Literary Texts and Language Development: Piloting a Questionnaire Among Albanian Language and EFL Teachers in North Macedonia

Jehona Rushidi-Rexhepi[1], Jeta Hamzai[2], Bujar Adili[1]
[1]Faculty of Pedagogy, University of Tetova, North Macedonia
[2]Language Center, South East European University, North Macedonia
Email: bujar.adili@unite.edu.mk

Abstract: *This pilot study aimed to adapt and validate a questionnaire assessing teachers' perceptions of the role of literary texts in students' language development. The original instrument by Duncan and Paran (2017) was modified to suit the Albanian educational context and piloted with a diverse sample of 30 teachers in North Macedonia, including both Albanian Language and Literature and English as a Foreign Language (EFL) teachers. The 32-item instrument, organized into four theoretically grounded subscales, was evaluated using descriptive statistics and internal consistency analysis. Following expert review and item analysis, three items with weak psychometric performance were removed, resulting in a final 29-item version. The revised instrument demonstrated good internal consistency (Cronbach's α = .878), with subscale reliability coefficients ranging from .728 to .772. This pilot confirms the reliability and contextual relevance of the adapted questionnaire and supports its suitability for use in future large-scale research. Future studies should examine the instrument's construct validity through Confirmatory Factor Analysis (CFA) or Exploratory Factor Analysis (EFA), and explore potential differences between teacher groups or teaching contexts.*
Keywords: literary texts, language development, teacher perceptions, questionnaire validation, Albanian Language and EFL teachers

INTRODUCTION

The integration of literary texts into language education has received renewed scholarly attention for its capacity to enhance vocabulary, deepen grammatical awareness, and foster critical reading skills (Paran & Tupas, 2025; Sastre & García, 2022; Sulistiyo et al., 2022). Literature provides authentic, contextually rich input that promotes both emotional and cognitive engagement, key

drivers of meaningful language acquisition (Savvidou, 2020; Snow et al., 2021; Paran & Robinson, 2016). Moreover, literary texts facilitate the integrated development of reading, writing, speaking, and listening skills, while exposing students to diverse discourse styles and cultural frames (Carabantes & Paran, 2022; Paran, 2022; Duncan & Paran, 2017).

Teacher beliefs play a central role in shaping the interpretation and enactment of curricular materials, directly affecting classroom practices and student outcomes (Cunningham & Farmer, 2016; Li, 2020). However, there is limited evidence on how Albanian Language and Literature teachers, and EFL instructors, conceptualize literary texts as resources for language development, which this study aims to address.

Grounded in Duncan and Paran's (2017) framework, this pilot study initiates the development of a psychometrically sound instrument to investigate how Albanian Language and Literature and EFL teachers in North Macedonia perceive the role of literary texts in language education. The study seeks to generate foundational evidence for tool refinement and future application in curriculum and teacher development research.

LITERATURE REVIEW

Literature in Language Learning: Benefits and Debates
Literary texts have long held a valuable place in language education, primarily for their capacity to foster linguistic competence and critical engagement (Paran et al. 2020). Numerous studies underscore their role in enhancing vocabulary acquisition, syntactic awareness, and discourse competence (Paran, 2008; Hall, 2015; Duncan & Paran, 2017; Mart, 2018). Literature allows learners to encounter varied registers, idiomatic expressions, and rich lexical input in authentic contexts (Lazar, 2020; Ghosn, 2013). This immersion supports both receptive and productive language development across multiple skills (reading, writing, listening, and speaking) especially when texts are integrated through a task-based or communicative approach (Ghosn, 2018; Bland, 2013). Moreover, literature stimulates learner motivation by offering emotional resonance and narrative immersion, which help establish meaningful connections between language and personal experience (McRae, 2022). Stories humanize grammar and vocabulary, allowing abstract structures to be internalized more naturally (Bland, 2018). The affective and cognitive engagement triggered by narratives also supports deeper processing and retention (Van, 2009). However, debates remain regarding text complexity and accessibility for younger or lower-proficiency learners (Tomlinson, 2023; Mart, 2018). Yet, growing evidence advocates for carefully chosen and scaffolded literary texts that align with learners' linguistic and cultural readiness, making literature not only feasible but pedagogically enriching in both L1 and EFL contexts (Paran, 2008).

Teacher Beliefs and Perceptions of Literary Texts
Teachers' beliefs play a decisive role in shaping instructional decisions and pedagogical practices in literature teaching. According to Borg (2003), teacher cognition (what teachers know, believe, and think) profoundly influences how literature is integrated into language classrooms. Classroom applications of literature often reflect deeply held beliefs about language learning, rather than formal methodological training (Phipps & Borg, 2009). EFL teachers, for example, may focus

more on functional language use, whereas L1 teachers may prioritize literary analysis due to differing educational backgrounds and goals (Paran, 2008). Research indicates that teachers' attitudes toward literature impact their willingness to integrate it into the curriculum (Kim, 2004). Ghosn (2013) argues that when teachers view literature as a culturally enriching and cognitively engaging resource, they are more likely to use it meaningfully in class. Duncan and Paran (2017) found that even among similarly qualified teachers, beliefs about literature's role can vary widely and shape student experience in language learning.

Research Gaps and the Need for a Contextualized Instrument
Although teacher cognition in language teaching has been widely studied (Borg, 2003), research focusing on teachers' perceptions of literature as a language tool in Albanian-speaking regions, particularly in North Macedonia, remains scarce. While literature is often integrated into language curricula, little is known about how Albanian L1 and EFL teachers view its pedagogical value. Existing studies suggest that teachers' language background and training significantly influence their beliefs and practices (Duncan & Paran, 2017; Paran et al., 2020). However, such distinctions have not been systematically explored in the Albanian context. Moreover, despite North Macedonia's multilingual and diverse classrooms, the role of cultural or societal factors in shaping teachers' use of literature remains under-researched. These gaps highlight the need for localized studies that examine how literature is perceived and used by language educators in both L1 and EFL settings.

Accurate assessment of teacher perceptions regarding literature in language learning requires a context-sensitive, psychometrically sound tool (Li et al., 2022). Duncan and Paran (2017) developed a robust instrument for evaluating English teachers' beliefs in the UK, yet no equivalent exists for Albanian-speaking educational contexts. Given linguistic and curricular differences, this study adapts their framework to suit Albanian Language and Literature and EFL teachers in North Macedonia. By refining the original instrument, the current study addresses a regional research gap and provides the foundation for empirical inquiry into teacher beliefs, ensuring cultural and instructional relevance in a previously understudied context.

Purpose of the Study
This pilot study aimed to adapt and preliminarily validate a questionnaire for use with Albanian Language and Literature teachers and English as a Foreign Language (EFL) teachers. The specific objectives were:

– To assess the content validity of the adapted instrument;
– To examine the internal consistency (reliability) of its subscales;
– To refine the instrument for future large-scale implementation.

METHODOLOGY

Instrument Development and Adaptation
The instrument used in this study was based on the questionnaire developed by Duncan and Paran (2017) to investigate teachers' views on the use of literature in secondary English classrooms in

the UK. Their original instrument aimed to capture attitudes regarding literature's contribution to linguistic development and intercultural understanding. For the current study, the tool was systematically adapted to fit the linguistic, curricular, and cultural realities of Albanian Language and Literature (L1) and English as a Foreign Language (EFL) instruction in North Macedonia.

The adaptation process followed methodological guidelines for instrument modification (DeVellis, 2016; Boateng et al., 2018). Additional methodological guidance on questionnaire construction, Likert-scale formatting, and demographic design was informed by established sources in educational research (Cohen et al., 2018; Dörnyei & Taguchi, 2010). First, a comparative analysis of the UK and North Macedonian educational settings was conducted to determine which items retained conceptual relevance. Items addressing ethnic diversity or broad multicultural themes were excluded as beyond the study's scope. However, two items on intercultural awareness and cultural comparison were retained for their relevance to literature's pedagogical and linguistic functions. The instrument's focus thus centered on teachers' perceptions of literature's linguistic, motivational and instructional value.

The initial draft contained 32 items, categorized into four subscales derived from the original structure: (1) Language Suitability and Challenge, (2) Content and Literary Value, (3) Teaching Practice and Confidence, and (4) Beliefs About Student Outcomes (Language/Culture). For clarity and ease of reference, these subscales are hereafter referred to by the following abbreviated labels: (1) Linguistic Fit, (2) Text Relevance, (3) Teacher Practice, and (4) Student Outcomes. These shorter names preserve the core conceptual focus of each domain while enhancing readability in tables, analyses and discussion.

Items were measured on a 6-point Likert scale ranging from 1 ("Strongly disagree") to 6 ("Strongly agree"). A 6-point format was intentionally selected to avoid a neutral midpoint, thereby encouraging participants to take a clearer stance on each item. This forced-choice design is supported in educational research as a strategy to reduce central tendency and social desirability bias, while promoting more differentiated and interpretable response patterns (Lacko et al., 2022; Dörnyei & Taguchi, 2010; Brown, 2011). To minimize acquiescence bias, three negatively worded items were included, of which only one remained in the final version following item analysis.

Subsequent revisions were informed by expert feedback and pilot testing, ultimately producing a refined 29-item instrument. This version was designed to ensure contextual relevance, conceptual clarity, and empirical reliability in measuring both L1 and EFL teachers' perceptions of literature use in language instruction.

Content Validity
Content validity was assessed through expert review and participant feedback. Three university-level experts in literature and pedagogy evaluated the initial instrument, focusing on item clarity, cultural relevance, and alignment with the intended constructs. Their feedback led to minor linguistic adjustments to reflect terminology and phrasing commonly used by Albanian Language and Literature and EFL teachers in North Macedonia.

In addition, pilot participants were invited to comment on the clarity and relevance of the items. Their responses informed the refinement process, resulting in rewording of several items and the elimination of those considered ambiguous or redundant. This iterative process helped ensure that the adapted instrument captured perceptions accurately within the target educational and cultural context.

Participants

The pilot study involved 30 teachers of Albanian Language and Literature and English as a Foreign Language (EFL) from primary and secondary schools in North Macedonia. Participants were selected through purposive sampling to ensure variation across key demographic and professional variables. The sample included 25 female and 5 male teachers; 12 had over 20 years of teaching experience, while 18 had less than 20 years. Regarding educational qualifications, 12 held MA or PhD degrees, and 18 held bachelor's degrees.

Participants represented both rural (n = 13) and urban (n = 17) school settings. Half of the teachers taught Albanian Language and Literature, and half taught EFL. In terms of school level, 18 worked in primary schools and 12 in secondary schools. With respect to professional development, 13 participants reported having received formal training in literature teaching, whereas 17 had not.

Procedure

Participants were recruited through professional networks and invited to participate voluntarily in the pilot study. The questionnaire was administered online, requiring approximately 12 minutes to complete. Participation was anonymous, and informed consent was obtained prior to completion. No incentives were offered for participation. The instrument was distributed and monitored by the researcher, who also handled all data collection and management. Once completed, responses were downloaded and entered into SPSS. Reverse-coded items were recoded prior to reliability analysis. All submitted responses were complete, with no missing data or technical issues reported.

RESULTS

Reliability Analysis

Initial reliability analysis was conducted on the original 32-item questionnaire adapted from Duncan and Paran (2017), which comprised four theoretical subscales. The full scale demonstrated reliable internal consistency (Taber, 2018), with a Cronbach's alpha of .849. To examine the internal consistency of each subscale in greater depth, item-level analyses were performed using SPSS. The subscale Teacher Practice showed not satisfactory internal consistency (α = .512; Taber, 2018), prompting further item-level analysis.

Given the insufficient internal consistency of Subscale 3, further diagnostics were conducted to identify specific items contributing to this issue. Based on corrected item-total correlations and Cronbach's alpha if item deleted, three items—Q5 ("I find it difficult to find new literary texts to teach"), Q7 ("I prefer to use texts I have already used successfully in the past"), and Q27 ("Students should be challenged by the language used in the literary text")—were identified as underperforming and removed from the scale. The item-total statistics revealed that three items

(Q5, Q7, Q27) had corrected item-total correlations below the acceptable threshold of .300 (Pallant, 2020), with some showing very low or negative correlations with the total scale (e.g., Q5: $r = –.226$; Q27: $r = –.021$), indicating misalignment with both the subscale and the underlying construct. After removing them, the revised Teacher Practice subscale consisted of five items (Q2, Q9, Q13, Q15, Q18) and demonstrated good reliability ($\alpha = .728$; Taber, 2018). All retained items showed corrected item-total correlations above .40, supporting their contribution to the subscale.

This improvement supports the decision to refine the subscale by excluding items that may have introduced measurement noise or conceptual inconsistency. Following the item removals, the overall internal consistency of the revised 29-item scale was reassessed. The final 29-item version of the instrument demonstrated reliable overall reliability, $\alpha = .878$ (Taber, 2018). It retained four subscales: Linguistic Fit (9 items), Text Relevance (9 items), Teacher Practice (5 items) and Student Outcomes (6 items). Item means ranged from 3.07 (Q27, removed) to 5.83 (Q4), and standard deviations ranged from approximately .46 to 1.73, indicating sufficient variability across responses. While SPSS issued a warning regarding the determinant of the covariance matrix—often related to multicollinearity or redundancy—this did not affect reliability interpretation and was partially addressed through item deletion.

Table 1 displays descriptive statistics (means and standard deviations) for each item on the revised scale, providing an overview of item-level responses.

Table 1. Item Statistics and Reliability for the 29-Item Literature Teaching Questionnaire

Items	Item Description (Shortened)	r (Item-Total)	M	SD	α if deleted
Item 1	Improves students' vocabulary	.397	5.40	.932	.875
Item 2	Easy to integrate into syllabus	.496	4.53	1.196	.872
Item 3	Texts support grammar learning	.348	4.73	1.437	.877
Item 4	Themes should be age-appropriate	.504	5.83	.461	.875
Item 5	Supports critical thinking	.372	5.43	1.135	.875
Item 6	Vocabulary within students' ability	.573	5.03	1.273	.870
Item 7	Enough time to explore texts	.687	4.10	1.373	.866
Item 8	Supports students' creative use	.338	5.40	1.037	.876
Item 9	Supports multiple language skills	.425	5.67	.661	.875
Item 10	Offers opportunities for cultural comparison	.731	5.03	1.426	.865
Item 11	I enjoy teaching literature	.489	5.50	.861	.873
Item 12	Reflects current language usage	.441	4.87	1.383	.874
Item 13	Confident teaching complex texts	.174	5.07	.828	.879
Item 14	Develops intercultural awareness	.561	5.17	1.262	.870
Item 15	Lends itself to teaching specific points	.548	5.07	1.143	.871
Item 16	Lacking training or guidance	.055	3.20	1.730	.889
Item 17	Suitable for classroom discussion	.330	5.57	.679	.876
Item 18	Learn new vocabulary	.553	5.63	.556	.874
Item 19	Grammar within students' ability	.346	5.07	1.388	.877
Item 20	Relevance to students' lives	.311	5.00	1.050	.877
Item 21	Overall linguistic suitability	.517	5.37	.964	.872
Item 22	Explores global/social issues	.537	5.17	.986	.872
Item 23	Increases student motivation	.660	5.23	.898	.870

Item 24	Shorter texts work better	.530	5.13	1.008	.872
Item 25	Linguistically challenging	.386	5.07	1.015	.875
Item 26	High literary quality	.470	5.00	1.313	.873
Item 27	Offers aesthetic/emotional value	.368	5.23	.971	.875
Item 28	Represents national and world literature	.214	5.47	.776	.878
Item 29	Supports specific language skills	.275	5.60	.621	.877

Note: The full instrument is provided in Appendix A.

While both Q13 ("I do not feel confident using literary texts in my classroom") and Q16 ("I feel confident choosing appropriate literary texts") showed relatively modest corrected item-total correlations (r = .191 and r = .239, respectively), several theoretical and practical considerations support their retention. These items capture teacher self-efficacy, a critical component of the "Teacher Practice" subscale. Confidence in selecting and using texts is directly aligned with the construct being measured and is strongly supported in the literature as a key variable influencing pedagogical effectiveness (Bandura, 1997; Borg, 2015). Q13 and Q16 offer complementary perspectives: Q13 reflects lack of confidence (negatively worded), while Q16 expresses confidence (positively worded). Together, they help avoid response bias and ensure construct balance within the subscale. Although the correlations were lower, both items exceed the minimum acceptable threshold of r = .15 (DeVellis & Thorpe, 2021), indicating a meaningful—albeit modest—contribution to the overall scale. Furthermore, their removal would not substantially improve subscale reliability. Given the relatively small sample size (N = 30), slight fluctuations in item performance are expected. These items may demonstrate stronger psychometric properties in larger-scale applications, and premature deletion could compromise content validity.

Subscale Reliability Summary
The refined instrument's four subscales each demonstrated acceptable to good internal consistency, with Cronbach's alpha values ranging from .728 to .772 (Taber, 2018). The highest subscale reliability was observed for Student Outcomes (α = .772), followed closely by Text Relevance (α = .736), Linguistic Fit (α = .730), and Teacher Practice (α = .728). These results suggest that the adapted instrument reliably captures teachers' perceptions across thematically coherent domains. Table 2 presents the final composition and internal consistency of each subscale and the overall instrument.

Table 2. Subscale Structure, Item Composition, and Internal Consistency Estimates

Subscale Name	Items	Count	α	M	SD
Linguistic Fit	3, 6, 12, 15, 19, 21, 24, 25, 29	9	.730	5.10	.655
Text Relevance	4, 5, 10, 17, 20, 22, 26, 27, 28	9	.736	5.29	.626
Teacher Practice	2, 7, 11, 13, 16	5	.728	4.48	.861
Student Outcomes	1, 8, 9, 14, 18, 23	6	.772	5.42	.629
Literature Teaching Questionnaire	29	29	.878	5.12	.518

Note. *M and SD represent subscale-level mean scores and standard deviations based on participant responses (N = 30).*

DISCUSSION

The adapted questionnaire demonstrated strong internal consistency and acceptable subscale reliability, confirming its suitability for use with both Albanian Language and Literature and EFL teachers in North Macedonia. The overall Cronbach's alpha of .878 and subscale alphas ranging from .728 to .772 suggest the instrument reliably captures teachers' perceptions across key domains. These results align with previous research emphasizing the importance of contextually grounded tools in educational settings (DeVellis & Thorpe, 2021; Boateng et al., 2018). Comparable reliability findings were reported in Duncan and Paran's (2017) pilot study, where most question batteries demonstrated good to excellent reliability. Specifically, out of 17 batteries, two achieved $\alpha \geq .90$ (excellent), seven had $.90 > \alpha \geq .80$ (very good), three had $.80 > \alpha \geq .70$ (good), three had $.70 > \alpha \geq .60$ (questionable), and two had $.60 > \alpha \geq .50$ (poor).

The pilot also reinforced the necessity of refining items that lack clarity or conceptual alignment. Three items were removed from the "Teacher Practice" subscale based on low corrected item-total correlations, improving its internal reliability. Notably, both positively and negatively worded items were retained to balance acquiescence bias and capture nuanced perspectives (Lacko et al., 2022), even when some items had lower correlations. This reflects the complex nature of teacher self-efficacy and beliefs (Borg, 2003; Borg, 2015; Bandura, 1997).

Importantly, studies in similar contexts have also reported that well-structured, perception-based pilot questionnaires can yield reliable attitudinal data. For example, Talevski and Shalevska (2023) piloted a questionnaire with EFL teachers in North Macedonia focusing on literature integration, also finding generally positive perceptions and meaningful differentiation across belief domains. Similarly, a pilot of a Kazakhstani EFL teachers' survey on first-language usage and translation beliefs reported Cronbach's α values above .70 for all subscales, reinforcing that pilot-scale validation is feasible and informative (Smagul, 2024).

The relatively high mean scores across most items suggest a generally positive perception of literature's role in language development. Teachers viewed literary texts as beneficial for vocabulary, motivation, and critical thinking—findings that resonate with the theoretical literature (Ghosn, 2013; Duncan & Paran, 2017). This engagement further indicates that the constructs are culturally and professionally relevant.

CONCLUSION

This pilot study provides initial evidence of the reliability and contextual relevance of a 29-item adapted questionnaire designed to examine teachers' beliefs about the role of literary texts in language development. Administered to a small sample of Albanian Language and Literature and EFL teachers in North Macedonia, the instrument demonstrated good internal consistency, particularly following the removal of three underperforming items from the Teacher Practice subscale. Item-level means suggested that teachers broadly value literature for its linguistic, cognitive, and motivational benefits.

Although the limited sample constrains statistical generalizability, the findings offer compelling preliminary support for the instrument's contextual fit and psychometric soundness. Its successful piloting marks a critical step toward rigorous validation in larger and more diverse populations. The instrument is now well-positioned for deployment in broader studies, where advanced procedures—such as Confirmatory Factor Analysis (CFA) and measurement invariance testing—can be employed to substantiate its factorial structure and cross-group stability. Its robust design renders it a valuable tool for both empirical inquiry and targeted professional development in literature-based language instruction.

Implications

The results carry meaningful implications for both research and practice. From a research perspective, the adapted questionnaire shows promise as a diagnostic tool for broader application across varied instructional contexts, including comparative studies between first and foreign language teaching. Its potential to reveal stable belief patterns across diverse educational settings makes it a candidate for future cross-regional and cross-disciplinary validation efforts.

For practitioners and policymakers, the findings affirm the pedagogical value of literature in language education. Participating teachers viewed literary texts as instrumental in promoting student engagement, vocabulary acquisition, and critical thinking. The instrument may thus serve not only as a research tool but also as a practical resource in teacher education, helping to identify professional development needs related to the confident and effective use of literary texts in the classroom.

Limitations

Despite its contributions, this pilot study is subject to several limitations. The sample was small (N = 30) and non-random, limiting the statistical generalizability of the results. The modest sample size also precluded Exploratory Factor Analysis (EFA), which is typically recommended to investigate the dimensional structure of adapted instruments. Furthermore, while Cronbach's alpha values indicated acceptable internal consistency, they offer only preliminary psychometric validation. The exclusive reliance on self-reported data introduces the possibility of response bias and social desirability effects (Lacko et al., 2022).

Future research should address these limitations by involving a larger and more demographically diverse teacher sample. This would enable more robust validation through EFA or CFA and facilitate subgroup analyses examining how teacher perceptions may vary by experience, training, or institutional context.

ACKNOWLEDGEMENTS

REFERENCES

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: W. H. Freeman.

Bland, J. (2013). *Children's literature and learner empowerment*. Bloomsbury.

Bland, J. (Ed.). (2018). *Using literature in English language education: Challenging reading for 8–18 year olds*. Bloomsbury Publishing.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language teaching*, *36*(2), 81-109. https://doi.org/10.1017/S0261444803001903

Borg, S. (2015). *Teacher cognition and language education*. Bloomsbury Publishing. https://www.torrossa.com/it/resources/an/5213476

Brown, J. D. (2011). Quantitative research in second language studies. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning: Volume II* (pp. 190–206). Routledge.

Carabantes, L., & Paran, A. (2022). 'It may also be our own fault to think so, to limit them before even trying': Assuming Learner Limitations during Materials Design in English Language Teacher Education. *Tesol Quarterly*, *56*(4), 1266-1289. https://doi.org/10.1002/tesq.3102

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge. https://doi.org/10.4324/9781315456539

Cunningham, B., & Farmer, A. (2016). An investigation of the relationships between teacher beliefs and student achievement: Teacher beliefs. *The International Journal of Educational Organization and Leadership*, *23*(2), 1. https://doi.org/10.18848/2329-1656/CGP/v23i02/1-14

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage publications.

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). Routledge. https://doi.org/10.4324/9780203864739

Duncan, S., & Paran, A. (2017). The effectiveness of literature on acquisition of language skills and intercultural understanding in the high school context. *London: UCL Institute of Education, University College London*, 3-114. https://www.ibo.org/contentassets/1fcefe0df17448bebe6781ea0396adff/effect-of-literature-on-language-acquisition-final-report.pdf

Ghosn, I. K. (2013). *Storybridge to Second Language Literacy: The theory, research and practice of teaching English with children's literature*. IAP.

Ghosn, I. K. (2018). Materials for early language learning. *The Routledge handbook of teaching English to young learners*, 374-388.

Hall, G. (2015). *Literature in language education* (2nd ed.). Palgrave Macmillan.

Kim, M. (2004). Literature discussions in adult L2 learning. *Language and education*, *18*(2), 145-166. https://doi.org/10.1080/09500780408666872

Lacko, D., Čeněk, J., Točík, J., Avsec, A., Đorđević, V., Genc, A., ... & Subotić, S. (2022). The necessity of testing measurement invariance in cross-cultural research: Potential bias in cross-cultural comparisons with individualism–collectivism self-report scales. *Cross-Cultural Research*, *56*(2-3), 228-267. https://doi.org/10.1177/10693971211068971

Lazar, G. (2020). *A writerly trajectory: reflections on published classroom resources for learners of English and students of academic writing* (Doctoral dissertation, Middlesex University).

Li, C., Yi, T., Zhang, S., Ma, C., & Liu, H. (2022). Developing and validating a scale for measuring pre-service Chinese as an additional language teacher beliefs. *Frontiers in Psychology*, *13*, 989581. https://doi.org/10.3389/fpsyg.2022.989581

Li, L. (2020). *Language teacher cognition*. Palgrave Macmillan UK.

Mardiningrum, A., Larasati, G. F., & Nuraini, E. I. (2024). The Use of Literary Works in EFL Classroom. In *SHS Web of Conferences* (Vol. 202, p. 04010). EDP Sciences. https://doi.org/10.1051/shsconf/202420204010

Mart, C. T. (2018). Literature in language teaching: A recipe to maximize learning. *L1-Educational Studies in Language and Literature*, *18*, 1-25. https://doi.org/10.17239/L1ESLL-2018.18.01.09

McRae, J. (2022). *Literature with a small'l': Developing thinking skills in language teaching and learning*. Wayzgoose Press.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Paran, A. (2008). The role of literature in instructed foreign language learning and teaching: An evidence-based survey. *Language Teaching*, 41(4), 465–496. https://doi.org/10.1017/S026144480800520X

Paran, A., & Robinson, P. (2016). *Literature*. Oxford University Press.

Paran, A., & Tupas, R. (2025). The intersection of literature and English as an international language. In *The Routledge Handbook of Teaching English as an International Language* (pp. 297-311). Routledge.

Paran, A., Spöttl, C., Ratheiser, U., & Eberharter, K. (2020). Measuring literary competences in SLA. In *The Routledge handbook of second language acquisition and language testing* (pp. 326-337). Routledge.

Phipps, S., & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices. *System*, *37*(3), 380-390. https://doi.org/10.1016/j.system.2009.03.002

Savvidou, C. (2020). Second language acquisition through literature: Curriculum, practices and perspectives. In C. Savvidou (Ed.), *Second language acquisition – pedagogies, practices and perspectives* (pp. 1–18). IntechOpen. https://doi.org/10.5772/intechopen.90562

Selfa Sastre, M., & Falguera Garcia, E. (2022). From text on paper to digital poetry: creativity and digital literary Reading practices in initial teacher education. *Frontiers in Psychology*, *13*, 882898. https://doi.org/10.3389/fpsyg.2022.882898

Smagul, A. (2024). Developing, validating, and piloting a MultiTeachViews questionnaire on L1 and translation use: Attitudes of Kazakhstani secondary school EFL teachers. *MethodsX, 13*, 102886. https://doi.org/10.1016/j.mex.2024.102886

Snow, C. E., Perlmann, R., & Nathan, D. (2021). Why routines are different: Toward a multiple-factors model of the relation between input and language acquisition. In *Children's language* (pp. 65-97). Psychology Press.

Sulistiyo, U., Al Arif, T. Z. Z., Handayani, R., Ubaidillah, M. F., & Wiryotinoyo, M. (2022). Determinants of technology acceptance in ICT-based English language learning. *Journal of Language and Education, 8*(2), 17–30. http://dx.doi.org/10.17323/jle.2022.12467

Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education, 48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Talevski, D., & Shalevska, E. (2023). Exploring EFL Teachers' Perceptions: Integrating Children's Literature in the Language Classroom. *International journal of Education Teacher*, *26*, 26-30. https://doi.org/10.20544/teacher.26.04

Tomlinson, B. (Ed.). (2023). *Developing materials for language teaching*. Bloomsbury Publishing.

Van, T. T. M. (2009). The relevance of literary analysis to teaching literature in the EFL Classroom. In *English teaching forum* (Vol. 47, No. 3, p. 2). US Department of State. Bureau of Educational and Cultural Affairs, Office of English Language Programs, SA-5, 2200 C Street NW 4th Floor, Washington, DC 20037.

APPENDIX

Literature Teaching Questionnaire (LTQ)

This instrument was developed to assess teachers' perceptions of the role and impact of literary texts in students' language development. The questionnaire includes 29 items, each rated on a 6-point Likert scale:

1 = Aspak nuk pajtohem (Strongly Disagree)
2 = Nuk pajtohem (Disagree)
3 = Disi nuk pajtohem (Somewhat Disagree)
4 = Disi pajtohem (Somewhat Agree)
5 = Pajtohem (Agree)
6 = Plotësisht pajtohem (Strongly Agree)

1. Teaching literature improves students' vocabulary. [Mësimdhënia e letërsisë përmirëson vokabularin e nxënësve.]
2. I find it easy to integrate literary texts into my syllabus. [E kam të lehtë t'i integroj tekstet letrare në planin tim mësimor.]
3. Students should be able to learn some new grammar from the literary text. [Teksti letrar duhet t'u mundësojë nxënësve të mësojnë diçka të re nga gramatika.]
4. The themes of the text should be age-appropriate. [Temat e tekstit duhet të jenë të përshtatshme për moshën e nxënësve.]
5. The content of the text should support the development of students' critical thinking. [Përmbajtja e tekstit duhet ta zhvillojë të menduarit kritik të nxënësve.]
6. The vocabulary of the literary text should be within the linguistic ability of my students. [Fjalori i tekstit letrar duhet të jetë në korniza të aftësive gjuhësore të nxënësve të mi.]
7. The time available is sufficient to explore literary texts in depth. [Kam kohë të mjaftueshme në dispozicion për t'i trajtuar tekstet letrare në thellësi.]
8. Teaching literature supports students' creative language use. [Letërsia suporton përdorimin kreativ të gjuhës nga ana e nxënësve.]
9. Literary texts help me address multiple language skills (reading, writing, speaking). [Tekstet letrare më ndihmojnë të shqyrtoj aftësi të ndryshme gjuhësore te nxënësit (lexim, shkrim, të folur).]
10. The literary text should offer opportunities for cultural comparison. [Teksti letrar duhet të ofrojë mundësi për krahasime ndërkulturore.]
11. I enjoy teaching literature in my subject area. [E shijoj ligjërimin e letërsisë në lëndën time.]
12. The language of the literary text should reflect current usage. [Gjuha e tekstit letrar duhet të pasqyrojë gjuhën e cila flitet dhe shkruhet sot.]
13. I feel confident in teaching complex literary texts. [Ndihem i/e sigurt kur ligjëroj tekste të ndërlikuara letrare.]
14. Teaching literature helps students develop intercultural awareness. [Letërsia ndihmon nxënësit të zhvillojnë ndërgjegjësimin ndërkulturor.]
15. The literary text should lend itself to teaching a particular language point. [Teksti duhet të mundësojë ligjërimin e një strukture gjuhësore të caktuar.]
16. I often lack sufficient training or guidance on how to teach literary texts. (***Reverse-coded***) [Shpesh e shoh që kam mungesë trajnimi për mënyrën e mësimdhënies së teksteve letrare.]
17. The content of the text should be suitable for classroom discussion. [Përmbajtja e tekstit duhet të jetë e përshtatshme për diskutim në klasë.]
18. Students should be able to learn new vocabulary from the literary text. [Nxënësit duhet të jenë në gjendje të mësojnë fjalor të ri nga teksti letrar.]
19. The grammar of the literary text should be within the linguistic ability of my students. [Gramatika e tekstit letrar duhet të jetë në korniza të aftësive gjuhësore të nxënësve të mi.]
20. The literary text should be relevant to my students' lives. [Teksti letrar duhet të ketë relevancë për përvojat dhe kontekstin jetësor të nxënësve të mi.]
21. Overall, the literary text should be well within the linguistic ability of my students. [Në përgjithësi, teksti letrar duhet të jetë lehtësisht i kuptueshëm për nxënësit.]

22. The literary text should explore global or social issues. [Teksti letrar duhet të trajtojë çështje globale apo sociale.]
23. Literary texts increase student motivation to learn the language. [Tekstet letrare rrisin motivimin e nxënësve për ta mësuar gjuhën.]
24. For each genre, shorter texts work better than longer texts. [Për çdo zhanër (gjini), tekstet më të shkurtra funksionojnë më mirë se ato më të gjata.]
25. The literary text should challenge my students from a language point of view. [Teksti letrar duhet t'i sfidojë nxënësit e mi nga aspekti gjuhësor.]
26. The literary text should be of high literary quality. [Teksti letrar duhet të jetë i një cilësie të lartë letrare.]
27. The literary text should offer aesthetic or emotional value. [Teksti letrar duhet të ofrojë vlera estetike ose emocionale.]
28. The literary text should represent the diversity of national and world literature. [Teksti letrar duhet të përfaqësojë larminë e letërsisë kombëtare dhe asaj botërore.]
29. The literary text should support the development of specific language skills (e.g., reading, writing, grammar, speaking). [Teksti letrar duhet ta ndihmojë zhvillimin e shkathtësive specifike gjuhësore (p.sh. lexim, shkrim, gramatikë, të folur).]