's framework– Communications, Cultures, Connections, Comparisons, Communities. In Europe, the Common European Framework for Languages began to provide similar guidelines (Garrett, 2025).

Since the 1990s, intercultural competence in FLE has garnered significant attention from scholars calling for this outcome to "be examined and interpreted as a multifaceted process" (Stier 2006, p. 5). A range of pedagogical frameworks aimed at cultivating this skill were proposed via traditional classroom instruction, cultural immersion via study-abroad or experiential learning, and other comparable methods (Deardorff 2006; Stier 2006; Leask 2015). Along with increased interest in methodology came varied attempts to define this concept: Alvino Fantini defines intercultural competence as "the complex of abilities needed to perform effectively and appropriately when interacting with others who are linguistically and culturally different from oneself" (2005, p.1). On the other hand, Hammer's definition underlines the dynamic, nuanced process of interaction where the speakers acquire "the capability to shift one's cultural perspective and appropriately adapt behavior to cultural differences and commonalities" (2015, p. 483). Likewise, Spitzberg and Changnon define intercultural proficiency as "the appropriate and effective management of interaction between people who, to some degree or another, represent different or divergent affective, cognitive, and behavioral orientations to the world" (2009, p. 7).

Theoretical considerations on the concept of intercultural competence have followed a long trajectory of epistemic genealogy since Edward Sapir and Benjamin Lee Whorf introduced their foundational notion on the interrelationship between language and culture (Subbiondo, 2005). The development of Lev Vygotsky's Sociocultural Theory in psychology, for whom the individual cognitive competence is dialectically linked to their community's culture, is a prime example of the overlapping impact of the emerging ideas that began to circulate at the time (Vygotsky, 1978). Vygotsky's ideas later influenced key concepts in applied linguistics, including the notion of "comprehensible input" by Stephen Krashen (1982) and "comprehensible output" by Merrill Swain (1985), both of which underscore the significance of cultural and contextual knowledge in achieving effective L2 comprehension and production.

Acceptance and practice of intercultural competence were further validated and recognized thanks to its incorporation into the discipline of Intercultural Communication Studies (ICS). Until the early 1990s, the domain of cultural competence in FLE had been strongly influenced by comparativist and positivist models of cross-cultural psychology, where culture was primarily understood in terms of nationality and compared through generalized constructs (Hofstede, 1991; Triandis, 1990). However, the rise of ICS in the 1990s reinforced the understanding of human communication as a dynamic, interpersonal process for effective meaning making and negotiation (Liddicoat & Scarino, 2013). Michael Byram's intercultural framework was developed based on this notion, outlining practical and ethical objectives in achieving comprehensive linguistic and cross-cultural proficiency in five competencies, succinctly summarized by Schenker (2012) as "knowledge of self and other, attitudes of openness and curiosity, skills or interpreting and relating, skills of diversity and interaction, and critical cultural awareness" (p. 450). The new intercultural frameworks shifted the old paradigms of FLE towards a more holistic, integrated framework prioritizing meaningful intercultural engagement in language education, as well as its broad acknowledgement across L2 language programs and curricula. As Byram et al. (2002) suggest, the goal of L2 intercultural competence is "to cultivate learners' cross-cultural competence alongside linguistic proficiency; to prepare them for engagement with individuals from diverse cultures; to facilitate their understanding and acceptance of others as unique entities possessing different perspectives, values, and behaviors; and to assist them in recognizing that such interactions serve as enriching experiences" (2002, p. 10).

3. Empirical research on ChatGPT's intercultural performance and its impact on FLE

The use of ChatGPT for a designated purpose in FLE has incited substantial interest within the domain of applied linguistics (see, e.g., Huang et al., 2022; Xiao et al, 2023; Zou et al., 2023), garnering both excitement as well as concern for its adoption in educational settings. One area of concern is L2 learners' lack of awareness and uncritical engagement with the AI chatbots, utilizing them as a search engine or ultimate authority for the knowledge they are seeking. This new digital environment has redefined the L2 learners' educational experience and critical decision-making processes for language programs (Darvin, 2025). Kirschenbaum and Raley (2024) highlight the epistemological concerns deriving from the development of new GPT models; GPT-4 features approximately 1.8 trillion parameters compared to the 175

billion parameters of the previous model GPT-3. What this indicates "is not yet another novel computational application or approach but rather a general condition of language and life" (2024, p. 509). This condition raises ethical and epistemological concerns as the chatbots' competence in language mimicry may not inherently translate into output that is free from cultural bias or misinterpretation. For the purposes of FLE, the application of ChatGPT in L2 intercultural engagement carries profound implications for culturally sensitive and inclusive pedagogy, particularly concerning the representation of cultural perspectives and worldviews pertaining to ethnic minority cultures and languages. This section provides a selective review of empirical studies on ChatGPT's intercultural performance, guided by three criteria: 1) the studies experimented with multiple languages and/or (sub)cultures to measure the IC performance of ChatGPT; 2) the studies represented methodological diversity employing computational benchmarking against human survey data as well as qualitative analysis from direct AI-human interactive engagement; 3) the studies addressed the impact of ChatGPT's IC performance on educational settings.

3.1. Empirical research from computational benchmarking against pre-existing human survey data

Georgiou (2025) investigated the cultural performance of ChatGPT by evaluating the chatbot's response to a prompt eliciting a general description of a selection of developed and developing countries based on the Human Development Index (HDI). The study employed ChatGPT-3.5 to prompt cultural descriptions of 20 countries–10 developed or high HDI countries, such as Singapore, Sweden, Denmark, etc., and 10 developing or low-HDI nations, such as Sierra Leone, Niger, South Sudan, and others. Drawing on the chatbot's responses, Georgiou (2025) conducted a sentiment analysis using the quantitative discourse analysis (QDAP) package within the R programming environment and implemented the Bayesian regression model to analyze the data. The study found that although ChatGPT produced generally positive descriptions across all countries, Bayesian statistical analysis showed that ChatGPT used language that encompassed more positive sentiments for countries with elevated HDI scores–predominantly European nations–compared to their low-HDI counterparts, largely from Africa. ChatGPT's description of the different countries and their cultures indicated a higher sentiment score associated with high-HDI nations, while the language used for low-HDI countries generated a lower sentiment score. These findings have important implications for the cultivation of intercultural competence in FLE, as ChatGPT users' lack of critical awareness and holistic understanding of cultural diversity within developing nations "may perpetuate perceptions of superiority or inferiority based on national economic status" (Georgiou, 2025, p. 5), posing the risk of reinforcing cultural stereotypes. While Georgiou's research (2025) is somewhat limited due to its concentration on a single language (English), it is nonetheless a compelling study that highlights the potential effects of utilizing ChatGPT in cross-cultural educational contexts.

Cao et al. (2023) examined the cultural alignment of ChatGPT by evaluating its responses against human-generated data from pre-existing responses to the Hofstede Culture Survey. Their experiment included three standardized prompts with questions across five distinct languages and cultures (English, Chinese, German, Spanish, Japanese) and employed Hofstede's framework across six dimensions (e.g., Power Distance, Individualism, Masculinity, Uncertainty Avoidance, etc.). A multi-turn interaction strategy with three categories of knowledge injection was employed to evaluate the model's adaptability and consistency with culture-specific questions. Findings revealed that the type of interaction strategy used in multi-turn conversations can impact the responses from ChatGPT, uncovering the dependency of the model's performance on prompts and conversation structures. Moreover, the study found that ChatGPT's cultural alignment scores varied significantly across the different cultures. ChatGPT performed in higher alignment with contexts familiar to American cultural values, which increased even more when prompted in English. Consistency rates for English prompts exceeded 70% and tended to favor American norms. ChatGPT responses showed better alignment when prompted using the native language of the target culture rather than in English, suggesting the prompt language has a great influence on ChatGPT-generated responses. The study by Cao et al. (2023) also found that ChatGPT's cultural alignment decreased when prompted in non-English languages, but Chinese and German tended towards a stronger alignment. These findings are particularly telling for L2 education other than English, given that ChatGPT's training data is predominantly English (96%, according to Ouyang et al.), resulting in its tendency to default to American cultural norms and paradigms. ChatGPT's language dominance poses challenges for non-English L2 learners and their intercultural outcomes, underscoring concerns about culturally misaligned information regarding language minorities of the Global South.

Similar to the study by Cao et al. (2023), the research undertaken by Wang et al. (2023) assessed cultural alignment of LLM's outputs in relation with pre-existing human-generated data from the World Values Survey. They utilized two models of LLMs, OpenAI's ChatGPT-4 and text-davinci-003. The study developed a benchmark dataset, incorporating tangible cultural artifacts (e.g., holidays, films, literature, songs, etc.) as well as intangible cultural values (e.g., opinions and beliefs), and categorized the analysis into two cultural dimensions –survival versus self-expression values, and traditional versus secular-rational values. Wang et al. compared the different cultural responses in six distinct languages: English, Chinese, Russian, Indonesian, Hindi, and Arabic. Similar to the findings uncovered by Cao et al. (2023), this study found that ChatGPT's responses were consistently more congruent with the English World Values Survey data, even when prompted to respond within specific cultural frameworks, such as Chinese or Arabic. The study found that chatbot responses increased in cultural alignment with English-speaking cultures and nations, corroborating Cao et al.'s findings about the linguistic and cultural

dominance of the English language found in AI-assisted environments. For Wang et al. (2023), the predominant influence of English in data training of LLMs reinforces a systemic issue dominated by the English cultural paradigms. As for L2 education, the study suggests caution and awareness when interacting with ChatGPT for cultural learning outcomes in non-English language pedagogy, as the chatbots' cultural misrepresentations may hinder the acquisition and engagement of culturally sensitive skills. As a result, the study suggests the need for robust, human-assisted intercultural education in FLE and calls for a critical approach to AI-assisted technologies within controlled settings.

Resembling in methodology and findings with Wang et al.'s study (2023), the experiment conducted by Tuna et al. (2024) deployed two LLM models, GPT-3.5-turbo and GPT-4, and compared their responses against human subject responses derived from the World Values Survey. Similar to previous studies, this study compared chatbots' output against pre-existing data on cultural values across cultures and nations, rather than engaging in human responses in the conventional form of interviews or qualitative data. The study probed ChatGPT's cultural performance in five distinct languages: German, French, English, Spanish, and Portuguese. Moreover, the experiment included studies within ten subcultural variations: English-speaking Great Britain and the USA, German-speaking Germany and Austria, Spanish-speaking Spain and Mexico, French-speaking Canada and France, and Portuguese-speaking Brazil and Portugal. Tuna et al.'s assessment (2024) employed inquiries on topics such as Trust, Faith, and Happiness, and measured the proximity of the GPT's response to existing human cultural norms, using the Euclidean distance. Similar to the findings from previous studies, Tuna et al. (2024) found that the chatbots showed closer proximity to Euclidean distance when prompted in English within Western-centric cultures, and that a downgraded version of GPT-3.5-turbo performed better than GPT-4 in cultural alignment, especially within the German language setting. However, both models showed lower cultural alignment when engaging subcultures and languages, particularly in Mexican Spanish and Brazilian Portuguese. For instance, in Mexican Spanish, the GPT-3.5-turbo's average distance from human values was significantly higher, indicating a clear lack of cultural nuance essential for language interaction in this subculture. The chatbots' high degree of distance from human cultural values in a specific subculture suggests that learners using these models may be exposed to culturally inappropriate or oversimplified content that can hinder the development of intercultural communicative abilities. Without accurate cultural representations, language learners can adopt and internalize expressions or views that are misaligned with the sociocultural expectations of native cultures and speakers. Likewise, these findings underscore the imperative of teaching intercultural competencies so that students are equipped to engage with the cultural mindset of the native cultures and speakers.

Ahmad et al. (2024) further investigated the efficacy of ChatGPT in capturing the cultural competence and nuances inherent in languages other than English. They examined the chatbots' performance in Hausa, a low-resource language utilized primarily in the West African region. The research assessed ChatGPT's outputs and compared them to those provided by 18 native Hausa speakers located in Nigeria. It employed 37 culturally relevant prompts that elicited responses on cultural norms and emotional expectations. The methodology encompassed two distinct phases, initially with participants generating their own open-ended responses to the questions, and then having them evaluate the cultural and emotional authenticity of ChatGPT's output utilizing a Likert scale. This two-phase process allowed the researchers to evaluate both the semantic similarity and emotional resonance, which are essential elements for navigating intercultural language exchange. Ahmad et al. (2024) found significant cultural and emotional discrepancies when compared to human responses. While ChatGPT displayed semantic similarity, the authors found that the chatbots lacked emotional depth, defaulting frequently to neutral tones or perspectives. For example, when prompted on the question "How would you feel if your student calls you by your first name?", ChatGPT responded with a polite acceptance, while the majority of Hausa speakers regard it to be culturally inappropriate and disrespectful. An average of only 8.2 participants said GPT's responses were likely articulated by a native Hausa speaker, while 5.2 expressed otherwise. Ahmad et al.'s study (2024) demonstrates that overreliance on LLM models could inadvertently promote generic or culturally inappropriate notions, particularly within underrepresented linguistic communities.

3.2. Empirical research from direct human interaction with ChatGPT

The studies examined thus far provided insights utilizing standardized prompts and pre-existing human responses to static survey data that measured ChatGPT's cultural proficiency. However, within their limitations is the absence of direct user engagement with the chatbots in the form of iterative conversation, interaction, direct feedback, and provision of additional contexts that more closely resemble real-world exchange in a dynamic environment. Therefore, in what follows, we will review empirical studies that have been conducted by means of interactive user feedback and direct engagement with the chatbots to assess their impact in a more fluid, collaborative setting.

Masato Tahara's study (2024) brings significant insights regarding the application of ChatGPT in the domain of L2 translation and cultural competence. Tahara investigated ChatGPT-4's cultural performance in the context of collaborative translation and dialogic reading, instructing five Japanese and two Malaysian Chinese students to engage with AI-translated adaptations of the Japanese novel *Jimmy* by Aoumi in English and Chinese. The author instructed the students to analyze cultural interpretation and intricate complex meanings from AI-generated translation. This case study integrated a dual-phase translation framework that incorporated GPT's translation with human collaboration. The method followed AI-translated text (Translation 1), followed by a

refinement process with human-addressed critical reading questions. The chatbots' responses were uploaded as analytical texts to inform a subsequent translation (Translation 2). The research showed how the diversity of cultural backgrounds from participants influenced their interpretations. For example, the Japanese participants resonated with societal pressures represented in the narrative, while Malaysian Chinese participants gave multiple perspectives on multicultural norms. Their results found that Translation 2 was more effective in conveying cultural nuances and characters' psychology, suggesting that the collaborative framework built upon human input and interaction with ChatGPT can enhance cross-cultural understanding. Tahara's study (2024) on AI-human interaction shows that new meanings can arise from the confluence of varied cultural inputs generated by users, lending to the chatbots' potential for developing interpretive flexibility and adjustment based on direct human feedback.

In engaging ChatGPT's direct interaction with human subjects, Darvin (2025) examined the interactions of six secondary school students (Grades 8-12) in British Columbia who spoke languages other than English in their home, and explored the implications of machine-human interaction for cross-cultural competence. The research utilized a case study methodology and data consisting of semi-structured interviews and multimodal discourse analysis of GenAI interactions with ChatGPT, Copilot, and CharacterAI. This methodology enabled the author to observe the ways the students engaged with these tools, the digital literacies enacted during the interaction, and the cultural knowledge reflected in the generated outputs. The results revealed that, while the students' interaction with ChatGPT in areas specific to lexicality, grammatical correction, and rewriting was promising, a majority of students displayed a lack of critical awareness about the mechanisms by which ChatGPT generated its responses. Many students, lacking critical digital literacy, approached the tool as a neutral source of information, failing to think critically about its cultural assumptions inherent in those outputs. One example is that ChatGPT frequently aligned with prevailing discourses or cultural assumptions about academic writing, discouraging the use of first-person pronouns, which represents a specific cultural bias. These behaviors reflect potential hazards of promoting language skills without concurrent development of cultural competence. Another critique that the author makes illustrates how particular AI features like avatar customization or premium access influence learners' engagement levels that can open or restrict their participation in more complex writing tasks or cross-referencing contexts, depending on the type of electronic device in use. This is an important finding because it affects the students' learning experience, what they learn, and what they are exposed to. Moreover, the study brings a critical perspective to the AI tools' lack of ability to accommodate low-resource languages, resulting in misinterpretations of languages such as Vietnamese or Malayalam, which exacerbates the cultural bias against non-dominant cultural identities. As a result, while Darvin (2025) does validate the utility of ChatGPT for L2 learning as a practical tool for areas specific to lexical or grammar acquisition, the author critiques its potential risks for cross-cultural engagement. Among the limitations of Darvin's study (2025) is its relatively small pool of participants, which may not reflect adequately the diversity of L2 learners on a larger scale. Despite this, Darvin's study (2025) is compelling research that supports the argument that effective L2 learning must involve a culturally critical perspective beyond mere technical proficiency.

In assessing ChatGPT's impact on the development of culturally-proficient L2 curriculum and instructional design, Kim et al.'s study (2023) employed a virtual learner persona – a Korean undergraduate student learning business English writing competencies for employment. The simulated learner had intermediate English proficiency and five years of language experience. While the study is limited due to the lack of live interaction with real-human learners, the experiment nevertheless enabled the authors to assess and gain insight into the quality of GPT's response in the area of instructional design and cultural competence in educational scenarios. The study deployed a two-phase methodological framework: first, GPT was instructed to design a business English curriculum based on H. Douglas Brown's model for L2 course development. Then, the chatbot was prompted to engage in teaching this course using a Task-Based Language Teaching (TBLT) approach. The research measured GPT's performance in building lesson plans, task assignment, and feedback capability. Kim et al. (2023) found that the chatbot was successful in generating structure, topic-relevant curriculum, offering effective examples for writing assignments, and linguistic scaffolding. However, in the area of cross-cultural language proficiency, the authors found that ChatGPT lacked the ability to replicate real-world business scenarios that engage in nuanced, distinct cultural settings. While the chatbots performed the prompted tasks, the level of engagement was simplistic and superficial, displaying limited connection with sociocultural variation and context-specific considerations in real-life communication. For instance, the study found that ChatGPT feedback was highly confined to grammatical and lexical correction, undermining deeper communicative and cultural dimensions like politeness strategies, business cultural norms, and regional differences in business etiquette. Further, Kim et al. (2023) found that the chatbots are overly focused on grammatical accuracy rather than the ability to engage and perform interculturally and pragmatically. These shortcomings exposed the learner to missed opportunities for essential intercultural engagement and effective communication in multicultural business settings. While the GPT is a promising tool for self-directed learning, its cultural performance is limited when addressing sociolinguistic depth for authentic intercultural communication. Given the drawbacks, Kim et al. (2023) suggest that L2 language learners should use the tool under the supervision and guidance of educators who can deliver expertise on culturally-appropriate knowledge and critical engagement. Kim et al.'s research (2023) underscores the importance of human-led instructional oversight and guidance for adequate cultural skills to ensure L2 learners are equipped with the tools needed to

navigate diverse communicative contexts with cultural sensitivity and confidence in multicultural business settings.

4. Implications and recommendations for foreign language education

Empirical studies on the integration of ChatGPT into the acquisition of intercultural skills in SLA demonstrate limitations and opportunities for L2 learners and educators alike. While the chatbots' performance in supporting FLE in areas like linguistic scaffolding, grammar, and lexical acquisition are promising, challenges remain when adopting the tool to enhance L2 intercultural outcomes: among the persisting challenges are the prevalence of cultural and linguistic dominance of algorithmic contexts that resource the AI-pre-training datasets, which results in increased chatbot's alignment with dominant cultural paradigms (Wang et al., 2023; Cao et al., 2023). This has important implications for L2 intercultural pedagogy, for the teaching of cultural norms of subcultures and linguistic minority groups can be impacted by the overwhelming prevalence of linguistic dominance encountered in AI-enhanced environments. This, in turn, can lead to diminished cultural perspectives and visibility of minority languages and subcultures in FLE settings, course materials, curriculum, instructional design, and more (Broadhead, 2024). To tackle this problem, the current article calls for the imperative of cultivating critical artificial intelligence literacy (CAIL) among students and instructors to inform best practices in FLE. The purpose of CAIL is to raise awareness around the epistemic impact of LLMs and the algorithmic processes whereby knowledge is generated and disseminated in digital environments. As the empirical studies have demonstrated, algorithmic processes can often perpetuate prevailing cultural viewpoints while marginalizing alternative perspectives. Therefore, the current article offers four actionable insights or recommendations for cultivating CAIL in FL intercultural contexts:

1) Assist FL learners with specific guidelines and pedagogy to foster critical evaluation of cross-cultural information sourced by algorithmic systems. Supervise the use of ChatGPT for intercultural outcomes, facilitating access to human-led cultural expertise and cross-referencing practices.
2) Provide holistic guidance to facilitate AI-human collaborative learning and critical thinking practices through iterative, dialogic processes of human-generated input and feedback (Tahara, 2024).
3) Offer opportunities for reflective practices in the form of cultural response, interpretation, or reflective journal to address the discrepancies between AI-generated content and humans' lived experiences and cultivate metacognitive awareness about the process of knowledge construction and representation (Darvin, 2025).
4) Practice critical evaluation and cross-cultural comparison by contextualizing ChatGPT's culture-specific prompts and by examining its response within non-dominant cultural frameworks (Wang et al., 2023; Tuna et al., 2024).

While the chatbots may perform well syntactically, their output often fails to capture in-depth cross-cultural representations that reflect nuanced sensitivity to underrepresented linguistic communities. Educators and learners alike should be aware of these practical recommendations that can assist in achieving a more culturally inclusive and pedagogically relevant experience in FLE.

5. Conclusion

The seemingly boundless use of ChatGPT in educational contexts has generated a large corpus of research in FLE, noting its benefits in enriching lexical and grammar acquisition, L2 writing, and self-directed and adaptive learning. However, its impact on fostering intercultural competence continues to be developed and should be approached with caution and acute awareness about its potential risks. Critical thinking skills and methods should be promoted among L2 learners to equip them with the tools to evaluate knowledge production and cultural output generated by AI, particularly in assessing culturally sensitive knowledge of language minorities and subcultures. Algorithmic cultural (mis)representations can inadvertently entrench systemic inequities through their classification processes, and hence diminish the potential for sustainability and preservation of heritage and less-commonly taught languages. As Tahara (2024) suggests, GPT should be leveraged to advance language acquisition, but users should be critically aware of its best practices and risks involving L2 intercultural engagement.

Professor Yeon Mi Lee received her PhD degree from the University of California, Los Angeles. Dr. Lee is a faculty member in the Department of World Languages and Cultures at Samford University in the United States.
Email: ymlee1331@gmail.com

References

Ahmad, I. S., Dudy, S., Ramachandranpillai, R., & Church, K. (2024). Are generative language models multicultural? A study on Hausa culture and emotions using ChatGPT. *arXiv:2406.19504*. https://doi.org/10.48550/arXiv.2406.19504

Anjum, F., Raheem, B. R., & Ghafar, Z. N. (2024). The impact of ChatGPT on enhancing students' motivation and learning engagement in second language acquisition: Insights from students. *Journal of e-learning Research*, *3*(2), 1-11. https://doi.org/10.33422/jelr.v3i2.679

Broadhead, L. (2024). Insidious chatter versus critical thinking: Resisting the Eurocentric siren song of AI in the classroom. *Journal of Applied Learning and Teaching*, *7*(2). 28-37. https://doi.org/10.37074/jalt.2024.7.2.9

Byram, M., Gribkova, B, & Starkey, H. (2002). *Developing the intercultural dimension in language teaching. A practical introduction for teachers.*

Council of Europe Publishing, Language Policy Division.
Cao, Y., Zhou, L., Lee, S., Cabello, L., Chen, M., & Hershcovich, D. (2023). Assessing cross-cultural alignment between ChatGPT and human societies: An empirical study. *arXiv:2303.17466*. https://doi.org/10.48550/arXiv.2303.17466
Crawford, Kate. (2021). *Atlas of AI: Power, politics, and the planetary costs of artificial intelligence*. Yale University Press.
Darvin, R. (2025). The need for critical digital literacies in generative AI-mediated L2 writing. *Journal of Second Language Writing, 67*, 101186. https://doi.org/10.1016/j.jslw.2025.101186
Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education, 10*(3), 241–266. https://doi.org/10.1177/10283153062870
Fantini, A. (2005). *About intercultural communicative competence: A construct*. School for International Training.
Garrett-Rucks, P. (2025). Assessing intercultural competence. In C. Facke, X. Gao & P. Garrett-Rucks (Eds.), *The Handbook of plurilingual and intercultural language learning* (pp. 247-260). John Wiley & Sons.
Georgiou, G. P. (2025). ChatGPT exhibits bias toward developed countries over developing ones, as indicated by a sentiment analysis approach. *Journal of Language and Social Psychology*, *44*(1), 132-141. https://doi.org/10.31234/osf.io/49gzc_v1
Hammer, M. R. (2015). Intercultural competence development. In J. M. Bennett (Ed.), *The SAGE encyclopedia of intercultural competence* (pp. 483-486). Sage Publications.
Hofstede, G. (1991). *Cultures and organisations: Software of the mind*. McGraw Hill.
Huang, W., Hew, K. F., & Fryer, L. K. (2022). Chatbots for language learning—are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning, 38*(1), 237–257. https://doi.org/10.1111/jcal.1261
Kim, S.-Y., Shim, J., & Shim, J. (2023). A study on the utilization of OpenAI ChatGPT as a second language learning tool. *Journal of Multimedia Information System*, *10*(1), 79–88. https://doi.org/10.33851/jmis.2023.10.1.79
Kirschenbaum, M. & Raley, R. (2024). AI and the University as a service. *Pmla/Publications of the Modern Language Association of America, 139*(3), 504-515. https://doi.org/10.1632/S003081292400052X
Krashen, S. (1982). *Principles and practice in second language acquisition*. Oxford University Press.
Leask, B. (2015). *Internationalizing the curriculum*. New York: Routledge.
Liddicoat, A. & Scarino, A. (2013). *Intercultural language teaching and learning*. Wiley-Blackwell.
Long, M. H. (1980). *Input, interaction, and second language acquisition*. University of California, Los Angeles.
Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C. L., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., Schulman, J., Hilton, J., Kelton, F., Miller, L., Siemens, M., Askell, A., Welinder, P., Christiano, P., Leike, J., & Lowe, R. (2022). Training language models to follow instructions with human feedback. *arXiv:2203.02155*. https://doi.org/10.48550/arXiv.2203.02155
Schenker, T. (2012). Intercultural competence and cultural learning through telecollaboration. *CALICO Journal, 29*(3), 449–470. https://doi.org/10.11139/cj.29.3.449-470
Spitzberg, B., & Changnon, G. (2009). Conceptualizing intercultural competence. In D. K. Deardorff (Ed.), *The Sage handbook of intercultural competence* (pp. 2–52). SAGE Publications. https://doi.org/10.4135/9781071872987.n1
Stier, J. (2006). Internationalisation, intercultural communication and intercultural competence. *Journal of Intercultural Communication, 11*(1), pp. 1–12. https://doi.org/10.36923/jicc.v6i1.422
Subbiondo, J.L. (2005). Benjamin Lee Whorf's theory of language, culture, and consciousness: A critique of Western science. *Language and Communication*, *25*(2), 149-159. https://doi.org/10.1016/J.LANGCOM.2005.02.001
Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235–253). Newbury House.
Tahara, M. (2024). How does Generative AI function as an active cross-cultural communication agent? *INContext*, *4*(2), 63–83. https://doi.org/10.54754/incontext.v4i2.103
Thorne, S. L., & Tasker, T. (2011). Sociocultural and cultural-historical theories of language development. In *The Routledge handbook of applied linguistics* (pp. 487-500). Routledge. https://doi.org/10.4324/9780203835654.ch34
Triandis, H.C. (1990). Cross-cultural studies of individualism and collectivism. In J.J. Berman (ed.), *Cross-cultural perspective* (pp. 41–33). University of Nebraska Press.
Tuna, M., Schaaff, K., & Schlippe, T. (2024, November). Effects of language-and culture-specific prompting on ChatGPT. In *2024 2nd International Conference on Foundation and Large Language Models (FLLM)* (pp. 73-81). https://doi.org/10.1109/fllm63129.2024.10852463
Vygotsky, L. (1978). *Mind in society. The development of higher psychological processes*. (M. Cole, V. John-Steiner, S. Scribner & E. Souberman (Eds.). Harvard University Press. https://doi.org/10.2307/j.ctvjf9vz4
Wang, W., Jiao, W., Huang, J., Dai, R., Huang, J., Tu, Z., & Lyu, M. R. (2023). *Not all countries celebrate Thanksgiving: On the cultural dominance in Large Language Models. arXiv:2310.12481*. https://doi.org/10.48550/arXiv.2310.12481
Whorf, B. L. (1959). *Language, thought, and reality: Selected writings of Benjamin Lee Whorf* (J. B. Carroll, Ed.). MIT Press.

Wiseman, R.L., Hammer, M.R. and Nishida, H. (1989). Predictors of intercultural competence, *International Journal of Intercultural Relations*, *13*(3), 349–370. https://doi.org/10.1016/0147-1767(89)90017-5

Xiao, F., Zhao, P., Sha, H., Yang, D., & Warshauer, M. (2023). Conversational agents in language learning. *Journal of China Computer-Assisted Language Learning*. https://doi.org/10.1515/jccall-2022-0032

Zou, B., Reinders, H., Thomas, M., & Barr, D. (2023). Editorial: Using artificial intelligence technology for language learning. *Frontiers in Psychology, 14*, 1287667. https://doi.org/10.3389/fpsyg.2023.1287667

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2025 Lee. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

JLT Journal of Language Teaching

Peer-reviewed | Open Access | Google Scholar | MLA | Crossref | DOI

Call for Papers

Submit via https://jlt.ac/

Areas of Interest:

Language teaching intervention and experiments; Curriculum development; Language teacher education; Bilingual education; New technologies in language teaching; Testing, assessment, and evaluation; Educational psychology, and more.

We accept the following types of submission:

1. Research article: (6,000 to 8,000 words)
2. Review: (3,000 to 8,000 words)
3. Book review: (up to 3,000 words)
4. Features: (3,000 to 8,000 words)

Scan to submit your articles* & read more articles for free.

*Article Processing Charges Apply.

Contact: editor@jlt.ac

Pioneer Publications

ISSN (Online) 2770-4602

Journal of Language Teaching, *5*(2), 10-19, 2025
https://doi.org/10.54475/jlt.2025.007

Check for updates

Review OPEN ACCESS

A bibliometric analysis of research on foreign language classroom anxiety using CiteSpace[1]

Mengyao Xun, Ran Wei

School of Foreign Studies, Jiangsu Normal University, Xuzhou, Jiangsu Province, China

Received: May 1, 2025 / Accepted: June 18, 2025 / Published Online: June 24, 2025
© Pioneer Publications LTD 2025

Abstract

This study presented a bibliometric analysis of foreign language classroom anxiety (FLCA) research published between 2002 and 2024. Using the Web of Science (WoS) database as the primary data source, CiteSpace software was employed to generate scientific knowledge maps, illustrating the evolutionary trajectory of FLCA scholarship. Findings indicated a fluctuating yet overall upward trend in publication frequency, reflecting the growing academic interest in this domain. The analysis revealed a well-established core research team engaged in sustained contributions, fostering scholarly advancements. Three primary research clusters emerged: (1) Factors influencing FLCA, including intrinsic learner traits and external classroom dynamics; (2) Development and validation of FLCA measurement instruments, particularly the Foreign Language Classroom Anxiety Scale (FLCAS) and its adaptations; and (3) The impact of FLCA on language learning outcomes, highlighting its correlation with academic achievement and pedagogical strategies. Based on identified limitations in current research, this study finally proposed directions for future investigation.

Keywords foreign language classroom anxiety, bibliometric analysis, CiteSpace, research trends, language learning emotions

1. Introduction

Since the 1970s, scholars have increasingly recognized the significance of emotions in teaching effectiveness, with anxiety emerging as a critical factor influencing language learning. Research on foreign language anxiety has expanded substantially, following Brown's (1973) early investigation into emotional factors in language acquisition. Brown suggested that anxiety could significantly impact learners' language acquisition processes. In 1986, Horwitz et al. introduced the concept of foreign language anxiety, defining it as "a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process" (Horwitz, Horwitz, & Cope, 1986, p. 128) Anxiety is widely regarded as a major emotional variable affecting foreign language learning, influencing learners' language competence, academic achievement, and teacher assessments (Gardner, 1985; Steinberg & Horwitz, 1986; Young, 1991). Due to its pivotal role in language education, anxiety has become a central focus of research in foreign language teaching.

Horwitz et al. (1986) expanded on Foreign Language Classroom Anxiety (FLCA), establishing three connected components: (a) communicative apprehension, (b) dread of unfavorable assessment, and (c) exam anxiety. Communication apprehension refers to an individual's anxiety during actual or anticipated interactions (McCroskey, 1984). Fear of negative evaluation describes distress associated with perceived judgments from others (Watson & Friend, 1969), while test anxiety involves excessive worry regarding performance outcomes in evaluative situations (Sarason, 1984). Subsequent studies have reinforced the idea that FLCA encompasses a broad spectrum of anxieties linked to classroom learning and specific language skills such as speaking and reading (e.g., Pae, 2013). Horwitz (2016) created the Foreign Language Classroom Anxiety Scale (FLCAS) to measure language anxiety within educational environments. This scale has been extensively used in empirical studies focusing on second language acquisition.

Despite growing academic interest in FLCA, there remains a scarcity of bibliometric reviews examining this field comprehensively. CiteSpace, a widely recognized bibliometric tool, enables researchers to map knowledge structures and identify emerging trends. Applying CiteSpace to FLCA research published from 2002 to 2024 can offer valuable insights into the current state of scholarship, highlight prevalent research themes, and provide guidance for future studies.

[1] This work was supported by 2025 Postgraduate Research & Practice Innovation Program of Jiangsu Normal University (No. 2025XKT0430) and Humanities and Social Science Project of Ministry of Education of the People's Republic of China (Grant No. 24YJC740071).

2. Systematic Review Supported by CiteSpace Software

2.1. Data Collection and Selection Criteria

This study utilized the Web of Science (WoS) database, maintained by the Institute for Scientific Information in the United States, as the primary data source. To ensure comprehensive retrieval, search terms included "foreign language classroom anxiety", "second language classroom anxiety", "anxiety in foreign language classroom", and "anxiety in second language classroom". A subject-specific search was conducted within the SSCI and SSCI-EXPANDED sub-databases, covering the publication period from January 1, 2002, to December 31, 2024. Only journal articles and conference papers were considered. Following the initial retrieval, CiteSpace software was employed for deduplication, yielding a total of 659 articles. A subsequent manual screening process further refined the dataset, resulting in the selection of 123 relevant publications that met the study's inclusion criteria. The complete records and cited references from these papers were exported in plain text format. Titles, authors, abstracts, keywords, and reference information were saved as txt files for subsequent visual analysis using knowledge graph methodologies.

2.2. Research Methodology

This study employs bibliometric analysis and scientific mapping techniques, utilizing CiteSpace 6.3.R1 as the primary analytical tool to examine foreign language classroom anxiety (FLCA) research published between 2002 and 2024. CiteSpace 6.3.R1, developed by Professor Chaomei Chen, is a widely recognized software for information visualization (Mao, 2022). It enables temporal and dynamic visual analysis through scientific knowledge graph representation, assisting researchers in identifying emerging trends and core topics within a given field.

Through CiteSpace-generated visual representations, this study conducts analyses on keyword co-occurrence, author and countries collaboration networks offering insights into the current research landscape and future directions in FLCA scholarship. Specifically, the aim of this research is to explore the following research questions:

- What is the distribution of FLCA-related publications over time?
- Which countries have contributed most significantly to FLCA research?
- What are the most highly cited publications in FLCA studies?
- What are the key research hotspots in FLCA scholarship?

3. Results of the Bibliometric Analysis

3.1. Publication Trends

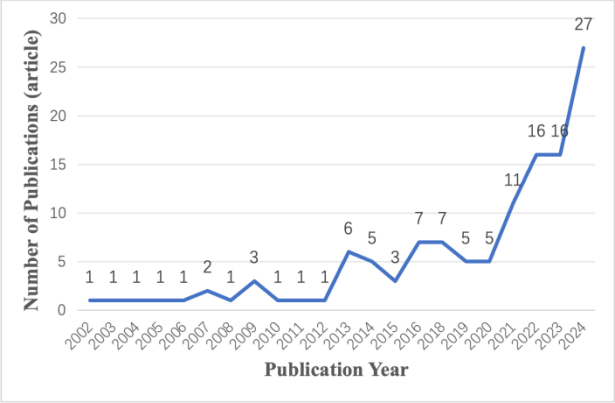


Figure 1. Annual publications in the research literature

Using CiteSpace and Excel, this study mapped the publication trends in foreign language classroom anxiety (FLCA) research since 2002. As illustrated in Figure 1, the evolution of research in this field exhibits fluctuations and can be broadly categorized into three distinct phases: (1) Slow Start Phase (2002–2007). During this initial stage, the number of annual publications remained exceptionally low, averaging only one article per year, except for 2006, when two articles were published. This indicates that FLCA had not yet emerged as a widely recognized research topic, and scholarly interest was still in the exploratory phase with limited research output. (2) Stable Fluctuation Phase (2008–2013). Although academic interest in FLCA increased, research output remained relatively modest. In 2008, three articles were published, and the number fluctuated between one and three in subsequent years, reaching six in 2013. This trend suggests that while scholarly engagement in FLCA was growing, research expansion remained gradual and inconsistent, with no large-scale surge observed. And (3) Rapid Growth Phase (2014–2024). A notable increase in research output began in 2014, when five articles were published, followed by fluctuations in subsequent years. Between 2015 and 2021, annual publication counts ranged from three to seven articles. However, a substantial rise was observed from 2022 onward, with 11 articles published in 2022, 16 in 2023, and a peak of 27 articles in 2024. This surge indicates growing academic recognition of FLCA, likely driven by advancements in language education research and evolving pedagogical needs. The increasing number of publications highlights FLCA's emerging significance and suggests sustained interest in this domain.

3.2. Top Frequency Co-citation References

Citation frequency is a crucial indicator of the academic influence and impact of scholarly contributions. To examine the most influential works in foreign language classroom anxiety (FLCA) research, this study identified the six most frequently cited articles in the field, as summarized in Table 1. Analyzing these studies provides valuable insights into the predominant research directions and emerging trends in FLCA scholarship.

Table 1. Top 6 Top frequency Co-citation references

No.	Article	Author	Citation frequency	Publication year
1	How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners?	Jiang, Y. & Dewaele, J. M	299	2019
2	Does the effect of enjoyment outweigh that of anxiety in foreign language performance?	Dewaele, J. M. & Alfawzan, M	256	2018
3	Recasts, language anxiety, modified output, and L2 learning	Sheen, Y.	197	2008
4	Effects of Language Anxiety on Three Proficiency-Level Courses of Spanish as a Foreign Language	Marcos Llinás, M. & Garau, M. J.	163	2009
5	An exploration of speaking-in-class anxiety with Chinese ESL learners	Mak, B	162	2011
6	The effect of classroom emotions, attitudes Toward English, and teacher behavior on willingness to communicate among English foreign language learners	Dewaele, J. M.	147	2019

The most highly cited study by Jiang et al. (2019) investigated the relationship between Foreign Language Enjoyment (FLE) and FLCA among Chinese undergraduate students learning English as a Foreign Language (EFL). Their findings revealed that FLE was significantly higher than FLCA in English classrooms; however, the overall anxiety level among Chinese learners remained higher than that of an international sample. A moderate negative relationship was identified between FLE and FLCA, indicating a substantial shared variance. While gender did not have a substantial influence on either construct, FLE was found to be predominantly influenced by teacher-related factors, such as interactions with instructors and peers. In contrast, FLCA was largely driven by learners' internal psychological variables and was more closely associated with self-oriented concerns. The second most commonly referenced study, carried out by Dewaele and Alfawzan (2018), examined how Foreign Language Enjoyment (FLE) and Foreign Language Classroom Anxiety (FLCA) impact language performance. Their results showed that FLE exerted a more significant positive influence on academic results, emphasizing the role of instructional methods in molding learners' emotions and study-related choices. Sheen (2008) investigated the link between anxiety and recasts during language learning. The study found that learners with lower anxiety levels showed better performance in speeded dictation and writing posttests, generating more revised output and self-corrections. These results imply that varying anxiety levels may affect how effective corrective feedback is in the process of language acquisition. Marcos (2009) investigated anxiety levels among 134 Spanish language learners, using questionnaire data. The study found that advanced learners reported higher anxiety levels; however, high anxiety did not necessarily correspond to lower academic performance. The majority of learners exhibited moderate anxiety, which appeared to have no significant impact on their grades. This suggests that a certain level of anxiety may serve as a motivational factor for language learning rather than a hindrance. Mak (2011) applied questionnaire-based factor analysis to a sample of 313 Hong Kong university students, identifying five key factors contributing to speaking anxiety in the classroom. The study emphasized the impact of variables such as unprepared speaking and teacher feedback, offering pedagogical implications for anxiety reduction in language instruction. Lastly, Dewaele (2019) examined the relationship between FLCA and communication willingness among 210 Spanish EFL learners. Their results indicated that FLCA negatively predicted learners' willingness to communicate, whereas FLE and teachers' use of the target language had a positive influence, underscoring the importance of instructional strategies that foster enjoyment and engagement.

3.3. Author Collaboration Networks

Using CiteSpace software, this study conducted an author-based network analysis to visualize the co-author relationships in foreign language emotion research, as illustrated in Figure 3. The co-author network provides valuable insights into the key researchers and prominent research teams who contribute to the area. In this network visualization, node size and label prominence correspond to each author's number of publications, while connecting lines show collaborative links. Thicker lines indicate stronger and more frequent cooperative efforts among scholars. Figure 3 reveals multiple sub-network structures, reflecting active scholarly communication and collaboration within the field. Notably, the network structure led by Dewaele stands out as a particularly significant cluster. Authors with high publication output have established stable collaboration networks, suggesting that FLCA research has achieved a certain level of academic maturity. The presence of these structured networks indicates a strong emphasis on academic exchange and cooperative research efforts among scholars. Dewaele, as the most prolific author in foreign language emotion research, has contributed sixteen publications to this domain. His extensive work and collaborative engagements highlight his influential role in shaping FLCA scholarship and advancing theoretical and empirical studies in the field.

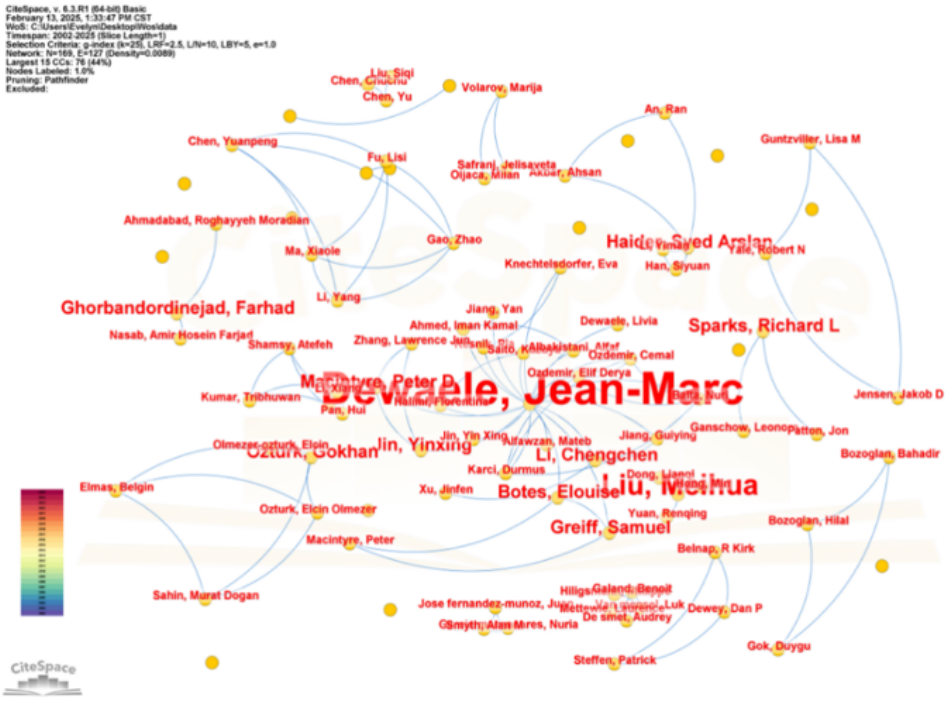


Figure 2. A visualization of the author collaboration network

Table 2. Top 10 authors

Ranking	Authors	Count	Centrality	Year
1	Dewaele, Jean-Marc	16	0.03	2013
2	Liu, Meihua	5	0	2021
3	Kim, Sung-Yeon	2	0	2009
4	Ozturk, Gokhan	2	0	2021
5	Greiff, Samuel	2	0	2022
6	Botes, Elouise	2	0	2022
7	Sparks, Richard L	2	0	2007
8	Park, Gi-Pyo	2	0	2013
9	Dordinejad, Farhad	2	0	2013
10	Haider, Syed Arslan	2	0	2022

3.4. Country Collaboration Networks

This study examines international collaboration in FLCA research, focusing on cooperative relationships among countries and the distribution of key contributors. The analysis of country-level collaboration reveals a network comprising 89 nodes and 145 links, as illustrated in Figure 3. Additionally, Table 3 presents the number of publications and betweenness centrality scores for the five most research-productive countries. The thickness and density of the links connecting nodes indicate the strength of collaborative ties among nations, with China demonstrating the most extensive cooperation with other countries. China holds a dominant position in FLCA research, with 44 publications - significantly surpassing other contributors. The United States and the United Kingdom follow as the second and third most productive countries, respectively. South Korea, ranked fourth, and Spain, ranked fifth in terms of publication output, also emerge as notable contributors within the collaboration network.

This study examines international collaboration in FLCA research, focusing on cooperative relationships among countries and the distribution of key contributors. The analysis of country-level collaboration reveals a network comprising 89 nodes and 145 links, as illustrated in Figure 3. Additionally, Table 3 presents the number of publications and betweenness centrality scores for the five most research-productive countries. The thickness and density of the links connecting nodes indicate the strength of collaborative ties among nations, with China demonstrating the most extensive cooperation with other countries. China holds a dominant position in FLCA research, with 44 publications - significantly surpassing other contributors. The United States and the United Kingdom follow as the second and third most productive countries, respectively. South Korea, ranked fourth, and Spain, ranked fifth in terms of publication output, also emerge as notable contributors within the collaboration network.

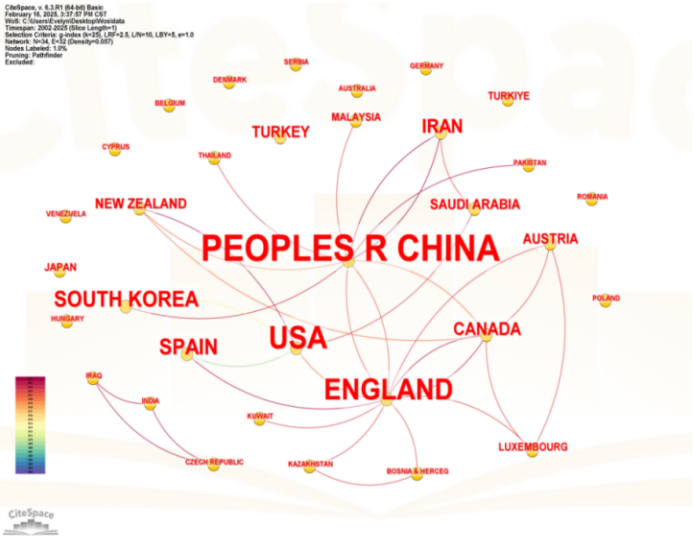


Figure 3. A visualization of the country collaboration network

Table 3. Top 5 most productive countries for foreign language classroom anxiety research

Ranking	Country	Count	Centrality	Year
1	People's Republic of China	44	0.22	2011
2	USA	23	0.06	2002
3	England	19	0.23	2013
4	South Korea	8	0.01	2009
5	Spain	8	0	2009

3.5. Analysis of Research Hotspots

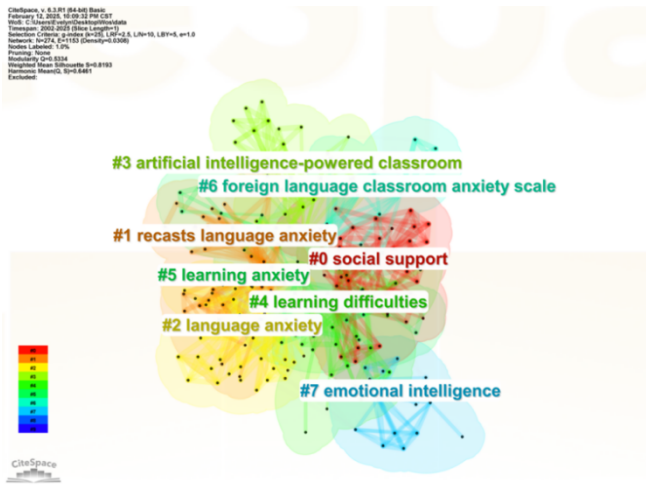


Figure 5. Keywords cluster analysis and visualization mapping

As crucial markers for encapsulating the central themes of academic papers, keywords play a vital role in scholarly communication. By conducting keyword analysis with CiteSpace, researchers can effectively identify emerging research hotspots and thematic evolution within a discipline. CiteSpace assesses the quality of clustering based on network architecture and cluster distinctiveness, relying on two key metrics: the modularity value (Q) and the average silhouette value (S). According to Chen et al. (2015), a Q value above 0.3 signifies a meaningful clustering structure, an S value over 0.5 indicates reasonable clustering validity, and an S value surpassing 0.7 confirms the robustness of the clustering results.

In this study, the keyword clustering module (Q) is 0.5334, while the average silhouette value (S) for the cluster is 0.6461. These metrics indicate that the clusters exhibit substantial internal connectivity and demonstrate strong thematic coherence. The co-cited literature keyword clustering knowledge map generated by CiteSpace identifies eight primary co-cited literature network clusters: #0 Social Support, #1 Recasts and Language Anxiety, #2 Language Anxiety, #3 Artificial Intelligence-Powered Classrooms, #4 Learning Difficulties, #5 Learning Anxiety, #6 Foreign Language Classroom Anxiety Scale, and #7 Emotional Intelligence (see Figure 5). Overall, FLCA research can be categorized into three primary thematic areas: (1) Factors influencing anxiety in foreign language classrooms, (2) Measurement scales for assessing foreign language classroom anxiety, and (3) The impact of anxiety on language learning outcomes. These findings offer valuable insights into prevailing research trends and emerging directions in FLCA scholarship.

3.6. Factors Influencing FLCA

The factors contributing to FLCA can be categorized into three primary dimensions: (1) the relationship between FLCA and individual intrinsic factors, (2) the influence of external environmental variables on anxiety levels, and (3) the impact of FLCA on academic achievement.

3.6.1. Individual Intrinsic Factors

Most studies examine the interplay between emotions and internal psychological traits. For instance, Dewaele (2013) explored the relationship among personality traits

(Psychoticism, Extraversion, Neuroticism), language proficiency, and FLCA in adult multilinguals. Results showed a notable association between FLCA and Neuroticism, with anxiety levels staying consistent across various languages. Moreover, multiple studies have examined the relationship between trait emotional intelligence and language anxiety.

Park and Brian (2013) employed the Foreign Language Classroom Anxiety Scale (FLCAS) to examine gender differences among Korean university students learning English. Their findings revealed that female students exhibited higher anxiety levels despite achieving better grades. And Teachers need to be aware of it and create supportive environments to reduce anxiety and encourage participation. Li (2019) reported a moderate association between students 'emotional intelligence, foreign language enjoyment (FLE), and foreign language anxiety (FLA). It is feasible for teachers to design intervention programs anchored in emotional intelligence, such as integrating the "ARGUER" positive psychology model with diary reflection activities. These initiatives enable students to enhance emotional awareness, recognition, and regulation skills while cultivating a positive emotional climate in the classroom. Chen et al. (2021) investigated the interconnections among trait emotional intelligence (EI), foreign language learning anxiety (FLA), and foreign language learning enjoyment (FLE) in foreign language oral classroom. The research findings indicated significant correlations among these three factors. Specifically, trait emotional intelligence had a stronger predictive effect on learning anxiety than on learning pleasure. And the researchers proposed that teachers ought to place great emphasis on cultivating students' characteristic emotional intelligence. For example, provide more targeted support and guidance for students with high anxiety to help them overcome anxiety disorders.

3.6.2. External Environmental Factors

Increasing research attention has focused on the relationship between FLCA and external influences. Ewald (2007) conducted a qualitative study on anxiety experiences among advanced Spanish learners, demonstrating that classroom dynamics - such as peer abilities, teacher behaviors, and fear of mistakes - contributed to heightened anxiety levels. The study emphasized the teacher's role in either mitigating or exacerbating student anxiety, advocating for a supportive classroom environment. Mak (2011) investigated classroom speech anxiety among Chinese ESL learners in Hong Kong, identifying five key contributors: speech anxiety, fear of negative evaluations, discomfort in speaking with native English speakers, negative attitudes toward the curriculum, and fear of academic failure. Additionally, classroom practices such as inadequate preparation time, immediate correction in front of peers, and restrictive language policies - were found to intensify anxiety levels.

Resnik et al. (2023) conducted a comparative study on the disparities in foreign language classroom anxiety (FLCA) experienced by students in online English classes before and during the epidemic. The results indicated that students' FLCA levels in online classes were remarkably lower compared to those in offline classes. Additionally, the sources of anxiety in online classes diverged from those in offline settings. Factors such as the utilization of technology and the ambiguity of course requirements were identified as the main contributing elements. Therefore, this study suggests that foreign language teaching should address the varying anxiety triggered by shifts in instructional modes. For offline teaching, cultivating an inclusive atmosphere is key to reducing students' anxiety during class discussions and peer comparisons. Timely feedback can also ease their concerns about understanding course content. By first addressing technical issues in the online teaching, teachers can then systematically outline course expectations and assessment criteria. Enhancing interaction through group discussions and peer collaboration further alleviates students' anxiety about technological unfamiliarity and ambiguous task guidelines.

A study by Wu et al. (2024) explored how teacher traits affect learners' foreign language learning enjoyment (FLE) and classroom anxiety (FLCA). The results indicate that learners taught by male teachers exhibit higher levels of FLE compared to those instructed by female teachers. Teacher characteristics demonstrate a stronger predictive capacity for FLE than for FLCA. Among these characteristics, teachers' positive emotions emerge as the most significant predictor of learners' FLE, whereas teachers' strictness is identified as having the most substantial influence on learners' FLCA. Lan (2024) found that both in-class flip instruction (IFI) and conventional flip instruction (CFI) can improve students' oral expression ability in business English more effectively. Among these two flipped teaching models, IFI has a better effect, which can reduce students' anxiety levels and improve students' participation and satisfaction. Teachers are advised to proactively explore and implement innovative teaching models like the flipped classroom, with particular emphasis on the IFI model. By optimizing the design of classroom activities and rationalizing time allocation, educators can provide more timely feedback and personalized support to students. This approach not only helps alleviate learning anxiety effectively but also enhances students' learning motivation and academic outcomes.

Regarding the reduction of students' anxiety in foreign language classes, Chinese English teachers in online classroom settings primarily employ five key strategy categories to alleviate learners' anxiety. These include fostering a non-threatening environment, boosting language learners' autonomy, involving students in learning tasks, increasing pair or group work, and attending to group dynamics Moreover, experienced and qualified teachers use these coping strategies more frequently than inexperienced and uncertified teachers. Liu (2023) found that interfere with cognitive performance the Hybrid SCMC (BYOD) model reduces student 'anxiety levels and is superior to the traditional FTF teaching and pure SCMC models in terms of interaction and learning experience.

3.7. Assessment of FLCA

The Foreign Language Classroom Anxiety Scale (FLCAS) is widely recognized as a reliable instrument for measuring FLCA. Empirical studies frequently employ questionnaires, assessments, and structured interviews to evaluate anxiety measurement tools across various contexts.

Despite its increasing use, the Short Foreign Language Classroom Anxiety Scale (S-FLCAS) required further validation. To address this, Botes et al. (2022) examined its reliability and validity using a sample of 370 language learners. Through exploratory and confirmatory factor analyses, they confirmed the S-FLCAS's single-factor structure, acceptable internal consistency, and evidence of convergent and discriminant validity.

In China, Dong and Huang (2023) set out to assess the psychometric properties—specifically reliability and validity—of the Chinese adapted version of the Short Foreign Language Classroom Anxiety Scale (S-FLCAS) in evaluating Foreign Language Classroom Anxiety (FLCA) among Chinese college students. The findings revealed that the Chinese version of S - FLCAS emerged as a reliable and valid unidimensional measurement tool. It is well - suited for gauging FLCA within the context of Chinese universities. And the scores are comparable among different genders and second language learner groups.

Öztürk et al. (2022) sought to integrate four skill-based foreign language anxiety scales into a unified framework and assess its predictive validity. Their study involved 385 Turkish EFL learners and employed five distinct anxiety scales. Structural equation modeling and confirmatory factor analysis demonstrated that the four skill-based anxieties could be merged into a single measure (SB-FLLAS), which effectively explained 88% of the variance in FLCA as assessed by FLCAS.

3.8. Impact of FLCA on Learning Outcomes

Empirical research predominantly indicates a negative correlation between FLCA and language performance (Horwitz et al., 1986). Ghorban Dordinejad and Nasab (2013) investigated FLCA as a mediating factor between perfectionism and English achievement among Iranian high school students. Their findings confirmed that FLCA was inversely correlated with English performance but positively associated with perfectionism. However, anxiety did not moderate the relationship between perfectionism and academic success.

During the COVID-19 pandemic, Liu (2023) examined the interconnections among FLCA, learning styles, and English achievement among 691 Chinese university students. Their research revealed a negative correlation between FLCA and academic performance, whereas learning styles showed a positive association with achievement. While gender and academic discipline influenced learning style preferences, they did not significantly affect FLCA levels. In foreign language teaching, teachers can foster a non-threatening online classroom environment by establishing social learning groups and integrating formative assessment with self-assessment. Granting students class autonomy, incorporating digital games, and teaching self-study skills can enhance their learning independence. For classroom activities, involving students in syllabus design, increasing collaborative tasks like group projects and online discussions, and leveraging peer interaction can reduce anxiety.

Some scholars explored the influences of learning motivation, anxiety and learning strategies on the Chinese academic performance of adult Chinese learners in Thailand during online learning. Anxiety is the most stable factor affecting learners' self-assessed Chinese proficiency and shows a negative prediction. Learning strategies have a certain positive predictive effect. The predictive effect of motivation is not significant. (Xu et al., 2022) In foreign language education, teachers should prioritize addressing learners' anxiety by creating a supportive environment and incorporating local culture to ease tension. It is also important to systematically train learners in self-regulated learning strategies like goal-setting and time management.

While FLCA research has historically relied on cross-sectional studies, an increasing number of longitudinal investigations provide deeper insights into temporal variations in anxiety. Sparks and Ganschow (2007) conducted a 10-year longitudinal study on 54 learners of Spanish, French, and German, finding that lower FLCA scores correlated with stronger native language skills and improved foreign language performance. Elahi Shirvan and Taherian (2018) applied potential growth curve modeling (LGCM) and data triangulation to examine changes in FLE and FLCA among university students over a semester. Their findings revealed that while FLE increased and FLCA decreased, initial anxiety levels did not predict long-term fluctuations.

4. Future Research Directions and Critical Analysis

Research visualization analysis indicates a growing academic focus on foreign language classroom anxiety (FLCA), with international journals increasingly publishing work in this domain. Both research themes and methodologies have evolved substantially, contributing to a more sophisticated understanding of FLCA. The progression of research in this field is primarily reflected in three key dimensions: research perspectives, research content, and research methodologies.

Early studies on FLCA primarily centered on language acquisition and pedagogical methods. However, as applied linguistics and second language acquisition research have advanced, interdisciplinary influences—particularly from cognitive science and psychology—have become integral to shaping the conceptual framework of FLCA. Increasingly, research in this field draws from cognitive psychology, demonstrating a clear trend toward integration with psychological theories.

Regarding research topics, contemporary studies emphasize diverse emotional experiences, examining the emotional states of language learners across different linguistic and cultural backgrounds, as well as various factors contributing to these emotions. Despite considerable scholarly output, systematic and multi-dimensional investigations remain relatively scarce. Additionally, mediating variables influencing FLCA are

often fragmented, limiting comprehensive insights into the complexity of foreign language anxiety. Another critical gap in existing research is the predominant focus on external anxiety-reduction strategies, with relatively few studies addressing self-regulation strategies rooted in learners' internal coping mechanisms. Future research could further examine self-regulation approaches to enhance understanding of how learners autonomously manage FLCA.

Various intrinsic and extrinsic factors contribute to FLCA, yet existing methodological approaches predominantly employ qualitative and quantitative techniques. Traditional data analysis relies on t-tests and analysis of variance (ANOVA) to examine FLCA under varying conditions (e.g., Park et al., 2013; Li & Xu, 2019). More recently, scholars have incorporated advanced statistical models, such as mixed-effects models and structural equation modeling, to explore interactions among multiple influencing variables (e.g., Öztürk et al., 2022). While efforts to develop innovative methodologies continue, empirical research in FLCA remains predominantly quantitative. Future studies should integrate a broader range of methodological approaches - such as longitudinal, experimental, and mixed-method designs to provide a more comprehensive perspective on FLCA dynamics.

Current FLCA research predominantly examines adult learners, particularly those from diverse linguistic and cultural backgrounds. However, research involving primary and secondary school learners remains limited, despite the relevance of foreign language anxiety at earlier stages of education. Similarly, the classroom anxiety experienced by foreign language teachers has not received sufficient scholarly attention. Future studies should broaden the scope of investigation to include younger learners, incorporating developmental perspectives to assess how FLCA manifests across different age groups. Additionally, expanding research on instructors' anxiety would offer deeper insights into the reciprocal emotional dynamics between teachers and students.

Most existing FLCA studies adopt cross-sectional designs, limiting the ability to track diachronic variations in foreign language anxiety over time. Given that language learning represents a dynamic process, future research ought to give priority to longitudinal studies for a deeper comprehension of the changing characteristics of FLCA. Although some longitudinal investigations have emerged, their number and scope remain restricted. For instance, Mariusz (2018) presents a valuable example of paired production studies, yet existing literature remains scarce, and the temporal range of analysis is relatively narrow. Long-term diachronic studies extending over several years could provide deeper insights into fluctuations in FLCA, offering more robust findings for pedagogical intervention.

5. Conclusion

This study presents a comprehensive visual analysis of foreign language classroom anxiety (FLCA) research published between 2002 and 2024. A systematic review of literature from the Web of Science database was conducted to identify research hotspots and emerging trends in the field. Findings indicate that FLCA research has undergone a three-stage developmental trajectory, beginning with an initial period of slow growth, followed by a phase of fluctuating expansion, and culminating in rapid scholarly advancement. Key research themes include factors influencing FLCA, the practical application of anxiety measurement scales, and the impact of classroom anxiety on learning outcomes. Methodologically, FLCA studies predominantly employ quantitative and qualitative approaches. To further advance this field, future research should integrate diverse theoretical frameworks and innovative methodologies, fostering a more comprehensive and nuanced understanding of FLCA dynamics.

In foreign language teaching, educators should address student individual differences, especially the diverse impacts of emotional intelligence facets. By doing so, they can design targeted teaching strategies to foster comprehensive foreign language proficiency. Considering the negative impacts of foreign language classroom anxiety, teachers should create a friendly and supportive learning environment to ease such anxiety. Providing positive feedback, using encouraging language, and setting realistic expectations can build students' confidence and promote active classroom participation.

To effectively reduce anxiety, teachers can adopt strategies such as creating a low-pressure learning environment, promoting student autonomy, designing varied activities, facilitating group collaboration, and monitoring classroom dynamics. Moreover, educators should acknowledge and differentiate for differences in gender, academic backgrounds, and learning styles (e.g., visual, auditory, and kinesthetic). Using varied teaching materials and activities to meet these diverse needs can enhance instructional effectiveness. Teachers should also monitor students' mental health, promptly identifying and providing intervention for those with severe anxiety or learning challenges. Offering psychological counseling, academic guidance, and emotional support helps students overcome obstacles and maintain a positive learning attitude. Lastly, novice and inexperienced teachers, in particular, require targeted training in skills like managing foreign language classroom anxiety, identifying diverse learning styles, and implementing differentiated instruction.

Mengyao Xun is a graduate student in English Language and Literature at School of Foreign Studies, Jiangsu Normal University, China. She is interested in second language acquisition and foreign language teaching.
Email: mengyao_xun@163.com

Ran Wei (Corresponding Author) is an associate professor at School of Foreign Studies, Jiangsu Normal University, China. Her research interests focus on second language acquisition, foreign language teaching and psycholinguistics. Her work appears in research journals such as *Assessing Writing*, *Language Teaching Research*, *Interactive Learning Environments*, *Frontiers in Psychology*.
Email: weiran@jsnu.edu.cn

References

Botes, E., van der Westhuizen, L., Dewaele, J.-M., MacIntyre, P., & Greiff, S. (2022). Validating the short-form foreign language classroom anxiety scale. *Applied Linguistics*, *43*(5), 1006-1033. https://doi.org/10.1093/applin/amac018

Brown, H. D. (1973). Affective variables in second language acquisition. *Language Learning*, *23*(2), 231-244.

Chen, T.-Y., & Chang, G.B.Y. (2004). The relationship between foreign language anxiety and learning difficulties. *Foreign Language Annals*, *37*(2), 279-289. https://doi.org/10.1111/j.1944-9720.2004.tb02200.x

Chen, Y., Chen, C. M., Liu, Z. Y., Hu, Z. G., & Wang, X. W. (2015). CiteSpace zhishi tupu de fangfalun gongneng [The methodological functions of the CiteSpace knowledge mapping]. Studies in Science of Science, *33*(2), 242–253.

Chen, Z., Zhang, P., Lin, Y., & Li, Y. (2021). Interactions of trait emotional intelligence, foreign language anxiety, and foreign language enjoyment in the foreign language speaking classroom. *Journal of Multilingual and Multicultural Development*, *45*(2), 374-394. https://doi.org/10.1080/01434632.2021.1890754

Dewaele, J.-M. (2013). The link between foreign language classroom anxiety and psychoticism, extraversion, and neuroticism among adult Bi- and multilinguals. *Modern Language Journal*, *97*(3),670-684. https://doi.org/10.1111/j.1540-4781.2013.12036.x

Dewaele, J.-M. (2019). The effect of classroom emotions, attitudes toward English, and teacher behavior on willingness to communicate among English foreign language learners. *Journal of Language and Social Psychology*, *38*(4), 523-535. https://doi.org/10.1177/0261927X19864996

Dewaele, J.-M., & Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance? *Studies in Second Language Learning and Teaching*, *8*(1), 21-45. https://doi.org/10.14746/ssllt.2018.8.1.2

Dong, Y., & Huang, J. (2024). Validating the Chinese version of short-form foreign language classroom anxiety scale. *Psychology in the Schools*, *61*(8), 3263-3283. https://doi.org/10.1002/pits.2321

Dordinejad, G. F., & Nasab, A.H.F. (2013). Examination of the relationship between perfectionism and English achievement as mediated by foreign language classroom anxiety. Asia Pacific Education Review, 14(4), 603-614. https://doi.org/10.1007/s12564-013-9286-5

Elahi Shirvan, M., & Taherian, T. (2018). Longitudinal examination of university students' foreign language enjoyment and foreign language classroom anxiety in the course of general English: Latent growth curve modeling. *International Journal of Bilingual Education and Bilingualism*, *24*(1), 31-49. https://doi.org/10.1080/13670050.2018.1441804

Ewald, J.D. (2007). Foreign language learning anxiety in upper-level classes: Involving students as researchers. *Foreign Language Annals*, *40*(1),122-142. https://doi.org/10.1111/j.1944-9720.2007.tb02857.x

Gardner, R. (1985). Social psychology and second language learning: *The Role of Attitude and Motivation*. Edward Arnold.

Horwitz, E. K. (2016). Factor structure of the foreign language classroom anxiety scale. *Psychological Reports*, *119*(1),71-76. https://doi.org/10.1177/0033294116653368

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125-132.

Horwitz, E.K. (2016), Reflections on Horwitz (1986), "Preliminary evidence for the validity and reliability of a foreign language anxiety scale". *TESOL Q*, *50*(4) 932-935. https://doi.org/10.1002/tesq.295

Jiang, Y., & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System*, *82*(59), 13-25. https://doi.org/10.1016/j.system.2019.02.017

Kruk, M. (2018). Changes in foreign language anxiety: A classroom perspective. *International Journal of Applied Linguistics*, *28*(1), 31-57. https://doi.org/10.1111/ijal.12182

Lan, E. M. (2024). Unveiling the silent struggle: Investigating the effects of flipped classroom instruction models on business English oral presentation development through online learning: A case of learner engagement, emotions, and anxiety. *Educational Information Technology*, *29*(7), 23299-23328. https://doi.org/10.1007/s10639-024-12746-4

Li, C. C., Dewaele, J. -M., & Jiang, G. Y. (2020). The complex relationship between classroom emotions and EFL achievement in China. *Applied Linguistics Review*, *11*(3), 485-510. https://doi.org/10.1515/applirev-2018-0043

Li, C., & Xu, J. (2019). Trait emotional intelligence and classroom emotions: A positive psychology investigation and intervention among Chinese EFL learners. *Frontiers in Psychology*, *10*, 2453. https://doi.org/10.3389/fpsyg.2019.02453

Liu, M. (2023). English classroom anxiety, learning style and English achievement in Chinese university EFL students. *Sustainability*, *15*(18), 13697. https://doi.org/10.3390/su151813697

Liu, X. (2023). The effect of hybrid SCMC (BYOD) on foreign language anxiety and learning experience in comparison to pure SCMC and FTF communication. *Frontiers in Psychology*, *14*, 1172442. https://doi.org/10.3389/fpsyg.2023.1172442

Liu, Y., & Wang, J. (2023). Strategies for reducing EFL learners' foreign language anxiety in online classes: Investigating teachers' teaching credentials and experience. *Heliyon, 9*(7), e17579. https://doi.org/10.1016/j.heliyon.2023.e17579

Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System*, *39*(2), 202-214. https://doi.org/10.1016/j.system.2011.04.002

Mao, W. (2022). Jiaoyu zhengce yu woguo daxue

yingyuyanjiu fazhan de guanxi—Jiyu CiteSpace de zhishi tupu fenxi [The relationship between educational policies and the development of college English research in China: A knowledge graph analysis based on CiteSpace]. Foreign Language Research, (3), 92-97.

Marcos-Llinás, M., & Garau, M.J. (2009). Effects of language anxiety on three proficiency-level courses of Spanish as a foreign language. *Foreign Language Annals*, *42*(1),94-111. https://doi.org/10.1111/j.1944-9720.2009.01010.x

McCroskey, J. C. (1984). The communication apprehension perspective. In J. A. Daly & J. C. McCroskey (Eds.), *Avoiding communication: Shyness, reticence, and communication apprehension* (pp. 13–38). Sage.

Öztürk, G., Şahin, M. D., Öztürk, E. Ö., & Elmas, B. (2022). Modeling skill-based foreign language learning anxieties as a single construct and testing its predictive aspect on foreign language classroom anxiety. *Educación XX1*, *25*(2), 337-361. https://doi.org/10.5944/educxx1.30453

Pae, T.-I. (2013). Skill-based L2 anxieties revisited: Their intra-relations and the inter-relations with general foreign language anxiety. *Applied Linguistics*, *34*(2), 232–252. https://doi.org/10.1093/applin/ams041

Park, G.-P., & French, B. F. (2013). Gender differences in the foreign language classroom anxiety scale. *System*, *41*(2), 462-471. https://doi.org/10.1016/j.system.2013.04.001

Resnik, P., Dewaele, J.-M., & Knechtelsdorfer, E. (2023). Differences in the intensity and the nature of foreign language anxiety in in-person and online EFL classes during the pandemic: A mixed-methods study. *TESOL Journal, 57*(3), 618-642. https://doi.org/10.1002/tesq.3177

Sarason, I. (1984). Stress, anxiety, and cognitive interference: Reactions to tests. *Journal of Personality and Social Psychology, 46,* 929-938. https://doi.org/10.1037/0022-3514.46.4.929

Sheen, Y. (2008). Recasts, language anxiety, modified output, and L2 learning. *Language Learning*, *58*(4),835-874. https://doi.org/10.1111/j.1467-9922.2008.00480.x

Sparks, R.L., & Ganschow, L. (2007). Is the foreign language classroom anxiety scale measuring anxiety or language skills? *Foreign Language Annals*, *40*(2), 260-287. https://doi.org/10.1111/j.1944-9720.2007.tb03201.x

Steinberg, F., & Horwitz, E.K (1986). The effect of induced anxiety on the denotative and interpretative content of second language speech. *TESOL Quarterly*, *20*(1), 131-136.

Watson, D., & Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology, 33*, 448-457. https://doi.org/10.1037/h0027806

Wu, H. (2024). How teacher characteristics shape Chinese undergraduate EFL learners' enjoyment and anxiety. *Asia Pacific Journal of Education*, *43*(1), 1-14. https://doi.org/10.1080/02188791.2024.2385842

Xu, W., Zhang, H. W., Sukjairungwattana, P., & Wang, T. M. (2022). The roles of motivation, anxiety and learning strategies in online Chinese learning among Thai learners of Chinese as a foreign language. *Frontiers in Psychology*, *13*, 962492. https://doi.org/10.3389/fpsyg.2022.962492

Young, D. (1991). Creating a low-anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal*, *75*(4), 426. https://doi.org/10.1111/j.1540-4781.1991.tb05378.x

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2025 Xun and Wei. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

JLT Journal of Language Teaching

Peer-reviewed | Open Access | Google Scholar | MLA | Crossref | DOI

Call for Papers

Submit via https://jlt.ac/

Areas of Interest:

Language teaching intervention and experiments; Curriculum development; Language teacher education; Bilingual education; New technologies in language teaching; Testing, assessment, and evaluation; Educational psychology, and more.

We accept the following types of submission:

1. Research article: (6,000 to 8,000 words)
2. Review: (3,000 to 8,000 words)
3. Book review: (up to 3,000 words)
4. Features: (3,000 to 8,000 words)

Scan to submit your articles* & read more articles for free.

*Article Processing Charges Apply.

Contact: editor@jlt.ac

Pioneer Publications

ISSN (Online) 2770-4602