rudn.tlcjournal.org
doi: 10.22363/2521-442X-2026-10-3-109-124

Original Research

Service-learning in language education: A scoping review (2018–2026)

Vladimir S. Nikolskiy[1*]

[1]Moscow Polytechnic University, Moscow, Russia

No transparent, reproducible synthesis has examined service-learning research in language education published after 2017 while jointly covering foreign-language, ESL/EFL and teacher-education settings. This scoping review mapped recent contexts, participants, research designs, documented outcomes and supporting evidence. Scopus and ERIC searches identified 47 eligible sources: 44 primary empirical reports and 3 reviews. The primary reports represented an estimated 42 distinct studies. Review conduct followed scoping-review guidance, reporting followed PRISMA-ScR, and all final eligibility decisions were made by the author after AI-assisted full-text screening. The primary corpus spanned 14 countries but remained concentrated in United States institutions, English and Spanish target languages, and higher education or pre-service teacher education. Qualitative designs predominated, and nearly all reports included self-reported or perceived evidence. Linguistic, professional, intercultural and civic outcomes commonly co-occurred. Only five reports used assessed measures, two included comparison groups, four followed participants for more than one course or project cycle, and eight collected direct community evidence. Fully online service appeared in three reports, and generative AI was directly investigated in one. The synthesis separates publication counts from underlying studies and outcome prevalence from evidential strength. It establishes that the literature documents several categories of participant learning and experience but provides insufficient evidence of consistent effectiveness, durable change or reciprocal community benefit. Reciprocity and socially mediated learning define service-learning, yet recipient perspectives remain largely absent from its empirical record.

Keywords: *service-learning, foreign language education, TESOL, community-based language learning, language teacher education, intercultural competence, scoping review*

Article history:
Submitted March 22, 2026
Revised August 5, 2026
Accepted September 16, 2026

CRediT Author Statement:
Vladimir S. Nikolskiy: Conceptualisation, Methodology, Validation, Formal Analysis, Investigation, Data Curation, Writing – Original Draft, Writing – Review & Editing, Visualisation.

Conflict of interest:
The author declared no conflict of interest.

Data availability statement:
The supplementary research materials for this review are available in Figshare at https://doi.org/10.6084/m9.figshare.33869971.

Funding:
No funding was reported for this research.

Acknowledgements:
AI tools supported language editing and AI-assisted screening under the author's sole adjudication. The author made all final eligibility decisions and assumes full responsibility for the work.

Doi:
10.22363/2521-442X-2026-10-3-109-124

For citation:
Nikolskiy, V. S. (2026). Service-learning in language education: A scoping review (2018–2026). *Training, Language and Culture, 10*(3), 109–124.

This is an open access article distributed under a Creative Commons Attribution-NonCommercial 4.0 International License (CC BY-NC 4.0), which allows its unrestricted use for non-commercial purposes, subject to attribution. The material can be shared/adapted for non-commercial purposes if you give appropriate credit, provide a link to the license, and indicate if changes were made.

1. INTRODUCTION

Service-learning (SL) has evolved over the past four decades from isolated volunteering initiatives into a recognised educational approach incorporated into the institutional policies of universities worldwide (Bringle & Clayton, 2022). SL is defined as a form of experiential education in which students engage in organised activity that addresses community needs while receiving structured opportunities for academic and personal growth through critical reflection (Jacoby, 1996). What distinguishes SL from volunteering, internships or practicums is the inseparable link between service and academic objectives, the requirement for structured reflection, and the presence of mutual benefit for the parties involved (Bringle & Clayton, 2022;

* Corresponding author
© Vladimir S. Nikolskiy 2026
Licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

Palpacuer-Lee et al., 2018). In language education, scholarly interest in SL grew from the late 1990s onwards. Early publications were largely descriptive accounts of individual programmes. In 2013, two special issues published in *Hispania* (Long, 2013) and *TESOL Journal* (Wurr, 2013) documented growing institutional recognition of the field (Long, 2013; Wurr, 2013).

During its first two decades, the area developed predominantly in the US and centred on Spanish: in the systematic review by Baker (2019), which covered publications through 2017, undergraduate Spanish programmes in the United States constituted the large majority of the corpus.

Two developments affected language education after 2017: the COVID-19 pandemic (2020—2022), which disrupted community-based projects requiring physical presence while accelerating virtual alternatives, and the rapid diffusion of generative artificial intelligence (GenAI) from 2022 onwards, which raised new questions about authenticity and the use of technology in language interaction. No transparent, reproducible map of the 2018—2026 literature covering both the foreign-language and the TESOL strands was available at the time of writing; among the reviews published during this period, Wurr (2018) addresses the TESOL strand and Anderson (2023) international service-learning, while Baker (2019) covers 1997 to 2017.

This study addresses that gap. Following the PRISMA extension for scoping reviews (Tricco et al., 2018) and current scoping-review conduct guidance (Levac et al., 2010; Peters et al., 2020), it maps research published between 2018 and 2026 to identify dominant contexts, methodological approaches, documented outcomes and gaps. Data extraction addressed the following research questions:

1. In what geographical, linguistic and educational contexts is SL in language education implemented, as documented in research published in 2018—2026?

2. Which participant categories and educational levels predominate?

3. Which methodological approaches and research designs are characteristic of the period?

4. What linguistic, educational and social-civic outcomes are documented, and with what kind of evidence?

5. What gaps and directions for future research are identified?

2. METHODOLOGY

2.1. Design

The study used a scoping review design to classify and tabulate publications on service-learning in language education according to publication year, country/context, target language, educational setting, participant group, research design, data sources, and outcome category (Arksey & O'Malley, 2005; Levac et al., 2010). Categories represented by few included publications were recorded as gaps in the evidence. This design accommodates heterogeneous evidence from multiple contexts. Methodological quality appraisal is not routinely required in scoping reviews and was not conducted in this study. The procedures were based on the methodological guidance of Levac et al. (2010) and Peters et al. (2020). Reporting followed the PRISMA-ScR checklist (Tricco et al., 2018).

2.2. Units of inclusion and analysis

Evidence-synthesis methodology distinguishes a publication ('report') from the study or project it documents (Higgins et al., 2024). The JBI Manual for Evidence Synthesis uses 'source of evidence' as the general designation for any item included in a scoping review (Peters et al., 2020). Accordingly, the report was the unit of inclusion in this review, while 'study' referred to a project, course, intervention, or participant sample documented in one or more reports. The corpus comprised 47 reports, each counted as one source of evidence: 44 primary empirical reports and 3 reviews. Counts and analyses of primary empirical evidence were restricted to the 44 primary reports. The three reviews were analysed separately. All reports were compared to establish whether two or more publications documented the same project, course, intervention, or participant sample. The assessment results are summarised in Section 4.8.

2.3. Information sources and search strategy

Two bibliographic databases were searched: Scopus and ERIC through the EBSCOhost interface. Both searches were completed on 4 August 2026. All retrieval counts in this section refer to searches completed on that date. Scopus was selected for multidisciplinary indexing in applied linguistics, education and area studies. ERIC was selected for education-specific indexing of teacher education, TESOL and language-programme research. In Scopus, the query was run in the title, abstract and keyword fields:

```
TITLE-ABS-KEY(
"service-learning" OR "service learning"
)
AND
TITLE-ABS-KEY(
"language education" OR "language learning"
OR
"language teaching" OR "language course" OR
"language programme" OR "second language" OR
"foreign language"
)
```

No date, document-type or language filters were applied through the Scopus interface. Date and document-type eligibility criteria were applied during screening. The search returned 41 records, all of which were exported.

In ERIC, the query was run using the EBSCOhost default keyword search:

```
(
"service learning" OR "service-learning"
)
AND
(
"language education" OR "language programme"
OR
"language program" OR "language course" OR
"language training" OR "language learning"
)
```

The query returned 381 records before interface filters. The EBSCOhost filters 'Peer Reviewed' and 'Publication Type: Journal Articles' reduced the result to 290 records. The rolling ten-year date filter reduced it to 154 records, all of which were exported. This filter included ten records published in 2017. These records were removed before screening under the 2018—2026 eligibility criterion and recorded individually in the screening and eligibility log.

Two checks tested whether known eligible reports were present in the corpus, and a third quantified database overlap. The reference lists of the three included reviews and the author's pre-search list of known reports were compared with the corpus. Neither check found an eligible report absent from the corpus. Deduplication removed 20 of the 195 retrieved records (10.3%). This percentage quantifies record overlap between the databases and does not measure search completeness.

Web of Science, Linguistics and Language Behavior Abstracts, the MLA International Bibliography and PsycINFO were not searched because institutional access was unavailable. Eligible reports indexed exclusively in these databases may therefore be absent from the corpus.

2.4. Eligibility criteria

Reports were eligible only when they met all five criteria: (a) they examined service-learning in language education or a community-based learning intervention that combined organised community activity with explicit language-learning or language-teacher-education objectives; (b) they were published from 1 January 2018 through 4 August 2026; (c) they were peer-reviewed journal articles reporting primary empirical research or a review synthesis; (d) publications in any language were eligible; and (e) the full text could be retrieved. For eligibility, a review synthesis had to identify a body of prior publications and synthesise findings from that body. Reviews eligible under criterion (c) were analysed separately, excluded from all denominators for primary empirical evidence, and reported in Section 4.9.

Reports were excluded when they had no language-education content, including service-learning projects in medicine or social work with no language-learning objective. Records in which 'service learning' denoted an unrelated application or platform, such as *LearningApps*, were also excluded. Reports that mentioned service-learning only in the discussion and presented no investigation of its implementation or outcomes were ineligible. Theoretical and position papers were excluded. Course, syllabus, practice and programme descriptions were excluded unless they reported an empirical objective, a participant or document sample, data-collection procedures and an analytical procedure. Books, book chapters, conference proceedings and dissertations were ineligible. Reports outside the specified publication period, reports not published as peer-reviewed journal articles, and reports for which full text could not be retrieved were also excluded.

2.5. Study selection: AI-assisted screening

The author completed study selection with AI assistance at the full-text stage. Eligibility criteria were fixed before screening. The author screened titles and abstracts without AI assistance. On 4 August 2026, Claude Opus 4.8 applied the predefined eligibility protocol to the full texts of all 82 reports and returned a recommended decision with a brief justification for each report. The author received these outputs before recording the human decisions and subsequently read every full text. The AI recommendations therefore formed part of the screening procedure and did not represent an independent second assessment. The author made all final eligibility decisions and retained responsibility for the final corpus, consistent with the RAISE guidance (Thomas et al., 2025) and the joint position statement on AI use in evidence synthesis (Flemyng et al., 2025).

The AI system recommended inclusion for 52 reports and exclusion for 30. The author included 47 reports and excluded 35. The decisions agreed for 77 reports: both assessments included 47 and excluded 30. Five reports received an AI recommendation for inclusion and a final human decision of exclusion; no report received an AI recommendation for exclusion and a final human decision of inclusion. Raw agreement was 93.9%, Cohen's κ was .873, and PABAK was .878. Because the author viewed the AI recommendations before recording the final decisions, these statistics describe concordance between the recorded outputs and cannot serve as estimates of independent reviewer agreement.

The five discordant reports were Sánchez-López (2019), Macknish (2019), Biagi and Bracci (2022), Ciriza (2023), and Rauschert and Byram (2018). All five were excluded under criterion (c). They reported no empirical research question or objective, participant or document sample, data-collection procedure, or analytical procedure, and none presented a review synthesis.

Eight reports were classified as borderline in the screening log. Feuerherm (2022), Lassila and Ahn (2022), Sharif et al. (2024), Omidire (2020), and Porto (2019) were excluded after a second full-text reading. Hernández

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

Varona and Gutiérrez Álvarez (2020) and Matheus and Gaugler (2020) remained in the corpus. Rauschert and Byram (2018) was excluded under criterion (c). The screening and eligibility log records the decision, reason and full bibliographic details for every excluded report. Excluded reports discussed in this section also appear in the reference list.

Two sensitivity analyses tested the borderline decisions. The first added the five discordant reports excluded by the author and regenerated Tables 2—7 using an alternative corpus of 49 reports. The second removed the two borderline reports retained in the corpus and regenerated the same tables with 42 primary reports. The complete recalculated results are recorded in the screening and eligibility log.

The modal categories were unchanged in both sensitivity analyses: North America for region, English for target language, higher education for educational setting, qualitative for research design, language and communication for outcome category, and self-report for evidence type. The United States accounted for 21/44 reports (47.7%) in the main analysis, 24/49 (49.0%) in the first sensitivity analysis, and 20/42 (47.6%) in the second. Fully online delivery accounted for 3/44 reports (6.8%), 3/49 (6.1%), and 3/42 (7.1%), respectively. Community or partner data were present in 8/44 reports (18.2%), 8/49 (16.3%), and 8/42 (19.0%).

The rank orders of target languages, educational settings, research designs, and evidence types remained unchanged. Changes were confined to categories separated by one to three reports. In the first sensitivity analysis, Europe rose from fifth to fourth place, ahead of the Middle East and North Africa. Professional and teacher development fell from second to fourth place among outcome categories, below civic and social justice and intercultural competence. Affective and motivational outcomes and digital and multimodal outcomes exchanged fifth and sixth places. The first sensitivity analysis also added school pupils as learners as a fifth educational setting. In the second sensitivity analysis, Latin America decreased from five to four reports and tied with the Middle East and North Africa. No modal category changed in either analysis.

2.6. Selection process and results (PRISMA-ScR)

Figure 1 presents the record-selection process. The database searches produced 195 records: 41 from Scopus and 154 from ERIC. Deduplication removed 20 records, and 10 records published in 2017 were removed under the 2018—2026 publication-date criterion. Apart from deduplication, publication year was the only basis for removal before title and abstract screening. No topical eligibility decision was made before that stage. The titles and abstracts of 165 records were screened. Eighty records were excluded because their bibliographic metadata indicated that they did not concern service-learning in a language-education setting. Full texts were sought for the remaining 85 reports. Three reports could not be retrieved by 4 August 2026 and are identified in the screening and eligibility log. One of these was the only record published in 2026 identified by either database.

The full texts of 82 reports were assessed, and 35 were excluded with one primary reason assigned to each report: 18 did not examine an intervention meeting the eligibility definition of service-learning, including telecollaboration, study-abroad programmes, teaching practica and ordinary classroom research; 3 contained no language-education content; 12 presented theoretical, position, practice or programme descriptions without an empirical investigation or review synthesis; and 2 described a service-learning activity unrelated to the language-education course or programme examined in the report.

The final corpus comprised 47 sources of evidence: 44 primary empirical reports and 3 reviews. The primary empirical analysis used the 44 reports, representing an estimated 42 distinct studies or projects.

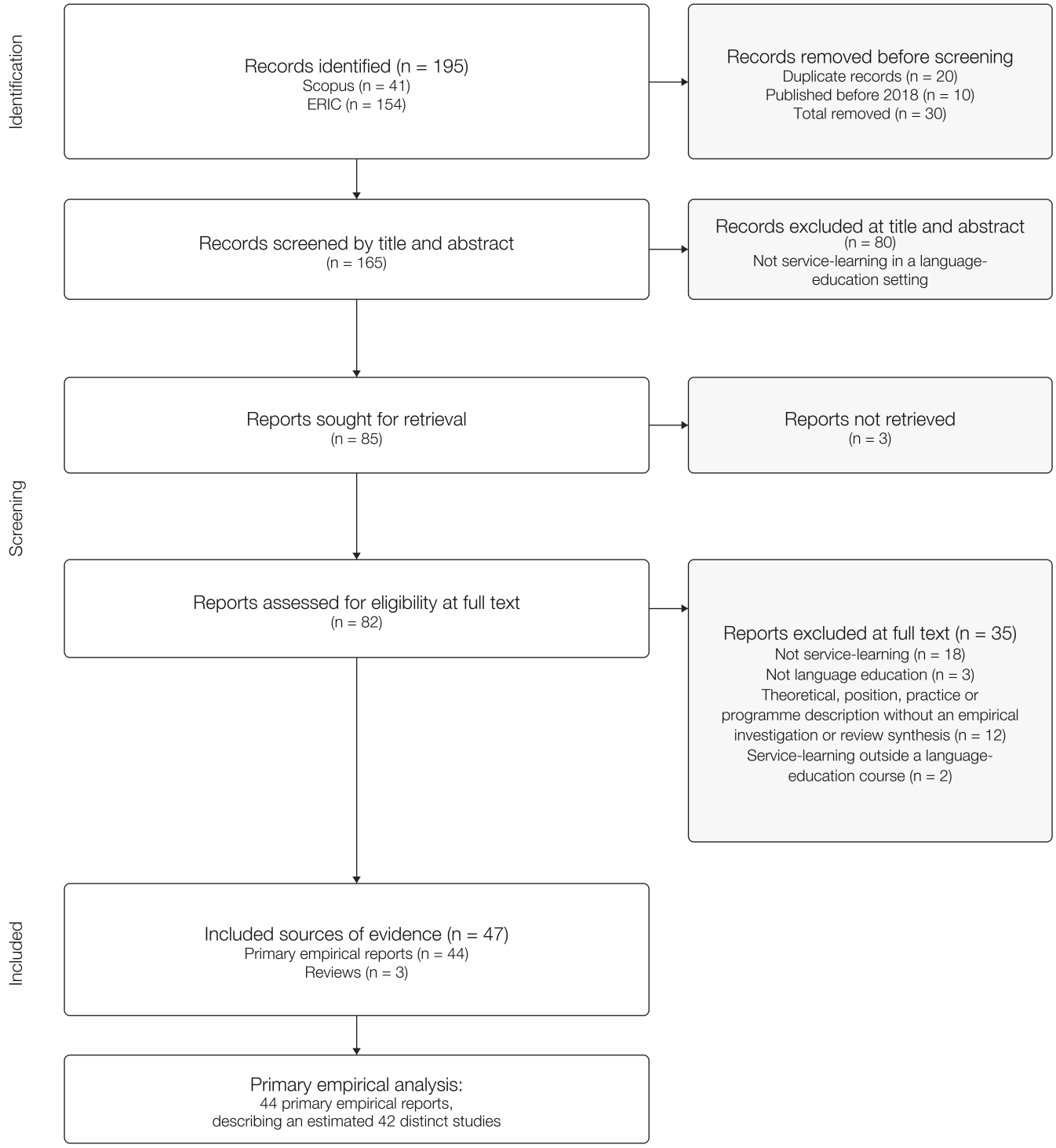
2.7. Data extraction and coding

A piloted form was used to extract data from the full text of each included report. Extracted variables comprised publication year; country of the participating institution; country of the service activity, if different; target language; educational setting; participant, recipient and data-provider groups; research design; presence of a comparison group; reported participant count and unit; principal data sources; documented outcomes; evidence type; service-delivery mode; collection of community or educator data; and links to publications reporting the same project or participant sample.

The matrix contained six binary fields for outcome categories and four for evidence types. Each report could receive multiple outcome and evidence-type codes. Evidence was classified as: (i) self-reported or perceived; (ii) assessed through a test, rubric or scored artefact; (iii) observed behaviour or analysis of an unscored product; and (iv) supported by a comparator. The report was the unit of coding. Since these categories were non-exclusive, their totals could exceed the number of reports. The matrix identifies every report contributing to each total.

Participant counts and units were transcribed as reported in the source. ‘NR’ was entered when no usable participant count was provided. Values referring to programmes, classes, publications or subgroups retained their original unit labels. All tabulated totals were calculated from the data extraction and coding matrix.

Service-delivery mode was determined from the reported location and method of service provision. ‘Online only’ was assigned when the service activity occurred online. Digital production of a podcast, digital story, bilingual



Note. One primary exclusion reason was assigned to each report excluded after full-text assessment.

Figure 1. PRISMA-ScR flow of records through the review

glossary or another deliverable did not qualify as online service delivery. Variables corresponding to RQ1—RQ5 were specified before coding. Online delivery and generative AI were added during analysis as topical codes. Result statements were worded according to evidence type: self-reported findings are described as reported or perceived, and non-causal relations as associations. Stronger terms are used only when supported by the research design.

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

2.8. Theoretical frameworks and coding criteria

The coding criteria were derived from three theoretical frameworks. Kolb's (2015) experiential-learning cycle provided criteria for coding reflection as a connection between service activity and learning outcomes. Reflective diaries and journals were examined for evidence of this connection. Byram's (Barili & Byram, 2021; Rauschert & Byram, 2018) account of intercultural citizenship provided criteria for the intercultural and civic/social-justice outcome categories. Civic/social-justice coding required an explicit objective or activity concerning civic action; a general reference to culture did not satisfy this criterion. Vygotsky's (1978) socio-constructivist theory of learning provided criteria for coding community interaction as a setting for learning and development. Evidence sources were classified separately as self-report, observed interaction, assessed performance or analysed artefacts.

Application of a theoretical category required evidence meeting its coding definition. Reports without such evidence received codes based on their reported characteristics, data sources and findings. The matrix records the resulting outcome and evidence codes; the three theories are not study-level variables. Section 3 presents their theoretical basis.

2.9. Supplementary research materials

The supplementary files comprise the database search strategy and search record; the original Scopus and ERIC search exports; the screening and eligibility log for all 195 records; the data extraction and coding matrix with its codebook; and the included-sources table. The included-sources table lists the 47 included reports and their principal characteristics. The screening and eligibility log records title-and-abstract decisions, full-text eligibility decisions and exclusion reasons, AI recommendations for the 82 full texts, the author's final decisions, and sensitivity-analysis results. The data extraction and coding matrix records report-level variables, outcome and evidence codes, and publications concerning the same project or participant sample. Report IDs are consistent between the data extraction and coding matrix and the included-sources table. Repository access is specified in the Data Availability Statement.

3. THEORETICAL BACKGROUND

3.1. Definition of relevant terms

SL is a course-based pedagogical approach that combines organised service, academic learning, structured reflection and reciprocal engagement with community partners (Bringle & Clayton, 2022; Palpacuer-Lee et al., 2018; Sreelakshmi & Vijayakumar, 2026). Reflection connects the service experience with course objectives and the analysis of civic issues. Reciprocity requires recognition of the knowledge, priorities and intended benefits of students and community partners. These criteria informed the operational definition used in this review. Volunteering, internships and practicums may also contain service or reflection. Classification depended on the stated educational objectives, service relationship and reflective procedure. Reports were eligible when they documented SL or an equivalent community-based learning approach meeting the criteria specified in Section 2.4.

Several related terms identify particular forms of SL in language education. Community-based service-learning (CBSL) refers to course-related learning conducted through direct engagement with a local community. Community-based language learning (CBLL) uses interaction with community participants in the target language as a source of language learning (Reisinger & Clifford, 2022). Intercultural service-learning (ISL) connects service activity with intercultural citizenship and civic participation in multilingual and multicultural settings (Barili & Byram, 2021; Rauschert & Byram, 2018). E-service-learning denotes SL in which the service activity is conducted through digital technologies (Fuchs & Ferguson, 2025; Park & Yi, 2023; Song, 2022). Digital production of materials without online service activity did not qualify as e-service-learning in this review.

Some accounts distinguish conventional SL, centred on academic objectives and competency development, from critical SL. Critical SL examines social justice, unequal power relations and community-defined priorities in service partnerships (Porto, 2023a, 2023b; Schneider, 2019). These questions are especially relevant to projects involving immigrants, refugees and minoritised communities, where unequal institutional authority may produce paternalistic assumptions about community participants (Biçer & Alan, 2022; Porto et al., 2022).

3.2. Theoretical foundations

Three theories provided the basis for the coding criteria. Kolb's (2015) experiential-learning cycle presents learning as a sequence of experience, reflection, abstraction and subsequent action. In SL, the service activity supplies the experience, and structured reflection connects that experience with academic knowledge and future action. This theory supported the coding of reflection and reflective data sources.

Byram's (Rauschert & Byram, 2018; Barili & Byram, 2021) account of intercultural citizenship connects intercultural learning with participation in civic action. It supported the intercultural and civic/social-justice outcome categories. Intercultural coding required evidence of engagement with linguistic or cultural difference. Civic/social-justice coding required an explicit objective, activity or finding concerning civic participation, inequality or social action. A general reference to culture did not satisfy the civic/social-justice criterion.

Vygotsky's (1978) sociocultural theory locates learning in socially mediated activity. Language use with students, educators and community participants can provide opportunities for learning through interaction. This theory supported the coding of community interaction and evidence derived from observed interaction or participant-produced artefacts.

3.3. Development of service-learning scholarship in language education

The pedagogical rationale for contemporary SL is associated with experiential education and the institutional development of university community engagement (Jacoby, 1996; Kolb, 2015). Baker (2019) identified 69 journal articles on community-based service-learning in second-language higher education published from 1997 to 2017. Publication frequency increased after 2008, and ten articles appeared in *Hispania*'s 2013 special-focus issue on community engagement. *TESOL Journal* published the special issue *Engaged Teaching and Learning: Service-Learning, Civic Literacy, and TESOL* in the same year (Hellebrandt & Jorge, 2013; Perren, 2013).

Spanish-language higher education in the United States and Canada accounted for 54 of the 69 publications reviewed by Baker (2019). Baker (2019) associated this concentration with established university partnerships involving Spanish-speaking communities. The literature also included heritage-language education, languages for specific purposes, English-language education and digitally mediated service activity.

3.4. Prior reviews and scope of comparison

Baker (2019) reviewed 69 publications from 1997 to 2017. Twenty-nine reported an empirical investigation or another systematic line of inquiry; the remaining 40 comprised programme descriptions, theoretical discussions and professional recommendations. Wurr (2018) examined service-learning research involving English language learners. Anderson (2023) examined linguistic difference in international service-learning. Baker (2019) reported a concentration on Spanish-language higher education in the United States and Canada, extensive use of qualitative methods and student self-reports, limited longitudinal evidence, few community perspectives and few comparison-group designs. The review recommended research on additional languages and locations, direct collection of community data, longitudinal observation and quasi-experimental designs.

The eligibility criteria used by Baker (2019) differ from those used here. Baker (2019) restricted inclusion to postsecondary second-language settings, excluded most TESOL and teacher-education publications, included only English-language articles and admitted many publications without an empirical investigation. The present review includes ESL/EFL, teacher education, school-level settings and educator-focused reports. Its primary analyses include only empirical reports. Baker's (2019) Spanish-language proportion of 78.3% represents 54 of 69 publications. The 25.0% reported in the present review uses the 44 primary empirical reports as its denominator. These percentages refer to differently defined corpora and are not interpreted as evidence of change over time. Baker (2019) described the predominance of student perspectives without reporting a percentage for exclusively student-focused publications. The present review therefore reports its community- and educator-data counts without a numerical comparison with Baker's (2019) corpus.

4. RESULTS

Unless stated otherwise, all counts and percentages in Sections 4.1—4.8 refer to the 44 primary empirical reports. Section 4.9 reports the three reviews separately. A summary of the 47 included reports and their principal characteristics is provided in the supplementary included-sources table.

4.1. Chronological distribution

From 2018 to 2025, the annual number of primary empirical reports ranged from three to eight and showed no monotonic increase (Table 1). No 2026 report was included. The searches returned one 2026 record, but its full text could not be retrieved for eligibility assessment.

Table 1
Included sources of evidence by year of publication (N = 47)

SOURCE	2018	2019	2020	2021	2022	2023	2024	2025	TOTAL
Primary empirical reports	7	5	8	4	7	4	3	6	44
Reviews	1	1	0	0	0	1	0	0	3
TOTAL	8	6	8	4	7	5	3	6	47

Note. One 2026 record could not be retrieved for full-text assessment.

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

4.2. Geographical contexts (RQ1)

The country of the participating institution was used as the geographical unit. North America accounted for 21 of the 44 reports (47.7%); all 21 concerned institutions in the United States. The remaining reports concerned institutions in Asia (10, 22.7%), Latin America (5, 11.4%), the Middle East and North Africa (4, 9.1%), Europe (3, 6.8%) and Oceania (1, 2.3%). Hong Kong accounted for four reports and Argentina for three. Thirteen countries other than the United States were represented, including six represented by one report each: Australia, Chile, Colombia, Jordan, Libya and Spain (Table 2).

Table 2
Geographical distribution of primary empirical reports (N = 44)

REGION	COUNTRIES OF THE PARTICIPATING INSTITUTIONS (N)	REGION TOTAL (%)
North America	United States (21)	21 (47.7%)
Asia	Hong Kong (4); Malaysia (2); South Korea (2); Vietnam (2)	10 (22.7%)
Latin America	Argentina (3); Chile (1); Colombia (1)	5 (11.4%)
Middle East / North Africa	Turkey (2); Jordan (1); Libya (1)	4 (9.1%)
Europe	Poland (2); Spain (1)	3 (6.8%)
Oceania	Australia (1)	1 (2.3%)
Total		44 (100%)

Note. Each report was assigned to the country of the institution whose learners participated. Six reports concerned service delivered in another country: Czop Assaf et al. (2019), United States–South Africa; Matheus and Gaugler (2020), United States–Dominican Republic; Shin and Jeon (2018), South Korea–Canada; Marklin-Gruber (2019), Hong Kong–Mainland China; Fuchs and Ferguson (2025), United States–Germany (virtual); and Baecher and Chung (2020), United States–Costa Rica.

4.3. Linguistic contexts (RQ1)

English was the target language in 25 of the 44 reports (56.8%), and Spanish in 11 (25.0%). These two languages accounted for 36 reports (81.8%). Four reports concerned multiple languages or multilingual projects. The remaining four concerned Russian, Korean, Turkish and German, with one report for each language (Table 3). Nguyen and Truong (2025) reported the Russian-language case.

Table 3
Target language of the service-learning course or project (N = 44)

TARGET LANGUAGE	N	PERCENTAGE
English (ESL/EFL and English as an Additional Language)	25	56.8%
Spanish	11	25.0%
Multiple or multilingual projects	4	9.1%
Other single languages	4	9.1%
Total	44	100%

Note. Other single languages: Nguyen & Truong (2025) Russian for specific purposes; Shin & Jeon (2018) Korean service language; Biçer & Alan (2022) Turkish as a host language; Fuchs & Ferguson (2025) German foreign language. Multiple or multilingual projects: Smith-Warshaw et al. (2020); Mocarz-Kleindienst (2025); Shi & Cheung (2025); Reisinger & Clifford (2022).

4.4. Settings and participant groups (RQ2)

Higher education was the principal educational setting in 24 reports (54.5%), and pre-service teacher education — in 17 (38.6%). Two reports examined practising teachers as participants, one was concerned with university educators. No report involved school pupils as learners enrolled in an SL course. Children or school pupils were recipients of university-student service in 31 reports (70.5%). Graybeal et al. (2025) measure language development among preschoolers receiving service from undergraduates. Adult immigrants or refugees were recipients in 14 reports (31.8%). Community members, partner staff or families supplied data in 8 reports (18.2%). Educator or faculty accounts were analysed in 5 reports (11.4%). Table 4 reports the educational settings and the overlapping participant, recipient and data-provider groups.

Table 4
Educational setting and participant, recipient and data-provider groups (N = 44)

CATEGORY	N	PERCENTAGE
EDUCATIONAL SETTING (mutually exclusive; sums to N)		
Higher education (students)	24	54.5%
Pre-service teacher education	17	38.6%
In-service teachers as participants	2	4.5%
University educators as participants	1	2.3%
Subtotal	44	100%
RECIPIENT, PARTNER AND DATA-PROVIDER GROUPS (overlapping; not additive with the block above)		
School pupils or children among the recipients	31	70.5%
Adult immigrants or refugees among the recipients	14	31.8%
Community members, partner staff, families or recipients as a data source	8	18.2%
Educators' or faculty members' own accounts analysed	5	11.4%

4.5. Research designs (RQ3)

Thirty-two of the 44 reports (72.7%) used qualitative designs, eight (18.2%) used mixed methods and four (9.1%) used quantitative designs (Table 5). The qualitative group included case studies, phenomenological studies, narrative inquiry, action research and thematic analyses of reflective diaries, interviews and artefacts. Two reports used a non-SL comparison group. Ayala Pazmiño (2019) compared randomly allocated SL and control sections. Bataineh and Obeiah (2022) conducted an EFL writing quasi-experiment with a control group receiving conventional training. Ene and Orlando (2022) compared self-directed and other-directed SL placements without a non-SL comparator. Olszak (2023) used a single-group pre/post design.

Table 5
Research designs of the primary empirical reports (N = 44)

DESIGN	N	PERCENTAGE
Qualitative	32	72.7%
Mixed methods	8	18.2%
Quantitative	4	9.1%

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

4.6. Outcomes and type of evidence (RQ4)

Six overlapping outcome categories were coded. Thirty-seven of the 44 reports (84.1%) documented outcomes in more than one category. The six category counts totalled 99 (Table 6). Evidence type was recorded through four non-exclusive binary codes, with the report as the unit. Reports could receive multiple evidence codes; the four counts therefore totalled 75 (Table 7). Forty-three reports (97.7%) included self-reported or perceived evidence. Lorenzen (2018) was the sole report without self-reported evidence. Twenty-five reports (56.8%) analysed observed behaviour, recordings or unscored artefacts. Five reports (11.4%) included a test, rubric or scored artefact. Two reports (4.5%) included a non-SL comparison group.

Table 6
Outcome domains by frequency and type of evidence (N = 44; domains overlap)

OUTCOME DOMAIN	REPORTS (N)	PERCENTAGE	EVIDENCE TYPES AND EXAMPLES IN THE CORPUS
Language and communication	24	54.5%	Predominantly self-reported or perceived outcomes; assessed evidence in five reports (vocabulary: Tocaimaza-Hatch, 2018; Walls & Tocaimaza-Hatch, 2018; writing: Bataineh & Obeiah, 2022; oral proficiency: Ayala Pazmiño, 2019; children's language: Graybeal et al., 2025).
Professional and teacher development	23	52.3%	Self-reported changes in beliefs and professional identity; analyses of lesson plans, recorded teaching and reflective artefacts.
Civic and social justice	21	47.7%	Self-reported outcomes and findings from reflective and narrative data (Porto, 2023a, 2023b; Porto et al., 2022; Ramos Leiva et al., 2021).
Intercultural competence	20	45.5%	Predominantly self-reported or perceived outcomes; pre/post survey evidence in Collopy et al. (2020) and Matheus and Gaugler (2020).
Affective and motivational	6	13.6%	Self-reported confidence and motivation.
Digital and multimodal	5	11.4%	Analysed digital products and self-reports; one report examined generative AI (Cheung & Shi, 2024).

Note. Outcome categories are not mutually exclusive; their counts sum to more than N.

Table 7
Type of evidence supporting documented outcomes (N = 44 reports; categories overlap)

TYPE OF EVIDENCE	REPORTS (N)	PERCENTAGE
Self-reported or perceived	43	97.7%
Observed behaviour or an analysed product, without scoring	25	56.8%
Assessed with a test, rubric or scored artefact	5	11.4%
Supported by a comparison group	2	4.5%

Note. The unit of analysis is the report. Evidence categories are not mutually exclusive. The two comparator-supported reports also received assessed and self-report codes.

4.7. Service delivery, generative AI (RQ4—RQ5)

Three reports concerned fully online service delivery: Song (2022), Park and Yi (2023) and Fuchs and Ferguson (2025). Two reports concerned partly online service: Barili and Byram (2021) and Pazilah et al. (2021). The remaining 39 reports concerned face-to-face service.

Five reports received the digital and multimodal outcome code. Song (2022) and Fuchs and Ferguson (2025) concerned online service delivery. Cheung and Shi (2024) and Shi and Cheung (2024, 2025) concerned face-to-face service with digital deliverables. Shi and Cheung (2025) report that educators in their division declined to transfer the service activity online because they considered face-to-face interaction necessary for the service relationship.

One report examined generative AI. Cheung and Shi (2024) examined undergraduates' use of generative-AI tools to co-create stories in a Hong Kong SL subject. Shi and Cheung (2025) mention generative AI among the tools used to prepare service deliverables in the same division but do not examine its use.

4.8. Reports, studies and overlapping samples

The 44 primary empirical reports represented an estimated 42 distinct studies. Three clusters of related publications were identified during full-text assessment.

Shi and Cheung (2024), Cheung and Shi (2024) and Shi and Cheung (2025) concern SL subjects in one division of a Hong Kong institution. Shi and Cheung (2024) report data from 12 students in the first cohort of a digital-storytelling subject. Cheung and Shi (2024) report data from nine students in a later cohort of the same subject. Shi and Cheung (2025) report data from seven educators teaching four subjects, including the digital-storytelling subject. The three reports concern distinct samples from the same institutional setting. Porto (2023a) and Porto (2023b) concern the same project, service site and cohort at a community centre in La Plata. The three pre-service teachers analysed in Porto (2023b) are a subset of the five student teachers reported in Porto (2023a). Porto et al. (2022) concerns a different institution, service site and sample in Patagonia. All three publications arose from the same funded initiative conducted in 2017—2018. Walls and Tocaimaza-Hatch (2018) state that another part of the same SL experience is reported in Tocaimaza-Hatch (2018). The two publications concern the same course and community partnership and probably report the same cohort.

The Porto pair and the Walls—Tocaimaza-Hatch pair account for the reduction from 44 reports to an estimated 42 distinct studies. All 44 reports were retained in counts of publication output. Four included reports also cite companion publications that fall outside the review corpus: Czop Assaf et al. (2019), Ene and Orlando (2022), Fuchs and Ferguson (2025) and Palpacuer-Lee et al. (2018).

4.9. Review-level findings

The three reviews were excluded from the denominators reported in Sections 4.1—4.8.

Baker (2019) synthesises 69 journal articles published from 1997 to 2017, including 54 concerning Spanish-language education. The review reports extensive use of qualitative methods and student self-reports, limited longitudinal data, few community perspectives and few comparison-group designs. Its abstract classifies 39 articles as empirical studies; its Results section classifies 29 as empirical or as containing another systematic line of inquiry.

Wurr (2018) surveys SL research involving English-language learners in language institutes, two- and four-year colleges and graduate teacher education. The author labels the article a meta-analysis. The article reports no effect sizes, eligibility criteria or screening counts.

Anderson (2023) analyses 53 articles concerning 49 international SL programmes. Twenty-five of the programmes addressed linguistic difference. The review identifies seven reported methods for addressing linguistic difference.

4.10. Areas with limited or absent coverage (RQ5)

Five areas had limited or absent coverage in the corpus. First, community members, partner staff, families or recipients supplied data in 8 of the 44 reports. Educator or faculty accounts were analysed in 5 reports. Second, four reports followed participants for more than one academic year or one project cycle: Czop Assaf et al. (2019), Pak (2020), Baecher and Chung (2020) and Fuchs and Ferguson (2025). Two reports used a non-SL comparison group. Third, English and Spanish accounted for 36 reports (81.8%). No participating institution was located in sub-Saharan Africa, South Asia, Central Asia or a state of the former Soviet Union. Fourth, no report examined school pupils as learners enrolled in an SL course. Children or school pupils appeared as service recipients in 31 reports. Fifth, generative AI was examined in one report.

5. DISCUSSION

5.1. Composition of the 2018—2026 corpus

The 44 primary empirical reports published from 2018 to 2025 concern institutions in 14 countries. The United States accounts for 21 reports (47.7%), while 23 concern institutions elsewhere. English is the target language in 25 reports and Spanish in 11; together they account for 81.8% of the corpus. The educational settings are concentrated in higher education and pre-service teacher education, which account for 41 of the 44 reports. The retrieved literature therefore combines international geographical coverage with pronounced linguistic and institutional concentration. Baker's (2019) corpus was concentrated in Spanish-language higher education in the United States and Canada, with Spanish represented in 54 of 69

Volume 10 Issue 3, 2026, pp. 109-124 doi: 10.22363/2521-442X-2026-10-3-109-124

publications. The present eligibility criteria also admit ESL/EFL research and teacher education, categories that Baker (2019) largely excluded. Differences in eligibility criteria, database coverage and units of analysis prevent measurement of temporal change between the two corpora. Both reviews nevertheless identify concentration in a small number of linguistic and institutional contexts. The current corpus establishes that this concentration is also present among empirical reports retrieved for 2018—2025, despite representation from thirteen countries outside the United States. The target-language categories identify the language associated with each report. They do not record the internal composition of multilingual or contact-language varieties. Research on mixed languages has examined lexical composition in relation to patterns of language emergence (Gridneva, 2026). Future syntheses with greater representation of contact-language settings may require more detailed language coding.

5.2. Outcomes and strength of evidence

Multiple outcome categories were common, while assessed and comparative evidence was scarce. Thirty-seven of the 44 reports documented outcomes in more than one category. Language and communication appeared in 24 reports, professional and teacher development in 23, civic and social justice in 21, and intercultural competence in 20. These frequencies show that service-learning publications commonly associate a single project with linguistic, professional, intercultural and civic outcomes.

The evidence supporting these outcomes was predominantly self-reported or perceived. Forty-three reports included self-reported evidence, 25 analysed observed behaviour or unscored products, 5 used a test, rubric or scored artefact, and 2 included a non-SL comparison group. The corpus supports conclusions about the outcomes documented in the literature and the evidence used to substantiate them. It supplies little basis for estimates of effectiveness, causal effects or comparative advantage over other forms of instruction.

Reflection journals and interviews document the participant meaning-making described in Kolb's (2015) experiential-learning framework. When used to substantiate learning outcomes, these sources remain participant reports and cannot independently establish linguistic attainment or durable behavioural change. The intercultural and civic/social-justice categories correspond to Byram's (Barili & Byram, 2021; Rauschert & Byram, 2018) account of intercultural citizenship. Most supporting evidence in these categories consists of reflections, interviews and narrative data. It establishes reported orientations, experiences and intentions; sustained intercultural performance and civic action remain weakly documented. The 25 reports analysing observed interaction or participant-produced artefacts supply evidence of situated activity relevant to Vygotsky's (1978) theory. The small number of assessed and comparator-supported reports limits conclusions about measurable change attributable to service-learning.

Baker (2019) also reported extensive reliance on qualitative methods and student self-reports, with limited longitudinal evidence, community perspectives and comparison-group designs. These methodological characteristics remain present under the empirical eligibility criteria applied in the current review.

5.3. Reciprocity, community evidence and longitudinal research

Reciprocity requires evidence concerning the experiences and intended benefits of students and community partners (Bringle & Clayton, 2022; Jacoby, 1996). Children or school pupils were service recipients in 31 reports, and adult immigrants or refugees were recipients in 14. Community members, partner staff, families or recipients supplied data in 8 reports. In the remaining 36 reports, community, partner or recipient data were not collected. Educator or faculty accounts were analysed in 5 reports.

The evidence base is therefore substantially stronger for student experiences than for community benefit. Student reflections can document learning, professional development and civic orientation. They cannot establish whether recipients regarded the service as useful, reciprocal or responsive to community priorities. Civic and social-justice outcomes appeared in 21 reports, while direct community evidence appeared in 8. The distinction separates evidence of student civic learning from evidence that a project benefited its community partner.

Several reports document ethical and practical reasons for omitting recipient data. Cheung and Shi (2024) could not obtain ethical clearance to interview refugee children, and Shi and Cheung (2024) excluded recipient interviews on privacy grounds. Fuchs and Ferguson (2025) and Ene and Orlando (2022) restricted data collection to students in their research designs. These cases show that the absence of recipient data partly reflects protections for vulnerable participants. The resulting limitation remains substantive: reciprocity cannot be evaluated from the community perspective in most reports.

Only four reports followed participants for more than one academic year or project cycle (Assaf et al., 2019; Baecher & Chung, 2020; Fuchs & Ferguson, 2025; Pak, 2020). The durability of reported linguistic, intercultural, professional and civic outcomes is consequently unknown. Claims about sustained development require follow-up beyond a single course or service project.

5.4. Online service delivery and generative AI

Fully online service delivery appeared in 3 reports, partly online delivery in 2, and face-to-face service in 39. Digital production and online service delivery constitute

distinct categories in this corpus. Several face-to-face projects required students to produce digital stories, podcasts or other digital materials, although the service relationship itself remained face to face.

A digital story, podcast or glossary documents a student-produced artefact. Its production alone does not establish whether recipients accessed or used it, or whether they found it useful. Evaluations of digital service projects could examine these questions directly through delivery records and recipient accounts. This would provide evidence about the community use of digital materials (Grishechko, 2026) alongside evidence about students' learning.

No fully online report was published in 2020 or 2021. The three fully online reports were published in 2022, 2023 and 2025. Publication dates and the small number of reports provide no basis for attributing online service-learning activity to the COVID-19 pandemic. The online reports document outcomes under their local conditions, but the corpus contains no comparative analysis of online and face-to-face delivery.

Generative AI was the subject of one report. Cheung and Shi (2024) examined undergraduates' use of generative-AI tools to co-create stories in one Hong Kong service-learning subject. Shi and Cheung (2025) mention generative AI as one tool used to prepare service materials but do not examine its use. Evidence concerning generative AI is therefore limited to one focal report from a single institution. No conclusion about its prevalence, effectiveness or influence on language service-learning can be made from this corpus.

5.5. Limitations

The search covered Scopus and ERIC and included reference-list checking of the three eligible reviews. Web of Science, Linguistics and Language Behavior Abstracts, the MLA International Bibliography and PsycINFO were not searched because institutional access was unavailable. Eligible reports indexed exclusively in those databases may be absent. The search terms were in English, which may have reduced retrieval of publications issued without English-language titles, abstracts or keywords. Grey literature was excluded by the eligibility criteria, so the findings apply to peer-reviewed journal articles.

Three reports that passed title-and-abstract screening could not be retrieved in full text. One was the only 2026 record identified by either database. The corpus consequently provides incomplete coverage of 2026 and may under-represent the most recent research.

No methodological quality or risk-of-bias appraisal was conducted. The evidence-type coding distinguishes self-reported, observed, assessed and comparator-supported findings. It does not establish study quality or risk of bias.

The synthesis reports the prevalence and characteristics of documented evidence and does not produce pooled estimates of intervention effects. Study selection used AI assistance with one human adjudicator. The author viewed the AI recommendations before recording the final decisions. The agreement statistics reported in Section 2.5 therefore describe concordance between the recorded AI recommendations and the author's decisions; they do not estimate agreement between independent assessors. Final eligibility decisions were made by one person, leaving the procedure susceptible to confirmation bias.

The unit of inclusion was the report. Two pairs of reports concern the same or overlapping participant samples, and the 44 primary reports represent an estimated 42 distinct studies. Publication frequencies may therefore include more than one report from the same project. Section 4.8 identifies these cases and separates publication counts from estimated study counts.

6. CONCLUSION

The empirical literature mapped in this review presents service-learning in language education as an internationally distributed area of research concentrated in tertiary English- and Spanish-language settings. It links service activity with linguistic, professional, intercultural and civic learning and commonly records several outcome categories from the same project. The research record remains predominantly qualitative and student-centred, including projects in which children, migrants and community organisations are the intended recipients.

The separation of outcome categories from evidence type, and of publications from underlying studies, establishes the evidential boundaries of this literature. Positive outcome frequencies do not constitute estimates of effectiveness, and publication frequencies do not represent independent studies. Reciprocity and socially mediated learning are defining principles of service-learning, yet recipient perspectives are rarely collected.

The available evidence therefore supports conclusions about participants' reported experiences and perceived learning. Evidence of consistent effectiveness, durable change and reciprocal community benefit remains insufficient.

Digital production is more established in the corpus than fully online service delivery, and evidence concerning generative AI is limited to one institutional case. Future research should collect ethically appropriate community and recipient evidence, use assessed and longitudinal outcomes, employ comparison designs when suited to the research question, and include school learners, underrepresented languages and locations, online delivery and generative-AI-supported service. Such evidence would support firmer conclusions about effectiveness, persistence and reciprocity.

References

Anderson, R. W. (2023). How do international service-learning programs attend to linguistic difference? A review of literature. *Journal of Service-Learning in Higher Education, 16*, 44—79.

Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*(1), 19—32. https://doi.org/10.1080/1364557032000119616

Ayala Pazmino, M. F. (2019). Service-learning education improvement for foreign language learners. *Journal of Language Teaching and Research, 10*(6), 1173—1180. https://doi.org/10.17507/jltr.1006.05

Baecher, L., & Chung, S. (2020). Transformative professional development for in-service teachers through international service learning. *Teacher Development, 24*(1), 33—51. https://doi.org/10.1080/13664530.2019.1682033

Baker, L. (2019). Community-based service-learning in language education: A review of the literature. *International Journal of Research on Service-Learning and Community Engagement, 7*(1), Article 2. https://dx.doi.org/10.37333/001c.11480

Barili, A., & Byram, M. (2021). Teaching intercultural citizenship through intercultural service learning in world language education. *Foreign Language Annals, 54*(3), 776—799. https://doi.org/10.1111/flan.12526

Bataineh, R. F., & Obeiah, S. F. (2022). Does service-learning training make a difference? An exploratory study of Jordanian EFL teachers' performance in and attitudes towards writing. *International TESOL & Technology Journal, 17*(2), 53—75.

Biagi, F., & Bracci, L. (2022). A democratic reflective syllabus in a second language acquisition context: In-person, digital and hybrid experiences. *Iranian Journal of Language Teaching Research, 10*(3), 43—60. https://dx.doi.org/10.30466/ijltr.2022.121225

Biçer, N., & Alan, Y. (2022). Service-based learning experiences of pre-service Turkish teachers: Language teaching to Syrian refugees. *South African Journal of Education, 42*(3), Article 2111. https://dx.doi.org/10.15700/saje.v42n3a2111

Bringle, R., & Clayton, P. (2022). Enriching global civic education in foreign language education through service learning. In C. Lütge, T. Merse, & P. Rauschert (Eds.), *Global citizenship in foreign language education: Concepts, practices, connections* (pp. 79—106). Routledge.

Cheung, L. M. E., & Shi, H. (2024). Co-creating stories with generative AI: Reflections from undergraduate students of a storytelling service-learning subject in Hong Kong. *Aust. Rev. Appl. Linguist, 47*(3), 259—283.

Ciriza, M. D. P. (2023). 'Teaching English as service' in Spanish language programs: A translanguaging approach. *L2 Journal, 15*(1), 1—14.

Collopy, R. M. B., Tjaden-Glass, S., & McIntosh, N. A. (2020). Attending to conditions that facilitate intercultural competence: A reciprocal service-learning approach. *Michigan Journal of Community Service Learning, 26*(1), 19—37. https://dx.doi.org/10.3998/mjcsloa.3239521.0026.102

Czop Assaf, L., O'Donnell Lussier, K., Furness, S., & Hoff, M. (2019). Exploring how a study abroad and international service-learning project shaped preservice teachers' understanding of humanizing pedagogy. *The Teacher Educator, 54*(2), 105—124. https://doi.org/10.1080/08878730.2018.1520376

Ene, E., & Orlando, H. (2022). International students and autonomy in service-learning English for academic purposes courses. *Michigan Journal of Community Service Learning, 28*(2), 1—21. https://dx.doi.org/10.3998/mjcsl.1162

Feuerherm, E., Hiramatsu, K., Miller, N. S., & Williams, K. (2022). Assessing outcomes of service learning: Student, instructor, and community reflections. *Journal of Community Engagement and Higher Education, 14*(2), 4—18.

Flemyng, E., Noel-Storr, A., Macura, B., Gartlehner, G., Thomas, J., Meerpohl, J. J., Jordan, Z., Minx, J., Eisele-Metzger, A., Hamel, C., Jemioło, P., Porritt, K., & Grainger, M. (2025). Position statement on artificial intelligence (AI) use in evidence synthesis across Cochrane, the Campbell Collaboration, JBI, and the Collaboration for Environmental Evidence 2025. *JBI Evidence Synthesis, 23*(11), 2162—2166. https://doi.org/10.11124/JBIES-25-00480

Fuchs, C., & Ferguson, H. (2025). Longitudinal civic engagement: Undergraduate students' reflections on an intergenerational virtual exchange. *AILA Review, 38*(2), 294—323. https://dx.doi.org/10.1075/aila.24045.fuc

Graybeal, L., Kim Thao, S., Porter, M., & Sims, M. (2025). Language learning and career readiness outcomes of an undergraduate speech pathology service-learning project. *Journal of Service-Learning in Higher Education, 20*, 73—91.

Gridneva, E. M. (2026). The hypothesis of dependence of the lexical nature of mixed languages on the patterns of their emergence. *Vestnik Tomskogo Gosudarstvennogo Universiteta, Filologiya, 100*, 38—52. https://dx.doi.org/10.17223/19986645/100/4

Grishechko, E. G. (2026). How to do things with code: An analytical framework for linguistic inquiry into programming environments. *Training, Language and Culture, 10*(1), 54—71. https://doi.org/10.22363/2521-442x-2026-10-1-54-71

Hellebrandt, J., & Jorge, E. (2013). The scholarship of community engagement: Advancing partnerships in Spanish and Portuguese. *Hispania, 96*(2), 203—214. https://doi.org/10.1353/hpn.2013.0058

Hernández Varona, W., & Gutiérrez Álvarez, D. F. (2020). English language student-teachers developing agency through community-based pedagogy projects. *PROFILE: Issues in Teachers' Professional Development, 22*(1), 109—122. https://dx.doi.org/10.15446/profile.v22n1.76925

Higgins, J., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M. J., & Welch, V. A. (Eds.). (2024). *Cochrane handbook for systematic reviews of interventions*. Cochrane.

Jacoby, B. (1996). *Service-learning in higher education: Concepts and practices*. Jossey-Bass.

Kolb, D. A. (2015). *Experiential learning: Experience as the source of learning and development* (2nd ed.). Pearson.

Lassila, E. T., & Ahn, R. (2022). Japanese pre-service teachers' learning experience using Moses' student-centered pedagogical framework. *International Journal of Teacher Leadership, 11*(2), 11—27.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: Advancing the methodology. *Implementation Science, 5*, Article 69. https://dx.doi.org/10.1186/1748-5908-5-69

Long, S. S. (2013). Editor's message: Focusing on the scholarship of community engagement. *Hispania, 96*(2), 201—202. https://doi.org/10.1353/hpn.2013.0052

Lorenzen, C. (2018). The potential of high-leverage teaching practices for pre-service language teacher development within community-based children's language programs. *Learning Languages, 23*(1), 15—20. https://eric.ed.gov/?id=EJ1187746

Macknish, C. J. (2019). Creating multimodal reflections in an IEP speaking-listening course. *TESL Canada Journal, 36*(3), 173—185. https://doi.org/10.18806/tesl.v36i3.1326

Marklin-Gruber, L. (2019). Service-learning benefits for English language learners: A case of China—Hong Kong cross-border English teaching. *Journal of Higher Education Outreach and Engagement, 23*(3), 21—36.

Matheus, C. C., & Gaugler, K. M. (2020). Fostering global mindedness in short-term community-based global learning programs: The importance of strategic design, collaboration, and reflection. *Frontiers: The Interdisciplinary Journal of Study Abroad, 32*(3), 156—181. https://doi.org/10.36366/frontiers.v32i3.583

Mocarz-Kleindienst, M. (2025). Implementing service-learning in student multilingual projects. *XLinguae, 18*(4), 162—170. https://doi.org/10.18355/XL.2025.18.04.11

Nguyen, N. T., & Truong, T. P. T. (2025). Developing professional Russian communication skills for hospitality internships through service-learning: A case study from Vietnam. *Russian Language Studies, 23*(4), 642—655.

Olszak, I. (2023). The impact of service-learning on English education. A socio-cultural project as an attempt to increase cultural awareness. *New Horizons in English Studies, 8*, 91—104. https://dx.doi.org/10.17951/nh.2023.8.91-104

Omidire, M. F. (2020). Experiencing language challenges in a rural school: Implications for learners' life aspirations. *Early Child Development and Care, 190*(10), 1619—1637. https://doi.org/10.1080/03004430.2019.1660963

Pak, C.-S. (2020). Exploring the long-term impact of service-learning: Former students of Spanish revisit their community engagement experiences. *Hispania, 103*(1), 67—85. https://doi.org/10.1353/hpn.2020.0004

Palpacuer-Lee, C., Hutchison Curtis, J., & Curran, M. E. (2018). Stories of engagement: Pre-service language teachers negotiate intercultural citizenship in a community-based English language program. *Language Teaching Research, 22*(5), 590—607. https://dx.doi.org/10.1177/1362168817718578

Park, E., & Yi, Y. (2023). Preservice teachers' experiences with e-service learning for middle school students in South Korea. *Language Teaching Research Quarterly, 35*, 21—36. https://doi.org/10.32038/ltrq.2023.35.02

Pazilah, F. N., Hashim, H., & Md Yunus, M. (2021). Service-learning in an English as a second language teacher training program: Exploring pre-service teachers' authentic learning experiences. *Arab World English Journal, 12*(2), 377—398. https://dx.doi.org/10.24093/awej/vol12no2.26

Perren, J. (2013). Strategic steps to successful service-learning in TESOL: From critical to practical. *TESOL Journal, 4*(3), 487—513. https://doi.org/10.1002/tesj.96

Peters, M. D. J., Marnie, C., Tricco, A. C., Pollock, D., Munn, Z., Alexander, L., McInerney, P., Godfrey, C. M., & Khalil, H. (2020). Updated methodological guidance for the conduct of scoping reviews. *JBI Evidence Synthesis, 18*(10), 2119—2126. https://dx.doi.org/10.11124/JBIES-20-00167

Porto, M. (2019). Affordances, complexities, and challenges of intercultural citizenship for foreign language teachers. *Foreign Language Annals, 52*(1), 141—164. https://doi.org/10.1111/flan.12375

Porto, M. (2023a). A community service learning experience with student teachers of English: Enacting social justice in language education in a community centre. *The Language Learning Journal, 51*(2), 178—194. https://doi.org/10.1080/09571736.2021.1971741

Porto, M. (2023b). Experientially grounded praxis of social justice language education: Pre-service teachers of English engage in field work in an Argentine NGO. *Teaching Education, 34*(2), 225—245. https://doi.org/10.1080/10476210.2022.2078299

Porto, M., Tavella, G., & Fernández, C. (2022). Undoing inequalities: Inclusive and transformative language practices in rural Argentinian Patagonia. *International Journal of Bilingual Education and Bilingualism, 25*(10), 3597—3613.

Ramos Leiva, L., Pérez Miranda, L., & Riquelme-Sanderson, M. (2021). Social justice in the preparation of English language teachers. *MEXTESOL Journal, 45*(2).

Rauschert, P., & Byram, M. (2018). Service learning and intercultural citizenship in foreign-language education. *Cambridge Journal of Education, 48*(3), 353—369. https://doi.org/10.1080/0305764X.2017.1337722

Reisinger, D. S., & Clifford, J. (2022). Community-based language learning as a transformative practice: An exploration of the modes of communication. *Foreign Language Annals, 55*(3), 668—683. https://dx.doi.org/10.1111/flan.12629

Sánchez-López, L. (2019). Transforming the community's linguistic landscape in a mixed-learner service learning capstone course in a Spanish for specific purposes program. *Revista Signos, 52*(101), 996—1012. https://doi.org/10.4067/S0718-09342019000300996

Schneider, J. (2019). Teaching in context: Integrating community-based service learning into TESOL education. *TESOL Journal, 10*(1), Article e00380. https://doi.org/10.1002/tesj.380

Sharif, M. N., Shaharuddin, A., Sirri, A., & Ramli, N. S. (2024). Toys for Friends (TFF): A community service learning (CSL) to cultivate social responsibility among primary school students. *Malaysian Journal of Qualitative Research, 10*(1), 53—63. https://dx.doi.org/10.61211/mjqr100105

Shi, H., & Cheung, L. M. E. (2024). Storytelling for understanding: A case study of an English-language digital storytelling service-learning subject for refugee children in Hong Kong. *Journal for Multicultural Education, 18*(1—2), 81—97. https://dx.doi.org/10.1108/JME-10-2023-0116

Shi, H., & Cheung, L. M. E. (2025). Service-learning as inclusive CALL: A Hong Kong case study. *Computer Assisted Language Learning Electronic Journal, 26*(2), 54—82. https://doi.org/10.54855/callej.252623

Shin, H., & Jeon, M. (2018). Intercultural competence and critical English language teacher education. *English Teaching, 73*(4), 125—147. https://dx.doi.org/10.15858/engtea.73.4.201812.125

Smith-Warshaw, J., Crume, P., & Pinzón-Pérez, H. (2020). Impact of service-learning on Latinx college students engaged in intervention services for the deaf: Building multicultural competence. *International Journal of Multicultural Education, 22*(3), 50—75. https://dx.doi.org/10.18251/ijme.v22i3.2413

Song, S. (2022). Digital service-learning: Creating translanguaging spaces for emergent bilinguals' literacy learning and culturally responsive family engagement in mainstream preservice teacher education. *TESL-EJ, 26*(3). https://doi.org/10.55593/ej.26103a5

Sreelakshmi, V. S. K., & Vijayakumar, S. (2026). Intercultural communicative competence in ESP research: A Scientometric Mapping of global trends, theoretical clusters, and pedagogical gaps (2003—2023). *Training, Language and Culture, 10*(2), 81—100. https://doi.org/10.22363/2521-442X-2026-10-2-81-100

Thomas, J., Flemyng, E., & Noel-Storr, A. (2025). *Responsible use of AI in evidence SynthEsis (RAISE): Recommendations and guidance* (Version 2, updated June 3, 2025). OSF. https://doi.org/10.17605/OSF.IO/FWAUD

Tocaimaza-Hatch, C. C. (2018). Speaking is doing and doing is learning: Vocabulary learning in service-learning. *Journal of Spanish Language Teaching, 5*(1), 66—80. https://doi.org/10.1080/23247797.2018.1458552

Tricco, A. C., Lillie, E., Zarin, W., O'Brien, K. K., Colquhoun, H., Levac, D., ... Straus, S. E. (2018). PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Annals of Internal Medicine, 169*(7), 467—473. https://doi.org/10.7326/M18-0850

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Walls, L. C., & Tocaimaza-Hatch, C. C. (2018). Exposing vocabulary learning in service-learning. *International Journal of Research on Service-Learning and Community Engagement, 6*(1), Article 17. https://dx.doi.org/10.37333/001c.6965

Wurr, A. J. (2013). Editorial. *TESOL Journal, 4*(3), 397—401. https://doi.org/10.1002/tesj.99

Wurr, A. J. (2018). Advances in service-learning research with English language learners. *Journal of Service-Learning in Higher Education, 8*, 84—99.

ABOUT THE AUTHOR

Vladimir S. Nikolskiy
DSc in Education, Professor
Center for Project Activity
Moscow Polytechnic University, Moscow, Russia
38 Bolshaya Semyonovskaya Str., 107023 Moscow, Russia
Email: v.s.nikolskij@mospolytech.ru
ORCID: https://orcid.org/0000-0002-4290-1443