Review

A Scientometric Review of Research Trends in Language Assessment Literacy

Xuan Wang *[ID], **Jingying Zuo** [ID], **Feiyu Liu and Ze Sun**

School of Foreign Studies, Northwestern Polytechnical University, Xi'an 710072, China
* Correspondence: xuan.emily.wang@nwpu.edu.cn

Abstract: This study provides a scientometric review of language assessment literacy (LAL) research published in the Web of Science (WoS) core collection between 2008, when LAL was first indexed as a combined phrase, and 2021. Six comprehensive analyses were conducted utilizing CiteSpace: publication and cited trends, dual-map overlay, most cited papers, co-authorship, co-occurrence, and co-citation. Results showed that the annual publications have shown a fluctuating upward trend that has undergone three stages: the initial (2008–2012), fluctuated (2013–2017), and steadily increasing stages (2018–2021). Professional conferences in language assessment and testing have significantly contributed to these development stages and are closely related to stage-critical milestones. "Teacher" appears as one of the most frequently researched stakeholder groups, and research on other stakeholders, especially students, has increased. "Psychology" and "Education" are major areas for both citing and cited journals and provide a solid foundation for LAL research. However, the interdisciplinary nature of LAL research still leaves room for exploration. This study provides a reference point for future research by identifying trends in LAL research and comparing existing studies. Furthermore, when authors and author linkage networks are used to build a research network, multi-dimensional indicators should be considered to guarantee the completeness of the network.

Keywords: language assessment literacy; scientometric review; CiteSpace

1. Introduction

Language assessment has been widely researched across varying educational and professional contexts [1–3] and is considered vitally important to students' academic growth [4–6]. Effective assessment enables language teachers to achieve and maintain high teaching quality [7,8], while inadequate assessment weakens language learning quality, resulting in a loss of student motivation and confidence [9,10]. Assessing language skills is challenging because many factors can influence the reliability and validity of designing, administering, interpreting, utilizing, and reporting procedures. Therefore, language assessment literacy (LAL) is vital to solving these problems and benefiting students [11–13].

Academic focus on "assessment literacy" stems from the test-based accountability system that emerged in many educational settings in the 1980s [14,15]. In these educational settings, teachers had to monitor and report all student learning outcomes according to established standards [9,16]. After comprehensively considering informal language assessments, such as classroom observation, random testing, and oral feedback, Brindley [17] distinguished LAL to reflect language disciplines' unique characteristics and requirements.

The LAL research that has emerged since has explored conceptual frameworks and clarifying relationships among the elements examined. Davies [18] and Inbar-Lourie [19] developed and explained the earliest conceptual framework, consisting of principles (why to evaluate), knowledge (what to evaluate), and skills (how to evaluate). Taylor [20] constructed a LAL appraisal framework for different stakeholder groups that included eight elements: knowledge of theory, technical skills, principles and concepts, language pedagogy, sociocultural values, local practices, personal beliefs/attitudes, and scores and decision

making. Baker and Riches [21] further explored this LAL framework with Haitian teachers and narrowed its elements down to seven: theoretical and conceptual knowledge, task performance, language pedagogy, collaboration, awareness of local practices, awareness of personal beliefs/attitudes, and decision making. With empirical data's continuous enrichment and validation in different assessment contexts, new conceptual frameworks are constantly being constructed [21–23].

Different stakeholder groups have varying interests, needs, and expectations about language testing and assessment; therefore, their understanding of LAL frameworks and mastery levels of components are different [20,22,24]. Taylor [20] categorized LAL stakeholders into four groups: test writers, classroom teachers, university administrators, and professional language testers. Stakeholder groups' LAL profiles have different components and values because they do not participate in assessment activities identically. Brindley [17] proposed that, compared to other stakeholder groups, language teachers must master three extra components: the ability to evaluate language tests, develop language curriculum assessments, and put assessment into practice. Taylor [20] stated that although the four stakeholder groups need to master the same components of LAL, the requirements in each component are different because of the characteristics of the stakeholder group. For example, professional language testers are expected to have a thorough knowledge of every aspect of the assessment process and to achieve value 4 (the highest level) in components of knowledge of theory and principles and concepts. At the same time, classroom teachers focus more on designing practical assessment methods and are expected to only achieve level 2 in the same components. Because of these differences, LAL is not usually researched as a holistic concept and has been researched separately within different stakeholder groups. Classroom teachers are the most investigated group in the literature.

Stakeholders' LAL is the product of social influence and numerous factors affect it. Previous studies have found that internal factors closely related to stakeholders and external factors dominated by the environment jointly affect the development of LAL [9,25]. Internal factors, such as teachers' self-confidence [26], willingness to participate in assessment training [27,28], and teaching experience, [9] are those that directly affect their LAL. As for external factors, LAL is easily influenced by educational environments, teaching and assessment administrative orders, educational policies, and socio-cultural values of language teaching [8,23,29].

The existing literature significantly improves the understanding of LAL's conceptual nature, analytical framework, and social contexts; however, mainstream research papers focus on empirical investigation and theoretical discussion. New systematic review articles on LAL are relatively limited but necessary. First, emerging trends and thematic patterns should be scientifically explored because LAL is a complex research topic that keeps changing [30]. Second, more reviews have been conducted on assessment literacy (AL) in general education [31–34] than LAL in language assessment and testing, but LAL should be researched separately due to its complexities regarding assessment knowledge and communicative competence [4,35,36]. Third, existing reviews on LAL employ a conventional scoping approach that assists with discovering foci, themes, and contributions of LAL research. Still, few studies adopt the scientometric method, which has excellent benefits when reviewing a field of research [37]. Compared to the conventional approach, scientometrics can logically investigate emerging trends and thematic patterns and dynamically explore the intellectual landscape and citation linkage by analyzing metrics and indicators in bibliographic records. Additionally, broader literature can be reviewed in a visual format, enabling researchers to find influential patterns in complex networks more precisely and efficiently than a conventional review, which tends to include fewer than 100 articles [35,36].

While existing literature reflects that there is increased attention on LAL, there is no consensus on an optimal LAL theoretical framework. LAL is highly contextual, and its stakeholders are significantly involved, which means that LAL is a complex research topic

with complex networks. In this way, scientometric reviews should be used to explore its emerging trends, intellectual landscape, critical changes, and thematic patterns.

2. Methods

This study employed a scientometric approach to review emerging trends, intellectual landscape, critical changes, and thematic patterns in LAL research. Within the scope of this study, data was retrieved from one of the most widely used library resources, the Web of Science (WoS) [38]. To ensure the visibility of state-of-the-art research, the databases used for scientometric analyses in this article include the Social Science Citation Index (SSCI), the Art & Humanities Citation Index (A&HCI), and the Emerging Sources Citation Index (ESCI). SSCI is a multidisciplinary index including over 3400 journals across 58 social science disciplines from 1985 to the present. A&HCI features abstracting and indexing for more than 1700 arts and humanities journals and covers fields including social and natural science journals. Finally, ESCI extends the scope of publications in WoS to include high-quality and peer-reviewed publications in all disciplines. Therefore, this study included these three databases comprised of almost all authoritative academic journals in the social sciences and arts to retrieve comprehensive literature on LAL.

Researchers queried the 'topic' section on WoS, which searches titles, abstracts, keywords, and Keywords Plus, to search for data for the scientometric review. Publications on LAL from the WoS core collection on the SSCI, A&HCI, and ESCI databases were searched using both the combined phrase "assessment literacy" (AL) and the word "language" as search items. This is because "LAL" is written as either "language assessment literacy," "assessment literacy of teachers of languages," or "assessment literacy of EFL teachers" in the literature. Initially, the period was set to 2021, 20 years from when Brindley [17] first differentiated LAL from AL in 2001. However, LAL was first indexed into the WoS core collection via Inbar-Lourie's [19] journal article in 2008; therefore, 14 years from 2008 to 2021 were selected on WoS via "articles" and "reviews" as document types. Next, the language type was set to English. Concerning literature categories, multidisciplinary studies were included due to the interdisciplinary nature of LAL and the increased willingness of journals to accept articles from various fields. For instance, Sustainability has been categorized as a journal of environmental sciences and green technology; however, many of its social science articles have been cited in applied linguistics. Thus, to comprehensively understand LAL's overall dynamics and disciplinary structure, no categories or citation topics were excluded while searching and collecting data sources.

163 articles and review articles were collected from 77 journals across 10 WoS categories. These included education and educational research; language and linguistics; educational psychology; interdisciplinary applications of computer science; cultural studies; multidisciplinary psychology; environmental sciences; environmental studies; and green and sustainable science and technology. The final three categories were sourced from Sustainability. None of the 163 studies were duplicated or removed.

CiteSpace, a Java application created by Chaomei Chen, was employed to analyze and visualize the data. CiteSpace is widely applied in analyzing and visualizing co-citation networks, enabling researchers to identify co-authorship, co-occurrence, and co-citation network algorithms [39]. To thoroughly review LAL's emerging trends, intellectual landscape, critical changes, and thematic patterns, the following CiteSpace analyses were conducted: (1) dual-map overlay; (2) co-authorship analysis among researchers, organizations, and countries; (3) co-occurrence analysis among terms, keywords, and categories; and (4) co-citation analysis among references, authors, and journals. Publication and citation trends and highly cited papers and authors based on analyses from WoS and CiteSpace were also adopted to understand emerging trends in LAL and its intellectual landscape profoundly and comprehensively.

To comprehensively explore LAL's emerging trends, intellectual landscape, critical changes, and thematic patterns, this study examined publication and citation trends, dual-map overlay, highly cited papers, co-authorship, co-occurrence analysis, and co-citation

analyses. Specifically, publication and citation trends and dual-map overlay primarily help clarify emerging trends and critical changes; highly cited papers and co-authorship help understand the intellectual landscape; and co-occurrence analysis and co-citation analyses help understand thematic patterns. However, these analyses are not independent of each other nor only belong to one function; the analyses can support each other. For example, author co-citations in co-citation analyses can also support an understanding of the intellectual landscape.

3. Results

3.1. Publication and Citation Trends

A document's annual distribution and the times it has been cited are valuable indicators for analyzing, evaluating, and predicting a research field's development [40,41]. 163 LAL publications (articles and reviews) published between 2008 and 2021 were retrieved on WoS using the searching approach described in the research methods. The yearly quantitative distributions of these publications' literature and citations used by separate and cumulative techniques were investigated.

Figure 1 displays the number of LAL publications annually published and cited. Generally, LAL publications underwent three stages: initial (2008–2012), fluctuated (2013–2017), and steadily increasing (2018–2021). In the first stage, with the introduction of LAL, scholars explored concept definitions and theoretical framework construction [18,19,42,43]. In 2011, the 33rd Language Testing Research Colloquium (LTRC) held its first workshop on LAL. Conference reports were published in Language Testing, an international journal, in 2013. Based on this, LAL became an increasingly important issue for scholars and publication progressed to the second stage, during which scholars continued exploring LAL theoretical frameworks applicable to different testing stakeholders and contexts [6,20]. As a result, LAL research has become more refined, expanding empirically. Scholars have conducted regional analyses and specific case studies centered on teachers' LAL [6,9,27,28] and investigated multiple aspects of LAL training [12]. In 2017, the International Language Testing Association (ILTA) conference focused on developing relevant research, with LAL as the theme. Thus, LAL has attracted considerable attention and the number of publications has increased rapidly. During the third research stage, from 2018 to 2021, scholars published richer regional and empirical studies [21,29] and critical reviews of LAL research development [22,44,45]. The number of citations is lower than the number of published articles, reflecting the research field's developmental stage.
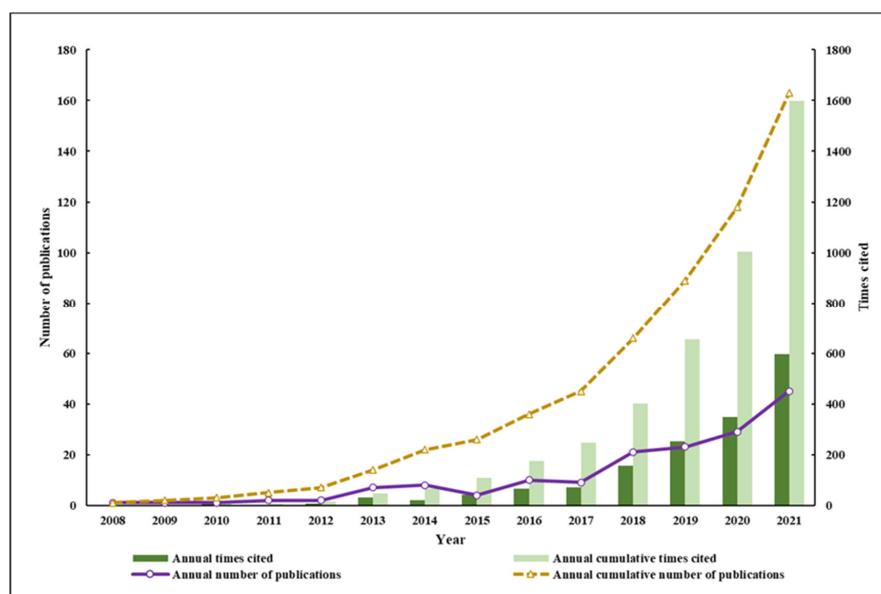

Figure 1. Yearly quantitative distribution of LAL literature in WoS from 2008 to 2021.

Figure 1 also shows the consistency of citations and the number of annual published articles. From 2008 to 2012, when language test literacy research was in its infancy, the number of citations was low. From 2013 to 2017, citations fluctuated and increased as language test literacy research became more abundant and recognized. From 2018 to 2021, an increase in the number of citations reflects the explosive growth of LAL research as the results of the previous stages were comprehensively deepened and cross-analyzed. In summary, LAL research has shown continuous growth.

3.2. Dual-Map Overlay

A dual-map overlay was formed by superposing one map drawn by CiteSpace onto another to reflect the intersection of the citing and cited clusters. The dual-map overlay facilitates researchers' understanding of the data's knowledge sources and their interactions [46]. The citing and cited overlay are displayed, showing the discipline distribution of LAL research, clustered by the Blondel algorithm in the same graph. Figure 2 presents the map after the operation. The citing map is shown on the left, and the cited map is on the right. Arcs between the two maps represent citation links, indicating the citation relationship between citing and cited journals. Citation links are depicted when the colors of arcs are consistent with the colors of their source clusters. Connecting lines represent co-citation links across cross-disciplinary boundaries. The ellipses (at the bottom right half of the graph) indicate the number of publications and authors in the LAL research field.
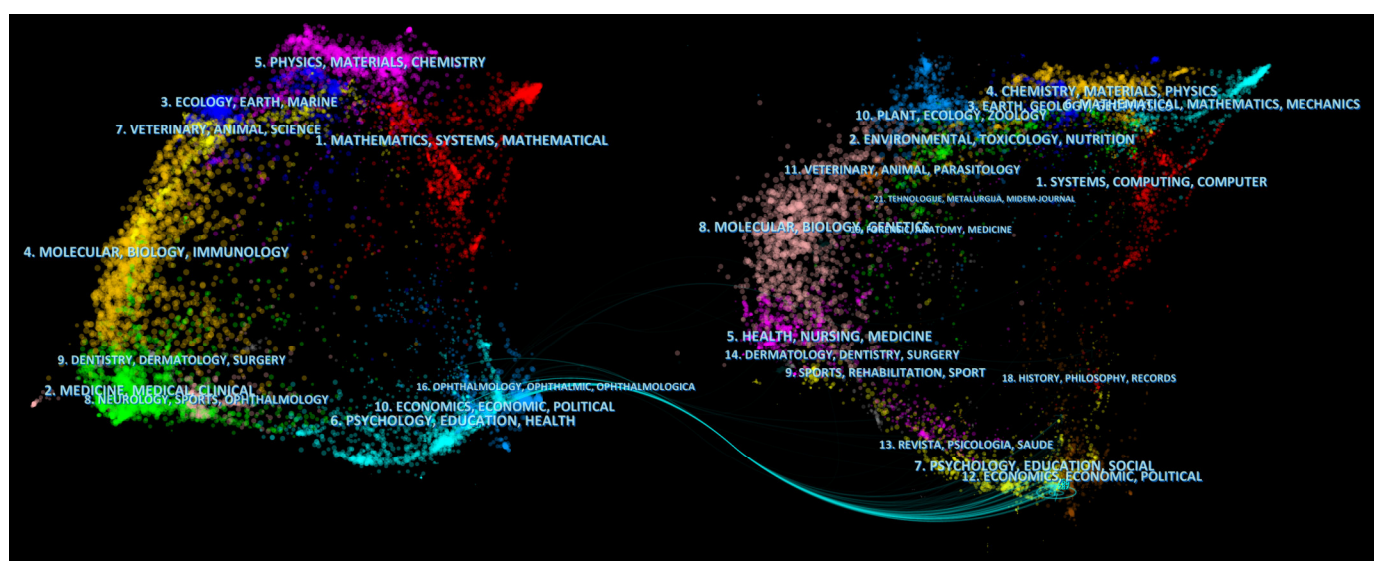

Figure 2. Dual-map overlay for LAL research.

For the number of publications, the LAL academic research results are primarily concentrated on the left side ("4 MOLECULAR, BIOLOGY, IMMUNOLOGY," "2 MEDICINE, MEDICAL, CLINICAL" and "6 PSYCHOLOGY, EDUCATION, HEALTH"). For the citation relationship, the arcs are concentrated on the left side ("6 PSYCHOLOGY, EDUCATION, HEALTH" and "10 ECONOMICS, ECONOMIC, POLITICAL"), and the right side ("7 PSYCHOLOGY, EDUCATION, SOCIAL" and "12 ECONOMICS, ECONOMIC, POLITICAL"). These concentrations indicate that LAL's impact has primarily been in the social sciences. Psychology, education, political science, and economics are the main LAL research fields and knowledge sources.

3.3. Most Cited Papers

The citation frequency of an article is an essential criterion for reflecting its value, since highly cited articles can become the knowledge base of research in a particular field. A review of these articles serves to clarify research trends in LAL [47]. Table 1 provides

the citation information for the ten most-cited articles. Xu and Brown [6] provided a scoping review of LAL studies by connecting educational assessment and teacher education and proposing putting a new conceptual framework for teacher assessment literacy into practice. Three articles published between 2008 and 2012 underpinned LAL studies by providing unique content and frame materials for LAL studies [19,48] and reviewing LAL's and stakeholders' roles in applied linguistics, education, and society [42]. Since 2013, regional studies in specific social contexts and studies in practice have emerged. Vogt and Tsagari [28] gathered data from seven European nations to determine the existing proficiency of foreign language teachers in LAL and to pinpoint the areas in which they still require training. Additionally, Lam [27] collected data from five Hong Kong teacher-training institutions to look at the state of LAL instruction. Moreover, research with a practical orientation refines research objectives. For example, Scarino [25] used sociocultural theories of learning to evaluate the process of developing LAL; Crusan et al. [9] concentrated on teachers' literacy in writing assessments; and Malone [49] used a control test with a group of US language instructors and testers to demonstrate the difficulties of integrating the technical material deemed necessary by testers while fulfilling the actual and practical demands of teachers.

Table 1. Most-cited publications from 2008 to 2021.

Rank	Publications: Author (Year) Title	Cites	Ref.
1	Xu, Y. and Brown, G. 2016. Teacher Assessment Literacy in Practice: A Reconceptualization	140	[6]
2	Fulcher, G. 2012. Assessment Literacy for the Language Classroom	111	[48]
3	Inbar-Lourie, O. 2008. Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses	99	[19]
4	Taylor, L. 2009. Developing Assessment Literacy	92	[42]
5	Scarino, A. 2013. Language Assessment Literacy as Self-Awareness: Understanding the Role of Interpretation in Assessment and in Teacher Learning	78	[25]
6	Vogt, K. and Tsagari, D. 2014. Assessment Literacy of Foreign Language Teachers: Findings of a European Study	78	[28]
7	Taylor, L. 2013. Communicating the Theory, Practice and Principles of Language Testing to Test Stakeholders: Some Reflections	63	[20]
8	Crusan, D., Plakans, L. and Gebril, A. 2016. Writing Assessment Literacy: Surveying Second Language Teachers' Knowledge, Beliefs, and Practices	61	[9]
9	Lam, R. 2015. Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy	50	[27]
10	Malone, M. 2013. The Essentials of Assessment Literacy: Contrasts between Testers and Users	45	[49]

3.4. Co-Authorship

3.4.1. Author

To a certain extent, author co-operation analysis can reveal the core authors in a field and the degree of association between authors [50]. CiteSpace was used analyze the authors of document publications by selecting the node type as "Author" and setting the configuration parameters to g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0. Figure 3 shows the map created by the applied operations. In the CiteSpace author co-operation graph, the nodes represent the authors, the size of the nodes reflects the number of the authors' articles, the thickness of the lines between the nodes reflects the strength of the co-operative relationship, and the color of the nodes reflects the years in which authors publish. Figure 3 shows that there are 157 nodes (authors) with 90 links and a density of 0.0073. The color distribution of nodes represents the year of publication. The ten most highly cited authors from 2008 to 2021 are led by Frank Giraldo, Dina Tsagari, Karin

Vogt, and Ali Isik (see Table 2). According to Price's law and the formula for core authors (m = 0.749 (nmax $\frac{1}{2}$) = 3.745), authors with four or more publications are core authors [51]. When Figure 3 is combined with Table 2, it is clear that there are four considerably larger nodes in Figure 3 represented by Frank Giraldo, Dina Tsagari, Karin Vogt, and Ali Isik. These are the core authors. The rest of the map shows independent author groups and two- and three-person collaborations. However, the linkage strength between the nodes is weak. There are more isolated nodes, and there is no linkage between author groups and two-person collaborations, indicating that the collaboration strength between LAL-domain authors needs to be enhanced. At the same time, the figure shows that Lynda Taylor and Ofra Inbar-Lourie were the earliest to study the field, with more authors publishing later, indicating that LAL's research interval is not long and that there has been considerable progress and development in recent years.
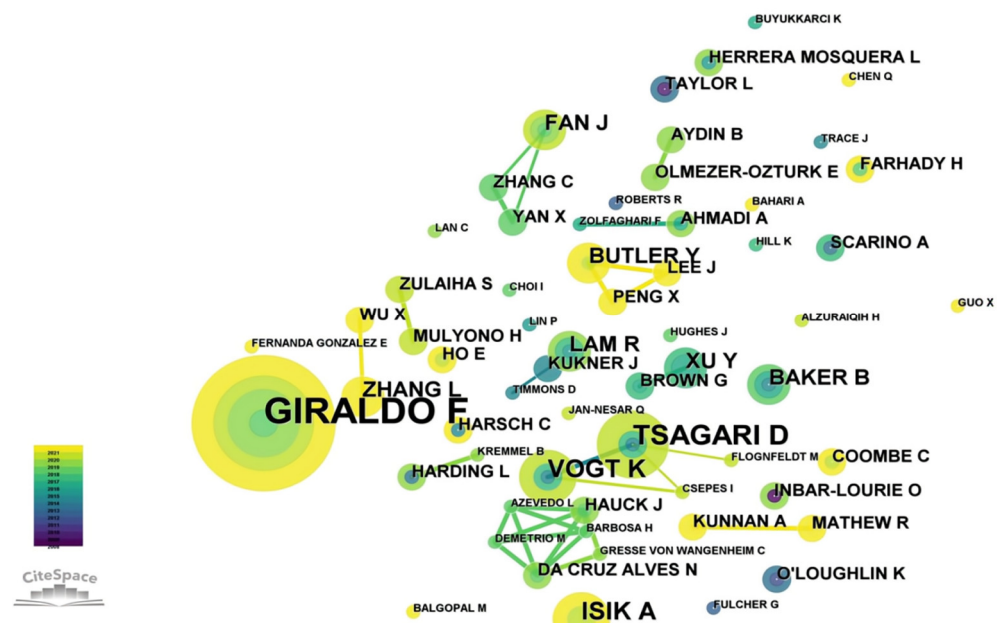

Figure 3. Author network map.

Table 2. Authors with the most published articles from 2008 to 2021.

Count	Year	Authors
10	2018	Frank Aristizábal Giraldo
5	2014	Dina Tsagari
4	2014	Karin Vogt
4	2020	Ali Isik
3	2020	Yuko Goto Butler
3	2021	Lawrence Jun Zhang
3	2018	Jason Jinsong Fan
3	2016	Yueting Xu
3	2015	Ricky Lam
3	2014	Baker Beverly

3.4.2. Institution

An author studies a subject by publishing papers or participating in academic activities, while institutional articles reflect the institution's research achievements [52]. An institutional analysis of the published documents can show the research directions of different institutions in a certain field in a given period and the degree of attention an institution's scholars paid to a domain field. Additionally, the number of publications and the frequency of citations can reflect the institution's research status and results in the discipline.

CiteSpace was used to select the node type "institution" and the configuration parameters were set to g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0, to analyze the institutions that published the documents. Figure 4 shows the map created after the operations were applied. The nodes in the CiteSpace institutional co-occurrence map represent the research institutions, the node sizes reflect the number of institutional publications, the linkages between the nodes reflect the strength of institutional partnerships, and the node colors represent year of institutional publication. Figure 4 shows 138 nodes (institutions), 88 links, and a density of 0.9655. In Figure 4, there are eight large visible nodes representing the Islamic Azad University, University of Melbourne, University of Caldas, Hong Kong Baptist University, and University of Auckland. The institutions centered on the Islamic Azad University and the University of Melbourne; institutions that have strong research capabilities and large talent pools in LAL fields. The eight most highly cited institutions are all universities from different countries, indicating that universities contribute considerably to international LAL research. In addition, from Figure 4, most institutions are isolated nodes, except for the University of Melbourne, Shiraz University, and the Chinese University of Hong Kong, which each constitute small and decentralized collaborative networks. Furthermore, the connections between the collaborative networks are not strong. This indicates that LAL research has yet to reach a high degree of international co-operation and that inter-institutional collaboration needs to be strengthened. In terms of node color, most of the research institutions started their research in later years. However, Tel Aviv University and the University of Melbourne began their research earlier and have continued to work on it more extensively. This result aligns with LAL's late birth in 2001.
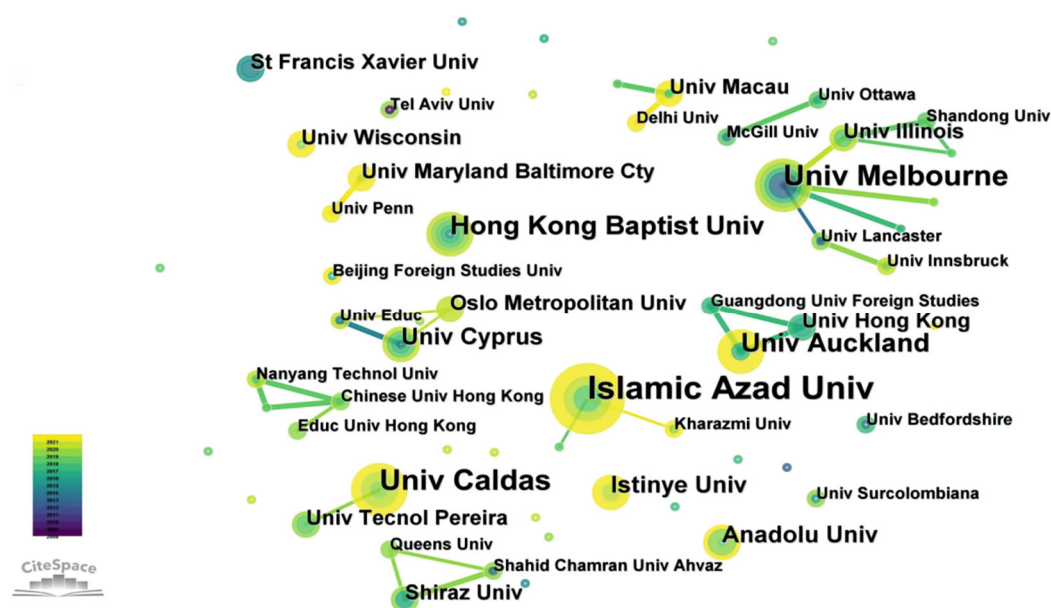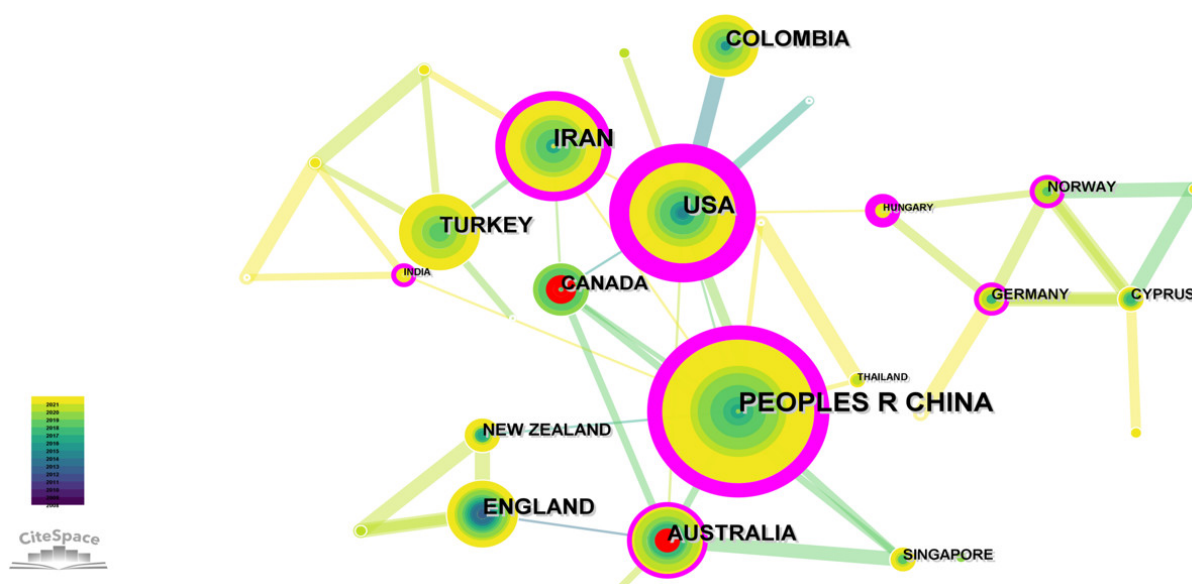

Figure 4. Institution network map.

3.4.3. Country/Region

An analysis of the countries and regions where publications originate can show the focus of scholarly emphasis in specific fields in different countries [53]. CiteSpace was used to select the node type "Country", and the configuration parameters were set as: g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0. Figure 5 shows the map created after the operations were set. There are 39 nodes (countries/regions) and 49 connecting lines with a density of 0.0661. In the CiteSpace country co-occurrence map, the nodes represent countries, the node sizes reflect the number of country/region issuance, the connecting lines between the nodes reflect the strength of the co-operative relationship between countries/regions, the node colors represent the year of country/region issuances, and the nodes' purple outer circles indicate a significant intermediary centrality the country

or region that has strong influence. The red dots indicate high burst intensity, marking whether the country or region has developed quickly in the related field. The nodes in the chart representing China, the USA, Iran, and Turkey are sizably significant, showing that although their publication record begins later than England's, their development is fast with a lead volume of publication. China, the USA and Iran have purple outer circles, indicating their significant intermediary centrality, with the highest intermediary centrality of 0.71 observed in the USA, indicating a broader scope of co-operation in LAL research. Additionally, Australia and Canada are emergent points with high emergent intensities, suggesting the rapid development of LAL research in both countries over a certain period. When compared to the institution network map, the figure's co-operation network is clearer and some of the connecting lines are thicker, indicating that LAL research has many levels of international co-operation. However, there are still some countries that appear less engaged in international collaboration than others, indicating a need to strengthen collaborative research in LAL. Geographically, LAL research involves Asia, North America, Europe, South America, and Oceania, showing that LAL research is valued in different countries and regions. Most of the countries/regions with a higher publication volume have developed economies or have economies that are undergoing rapid development with considerable momentum.


Figure 5. Country/region network map.

3.5. Co-Occurrence Analysis

3.5.1. Terms

A co-occurrence analysis of terms can show common LAL themes and reveal the evolution of research topics, hotspots, directions over time, and weaknesses [54].

CiteSpace was used to conduct a term co-occurrence analysis based on the dataset of 163 documents. The node type was set as "Term" with the following configuration parameters: g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0. Figure 6 shows the map after the operations. There are 386 nodes and 1509 connected lines with a density of 0.0203. In Figure 6, the nodes represent terms, the node sizes represent the number of occurrences, the colors represent the year of co-occurrence, the nodes' purple outer circles represent their mesoscopic centrality, and the red dots represent emergent points. Figure 6 shows that LAL studies' main themes are "assessment literacy," "language assessment literacy," "language testing," "assessment," "language teachers," and "professional development". These indicate that LAL research topics are relatively concentrated and mostly related to "language," "assessment literacy," "testing," and "teachers". These findings reflect that the main subject of LAL research is "teachers," and that the locations are primarily

"school" and "classroom". "Semi-structured interviews" were the primary method used for empirical studies, showing that small-sample qualitative studies are relatively popular in the LAL field. Additionally, the "assessment literacy," "language assessment literacy," and "language testing" nodes have purple outer circles, indicating that these nodes are more central. In contrast, other nodes are scattered across the network map, suggesting that, with continued research development, researchers' concerns are becoming more diversified while several core themes are still maintained. The outstanding position of "teachers" on the map reflects that the main subject of LAL research is "teachers," a leading stakeholder group implementing LAL in actual practices. Considering time co-occurrence nodes, co-occurrence terms with large nodes received attention during 2014 and 2021. LAL rapidly developed in assessment literacy and language research after 2014. "Teachers" and "language testing" experienced similar growth.
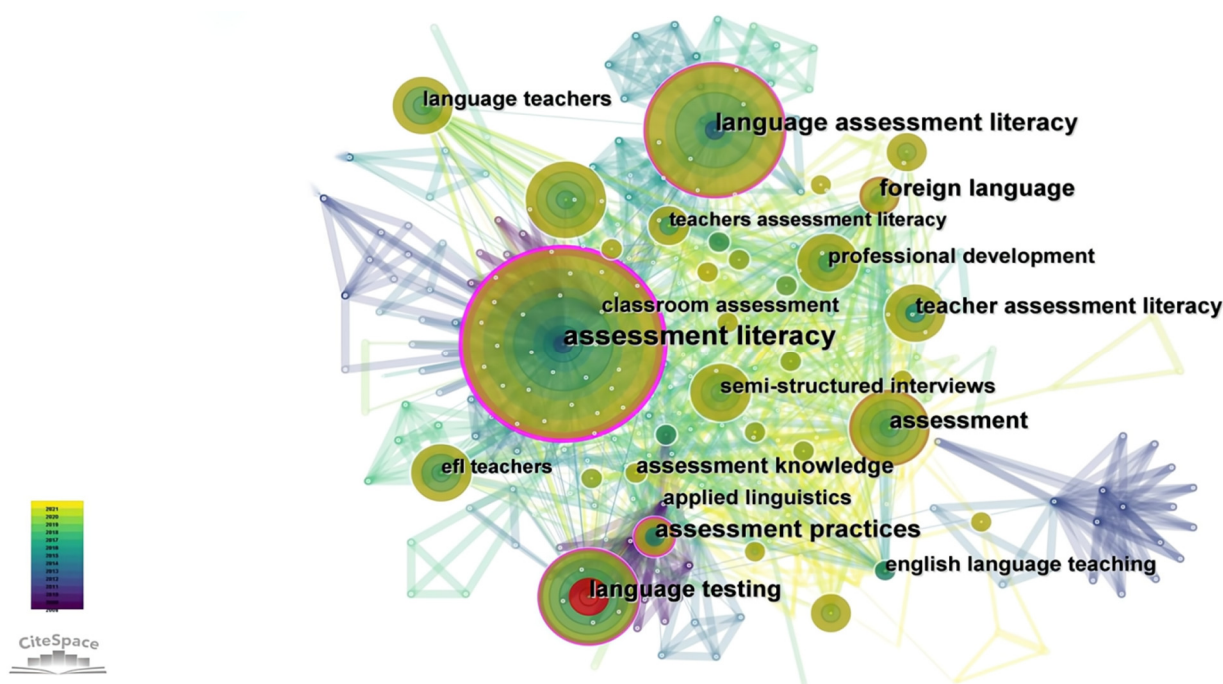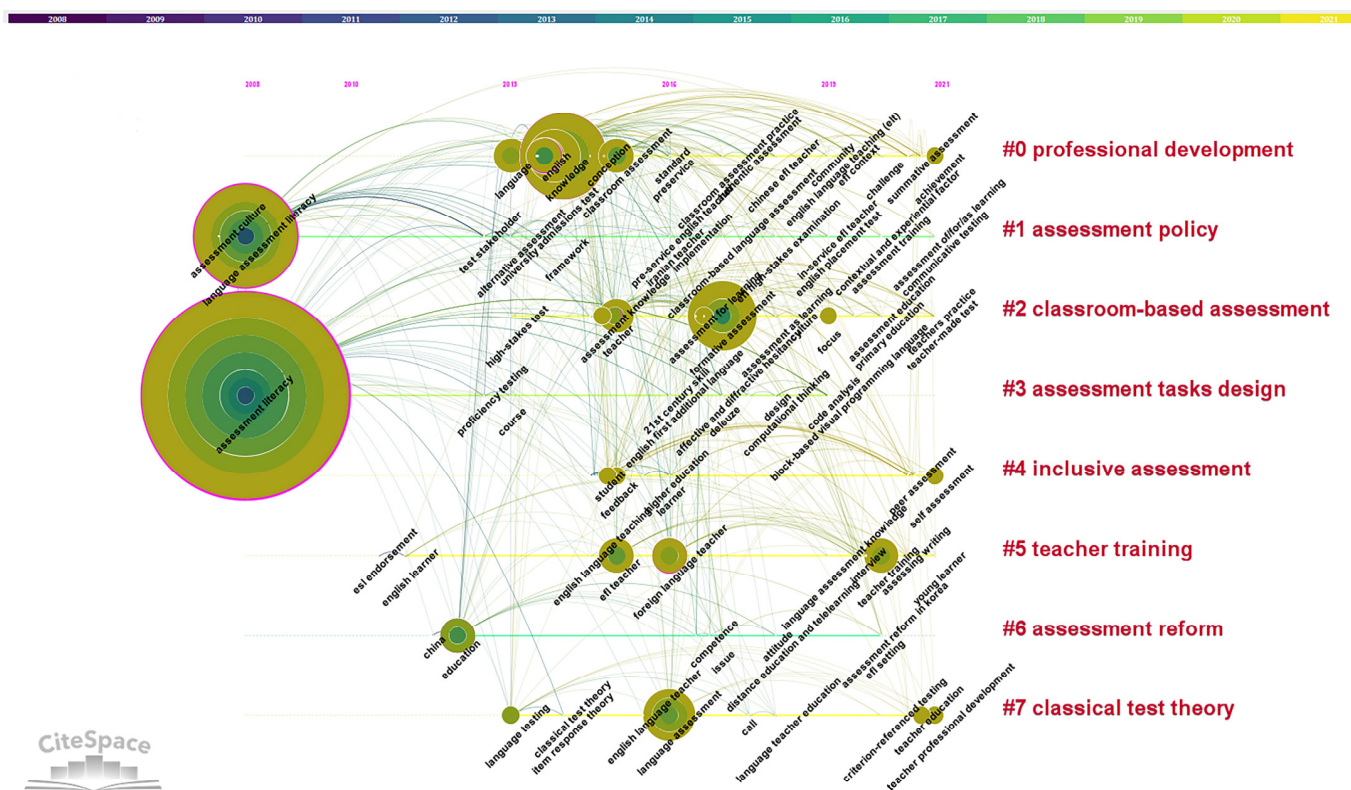

Figure 6. Term co-occurrence network map.

3.5.2. Keyword

Keywords symbolize highly condensed academic views, define an article's content, and reflect research status, concerns, and hotspots. Keyword co-occurrence analysis demonstrates the frequency of keyword co-occurrence and connections between keywords. This can help researchers analyze research hotspots in different periods, grasp changes in research hotspots and future research trends, and organize knowledge structures [55]. Furthermore, compared to term co-occurrence, keyword co-occurrence results could reveal more targeted information about what has been studied due to the specific role "keywords" play in a document. For example, term co-occurrence analysis fails to indicate the role of "students" as one of the stakeholders of LAL, but keyword co-occurrence analysis does. This study used the "co-word analysis method" in CiteSpace. The node type chosen was "Keyword" and the parameters were set as: g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0. Figure 7 shows the map created after the operations. There are 230 nodes and 970 links with a density of 0.0368. In the CiteSpace keyword co-occurrence mapping, the nodes represent keywords, node sizes represent the frequency of keywords, links represent the co-occurrence of keywords, and node colors represent the year of keyword co-occurrence. Figure 7 shows that the keyword nodes and the lines between them are dense, indicating that the LAL research (2008–2021) is relatively rich, the links between the keywords are close, and the research focus is relative concentrated and in the stage of

vigorous development. Meanwhile, "assessment literacy," "language assessment literacy," "assessment," and "culture" nodes are large, indicating that these keywords signify essential LAL research areas. When analyzing the time of keyword co-occurrence node, LAL topics have been enriched and increasingly refined through continuous development since late 2001. For example, detailed assessment methods have been explored to enhance LAL knowledge structures and practices. The term "alternative assessment" appeared around 2014, "formative assessment" occurred about 2018, and "summative assessment" and "peer/self-assessment" appeared around 2021. In addition, "knowledge," which emerged after 2013, formed a large node indicating an emergent LAL research trend and a change in LAL research focus. Combined, LAL research is manifold, with this feature becoming increasingly apparent in recent years. Therefore, it is feasible to expect future disciplinary growth to be more multifaceted and to see potential LAL breakthroughs. Researchers can also consider keywords missing from the high-frequency terms shown in Figure 7 to fill specific research gaps. For instance, China and Korea are the only two countries displayed in Figure 7. Therefore, it is worthwhile to consider why they are referred to more frequently in LAL research and the research conditions than in other countries.
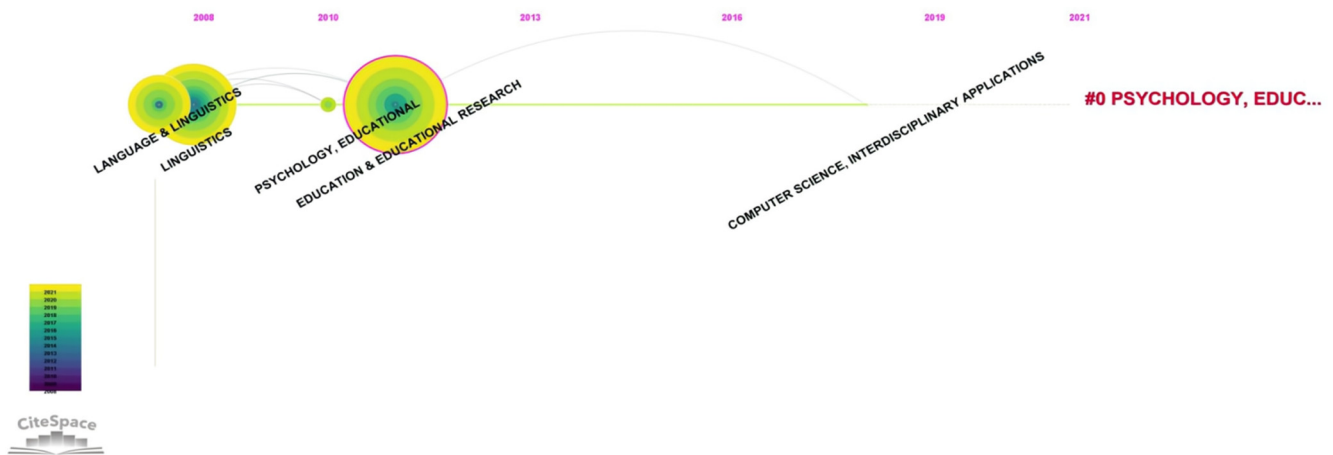

Figure 7. Keyword co-occurrence network map.

3.5.3. Categories

A co-occurrence analysis of scientific domains cited in the literature can reveal the structure of the research field's scientific domains. CiteSpace was used to make the scientific domain co-occurrence analysis more intuitive. The node type selected was "Category" and the configuration parameters were set as: g-index (k = 25), LRF = 3.0, LBY = 5, and e = 1.0. Figure 8 shows the map after the operations were set. The map contains 10 nodes with 12 lines and a density of 0.2667. Figure 8 shows several large nodes, including Education and Educational Research, Linguistics, Language and Linguistics, and Psychology, Educational, indicating that LAL research's core focus is in the language and education fields. Table 3 shows the co-occurrence count of these nodes. The table shows that LAL research is not limited to "Education and Educational Research," [44] "Linguistics," [8] "Language and Linguistics," [56] and other disciplines that are highly relevant to LAL but also involve

"Environmental Sciences," [57] "Green and Sustainable Science and Technology," [57] "Psychology Multidisciplinary" [58] and other disciplines that are not directly linked to the field. These indicate that LAL is an interdisciplinary concept and that researchers should consider this interdisciplinary nature when conducting their LAL research to ensure comprehensive results. Table 4 also shows LAL's development between 2008 and 2021. "Language and Linguistics" and "Linguistics" first emerged in 2008, and "Psychology, Educational, Education" and "Educational Research" followed suit. However, "Computer Science," "Interdisciplinary Applications," and other disciplines that are not directly or strongly linked to LAL only started to appear around 2018. As such, LAL started gradually developing into an interdisciplinary concept in 2001, when it became independent of general education fields. Though the cooperation of areas that seem irrelevant to LAL is limited, scholars have endeavored to enrich LAL study possibilities and expand scientific domains in which LAL can be applied or cooperated with. While LAL research is relatively young, it keeps evolving, meaning that researchers can look for new research hotspots and directions based on LAL's interdisciplinary nature and potential.

Figure 8. Category co-occurrence network map.

Table 3. Most highly cited categories from 2008 to 2021.

| Count	Centrality	Year	WoS Categories
111	0.17	2011	Education and Educational Research
78	0.00	2008	Linguistics
49	0.00	2008	Language and Linguistics
14	0.00	2010	Psychology, Educational
1	0.00	2021	Environmental Sciences
1	0.00	2021	Green and Sustainable Science and Technology
1	0.00	2021	Psychology, Multidisciplinary
1	0.00	2018	Computer Science, Interdisciplinary Applications
1	0.00	2017	Cultural Studies
1	0.00	2021	Environmental Studies

Table 4. Major clusters of co-cited LAL references.

Cluster ID	Size	Silhouette	Label (LLR)	Average Year
0	88	0.863	teacher training	2019
1	57	0.883	foreign language teachers	2016
2	55	0.948	International English Language Testing System	2010
3	40	0.966	assessment practices	2015
4	32	0.943	dialogic feedback	2020
5	27	0.999	teacher self-awareness	2008
6	26	0.995	knowledge base	2014
7	23	0.983	L2 writing	2018

3.6. Co-Citation Analysis

3.6.1. Document Co-Citation Network

Two documents that appear together in the reference list of a third citing document form a co-citation relationship [59]. Based on the co-citation relationship between documents, CiteSpace can create a cluster of documents with close co-citation relationships and macroscopically display several specific fields of academic research. The cited document node size and the thickness of links indicate the number of citations and co-citations, respectively [46]. In this study, a document co-citation analysis was conducted on 163 studies closely related to LAL to visualize and analyze the LAL field's research status, allowing for an understanding of LAL's knowledge foundations (co-cited literature collection) and research frontiers (cited literature collection). The utilized CiteSpace configurations are as follows: g-index (k = 25), Timespan (2008–2021), LRF = 3, LBY = 5, e = 1, and Pruning = none. Figure 9 shows the network of co-cited references in LAL research includes 431 nodes and 1,637 links. The map's modularity is 0.7833, which shows that the co-citation map is well structured and reflects LAL-related research areas. After clustering, the following eight categories emerge: #0 teacher training, #1 foreign language teachers, #2 International English Language Testing System, #3 assessment practices, #4 teacher self-dialogic feedback, #5 teacher self-awareness, #6 knowledge base, and #7 L2 writing. These categories represent the eight main research areas in the LAL field. The colors of each cluster in Figure 9 represent the average publication year of the relevant studies in a cluster. For example, the largest cluster (#0 teacher training) appears bright yellow, indicating that the studies in the cluster are newer, with an average publication year of 2019. The quality of co-citation clusters is supposed to meet the criteria of modularity and weighted mean silhouette [60]. The modularity metric represents a network's tight or loose internal structure. A silhouette score represents to what degree cluster members are lumped together based on several aspects of commonality. The higher the score is, the more meaningful a cluster. The results in Table 4 were obtained using the LLR algorithm. The lowest silhouette among the eight significant clusters formed by all the citations in the LAL field literature is 0.863, and the average silhouette reaches 0.9475, indicating that the formed clusters are highly reliable. Clusters #0, teacher training, and #4, dialogic feedback, were composed of 88 (2019) and 32 nodes (2020), respectively, making them the two most active research directions in the last three years of the study range.
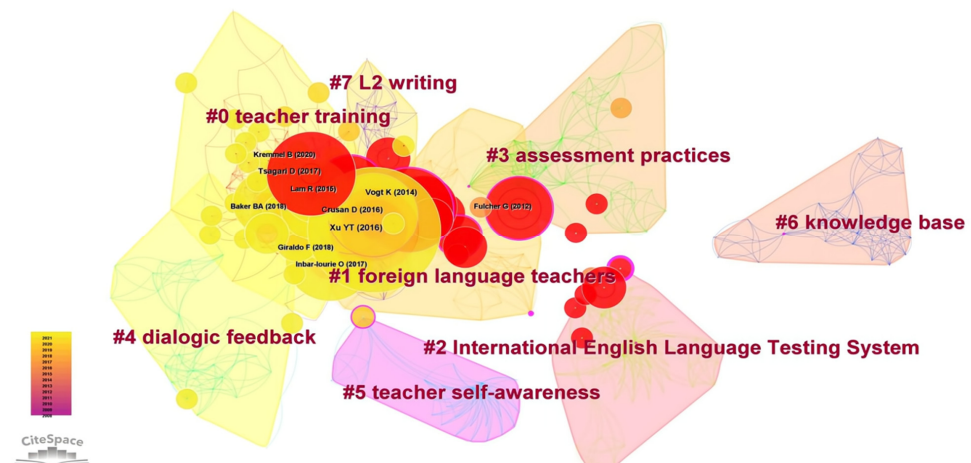
Figure 9. Co-citation LAL network clusters.

Citation counts of documents can show their importance and value in research fields and further indicate the significance of the clusters they belong to. Citation counts of the cited documents were calculated and analyzed, and detailed information about the ten most cited is displayed in Table 5. Six of these highly cited articles or monographs are clustered in #0 teacher training, which suggests that LAL training for teachers has received widespread academic attention. Three of the ten are clustered in the category #1 foreign language teachers and one of the three was cited the most often, demonstrating researchers' significant interest in foreign language teachers when exploring LAL. Though the citation count of the work of Crusan [9], clustering in #7 L2 writing, ranks second, Cluster #7 has no obvious advantage compared to Clusters #0 and #1 since only one document is linked. In summary, Clusters #0 and #1 attract more researcher attention. The focus on the two clusters proves the significance of "teachers" as one of the LAL stakeholders among LAL studies.

Table 5. Top 10 LAL citation counts.

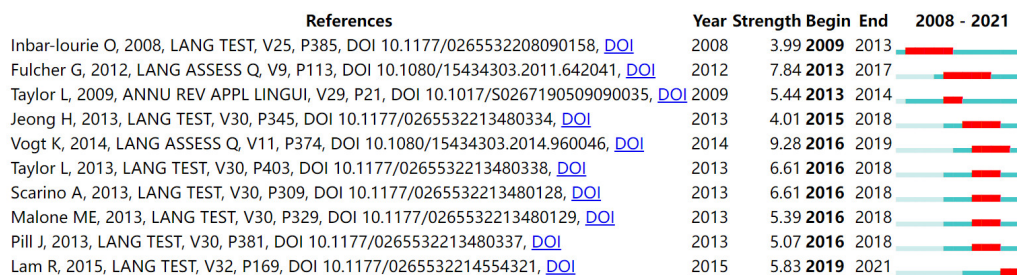| Citation Counts	Publications: Author Year Title	Cluster ID	Ref.
35	Xu, Y. and Brown, G. 2016. Teacher assessment literacy in practice: A reconceptualization	1	[6]
26	Crusan, D., Plakans, L. and Gebril, A. 2016. Writing assessment literacy: Surveying second language teachers' knowledge, beliefs, and practices	7	[9]
25	Tsagari, D. and Vogt, K. Assessment Literacy of Foreign Language Teachers around Europe: Research, Challenges and Future Prospects	0	[61]
24	Vogt, K. and Tsagari, D. 2014. Assessment Literacy of Foreign Language Teachers: Findings of a European Study.	1	[28]
19	Giraldo, F. 2018. Language Assessment Literacy: Implications for Language Teachers	0	[44]
17	Fulcher, G. 2012. Assessment Literacy for the Language Classroom	1	[48]
17	Lam, R. 2015. Language assessment training in Hong Kong: Implications for language assessment literacy	0	[27]
17	Kremmel, B. and Harding, L. 2020. Towards a Comprehensive, Empirical Model of Language Assessment Literacy across Stakeholder Groups: Developing the Language Assessment Literacy Survey	0	[22]
17	Inbar-Lourie, O. 2017. Language assessment literacy	0	[62]
17	Baker, B. A. and Riches, C. 2018. The development of EFL examinations in Haiti: Collaboration and language assessment literacy development	0	[21]

Centrality is depicted by establishing a "bridge" between two unrelated nodes. Different from citation counts, which indicate the significance of a single document in certain fields, centrality highlights a single document's structural importance; that is how closely it is connected with other documents of the same dataset. We calculated the centrality of the cited documents (see Table 6). These high-centrality articles serve as "bridges" in the LAL field and have practical guiding significance to and promotion effect on the construction or revision of LAL research framework. For example, high-centrality articles help efficiently identify research focuses and track related research.

Table 6. Top 10 LAL centrality scores.

Centrality	References: Author Year Title	Cluster ID	Ref.
0.21	Lantolf, J. P. and Poehner, M. E. 2008. Dynamic assessment	1	[63]
0.21	Malone M.E. 2008. Training in Language Assessment	2	[64]
0.19	Council of Europe. 2020. Common European Framework of Reference for Languages: Learning, Teaching, Assessment—Companion Volume	5	[65]
0.19	Fulcher, G. 2012. Assessment Literacy for the Language Classroom	1	[48]
0.18	Klinger et al. 2015. The Classroom Assessment Standards for PreK12 Teachers	3	[66]
0.17	Brown, J.D. and Bailey, K. M. 2008. Language Testing Courses: What are They in 2007?	2	[67]
0.16	Lam, R. 2015. Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy	0	[27]
0.16	Jeong, H. 2013. Defining Assessment Literacy: Is It Different for Language Testers and Non-Language Testers?	1	[5]
0.15	Borg, S. 2009. English Language Teachers' Conceptions of Research	6	[68]
0.13	Vogt, K. and Tsagari, D. 2014. Assessment Literacy of Foreign Language Teachers: Findings of a European Study	1	[28]

In CiteSpace, a burst represents a frequency surge of a language phenomenon, such as a surge of an article being cited. Sigma is a metric of a node in a network of cited references and is calculated based on burstiness [69]. Sigma can be used to identify innovative literature by highlighting significant nodes with rapid citation growth. Burst and Sigma analyses based on the dataset were conducted to discern the LAL field's research trends and to predict future creative research directions. An emerging field may signal a topic's direction and the results obtained using the burst algorithm could identify topic trends over time. Figure 10 reflects the ten references with the most robust citation bursts. The beginning of a blue line refers to the time the document was published, and the red bar depicts the burst period in which the document was cited. Innovative research provides the foundation for future research. As such, it tends to be highly cited. As a frontier scholar in the field of LAL, Taylor is the only researcher who authored two articles on the list. The duration of Taylor's bursts [42] is shorter than that of the other nine articles when comparing the burst citation periods of the listed articles with the most robust citation bursts. Taylor [42] demonstrates the significance of encouraging knowledge, skill sharing, and good quality assessment when examining the research content of the listed articles. Taylor's articles are micro-research focused when compared to the focuses of the other listed articles, which are more specific and practical. This phenomenon indicates that research with a more specific study orientation might have longer burst times in the LAL field. Another possible reason is that scholars may have more interest in teachers and elements related to teachers (see Sections 3.5.1 and 3.6.1). The articles with the longest burst citations are Inbar-Lourie [19] and Fulcher [48], whose research focuses on courses and classrooms, respectively. These are followed by Jeong [5] and Vogt and Tsagari [28], whose research focuses on testers and foreign language teachers, respectively. Among the

four, three articles are linked to the subject of "teacher". Lam [27] is the only citation still in the process of bursting, which indicates that training for teachers' LAL is still the topic of primary academic focus.

References	Year	Strength	Begin	End	2008 - 2021
Inbar-lourie O, 2008, LANG TEST, V25, P385, DOI 10.1177/0265532208090158, DOI	2008	3.99	**2009**	2013	
Fulcher G, 2012, LANG ASSESS Q, V9, P113, DOI 10.1080/15434303.2011.642041, DOI	2012	7.84	**2013**	2017	
Taylor L, 2009, ANNU REV APPL LINGUI, V29, P21, DOI 10.1017/S0267190509090035, DOI	2009	5.44	**2013**	2014	
Jeong H, 2013, LANG TEST, V30, P345, DOI 10.1177/0265532213480334, DOI	2013	4.01	**2015**	2018	
Vogt K, 2014, LANG ASSESS Q, V11, P374, DOI 10.1080/15434303.2014.960046, DOI	2014	9.28	**2016**	2019	
Taylor L, 2013, LANG TEST, V30, P403, DOI 10.1177/0265532213480338, DOI	2013	6.61	**2016**	2018	
Scarino A, 2013, LANG TEST, V30, P309, DOI 10.1177/0265532213480128, DOI	2013	6.61	**2016**	2018	
Malone ME, 2013, LANG TEST, V30, P329, DOI 10.1177/0265532213480129, DOI	2013	5.39	**2016**	2018	
Pill J, 2013, LANG TEST, V30, P381, DOI 10.1177/0265532213480337, DOI	2013	5.07	**2016**	2018	
Lam R, 2015, LANG TEST, V32, P169, DOI 10.1177/0265532214554321, DOI	2015	5.83	**2019**	2021	

Figure 10. References with the strongest citation bursts.

3.6.2. Author Co-Citation Network

An author co-citation relationship occurs when two scholars are cited in the same publication. This study established a network of author co-citation to identify widely recognized and highly cited scholars in LAL research. The CiteSpace configurations were set as follows: g-index (k = 25), Timespan (2008–2021), LRF = 3, LBY = 5, and e = 1. Networks were then pruned, sliced, and merged according to the pathfinder algorithm and the minimum spanning tree and pruning sliced networks, resulting in 399 nodes and 1021 links. The larger the size of each node, the more times the scholar is cited. The thicker the links between two nodes, the more citations both authors receive in the same paper.

Figure 11 shows that the ten most co-cited authors in the LAL research field are Fulcher (89), Inbar-Lourie (83), Taylor (75), Stiggins (65), Popham (65), Vogt (61), Davies (59), Scarino (55), Xu (53), and Malone (51). The high frequency of their works co-cited reflects that they stand out in LAL research, and that the value of their work is widely confirmed. The links are thicker in the left half of the network, which includes authors Fulcher, Davies, and Black, indicating that as two of the ten most-cited authors, Fulcher and Davies have closer relations with other authors in the LAL field compared to others on the list, whose link networks with others are relatively simple. Black is an exception. Although Black is not on featured on the list, their works are co-cited with others frequently. The conditions above inform researchers that extra attention needs to be paid to other significant authors if they want to look at or determine influential LAL research scholars based on the connection network of a particularly notable scholar. For example, although Inbar-Lourie is one of the most cited authors, links with other authors are limited. In addition, Fulcher might be ignored because they have no link line with Inbar-Lourie (see Figure 11). Influential scholars can also be identified using a citation burst analysis perspective, as seen in Figure 12. Figure 12 shows that Alderson (3.45, 2009–2014), Leung (3.28, 2014–2017), Black (3.85, 2015–2017), Stiggins (2.85, 2015–2018), Volante (2.73, 2015–2018), Hamp-Lyons (2.72, 2017–2018), and Harding (2.9, 2019–2021) have been heavily cited over time. Stiggins is the only author who overlaps when comparing the list in Figure 12 with the list in Figure 10, indicating that the authors leading new research bursts might not be widely recognized in the field of LAL. This informs researchers that what makes an author "important" should be defined and that said authors should be searched multi-dimensionally. Although the beginning year of the seven authors is 2008, Alderson and Harding's citation burst period differs from the others, whose citation burst periods are generally close. The burst of citing Alderson begins the earliest and continues for the longest time but concludes in 2014 with no sign of new burst momentum. The burst of citing Harding begins the latest but continues until the time the data was gathered. A burst continuing when this data was collected could reflect recent outstanding scholar(s) in the field whose research might include the latest research hotspots or trends that could be followed or advanced.
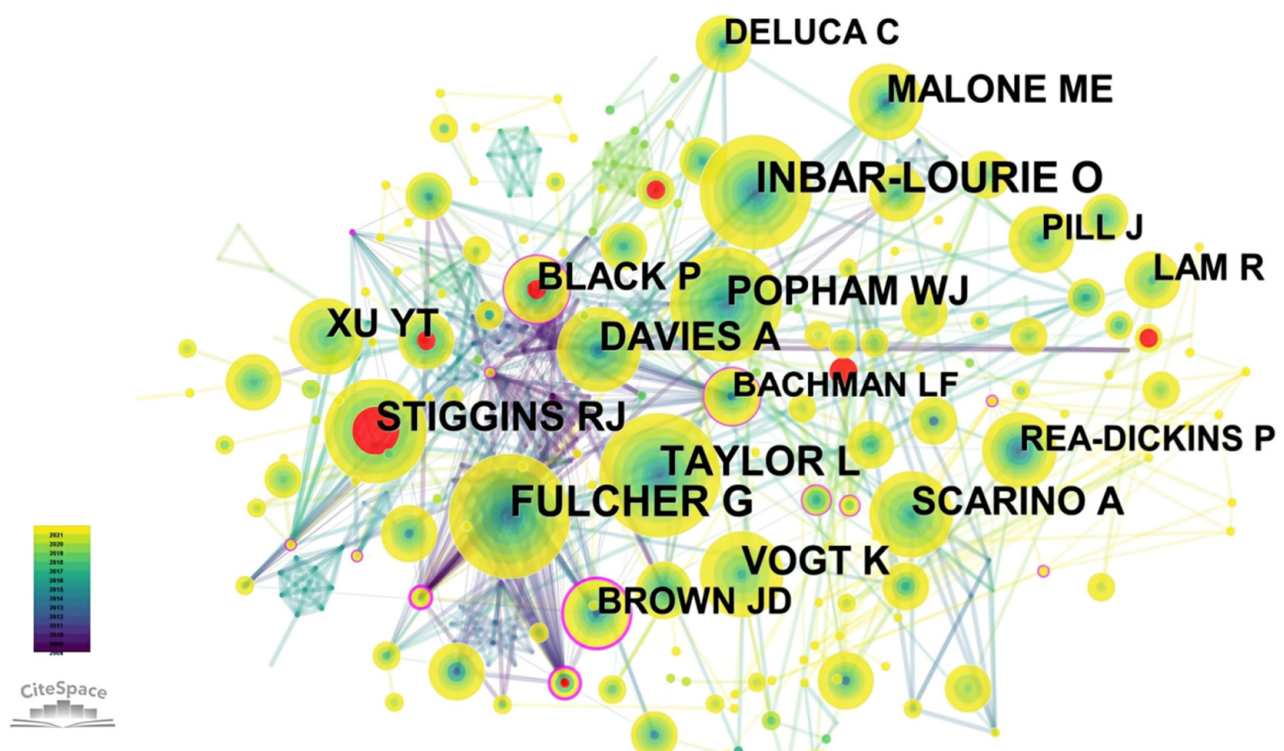
Figure 11. Author co-citation network.


Figure 12. Top cited authors with strongest citation bursts.

3.6.3. Journal Co-Citation Network

CiteSpace was adopted to generate a journal co-citation network (Figure 13) and understand the representative cited journals in the LAL field. CiteSpace configurations were set as follows: g-index (k = 25), Timespan (2008–2021), LRF = 3, LBY = 5, and e = 1. Networks were pruned, sliced, and merged using the minimum spanning tree, which resulted in 428 nodes and 722 links. The cited journal node sizes and link thicknesses indicate the number of citations and co-citations, respectively.
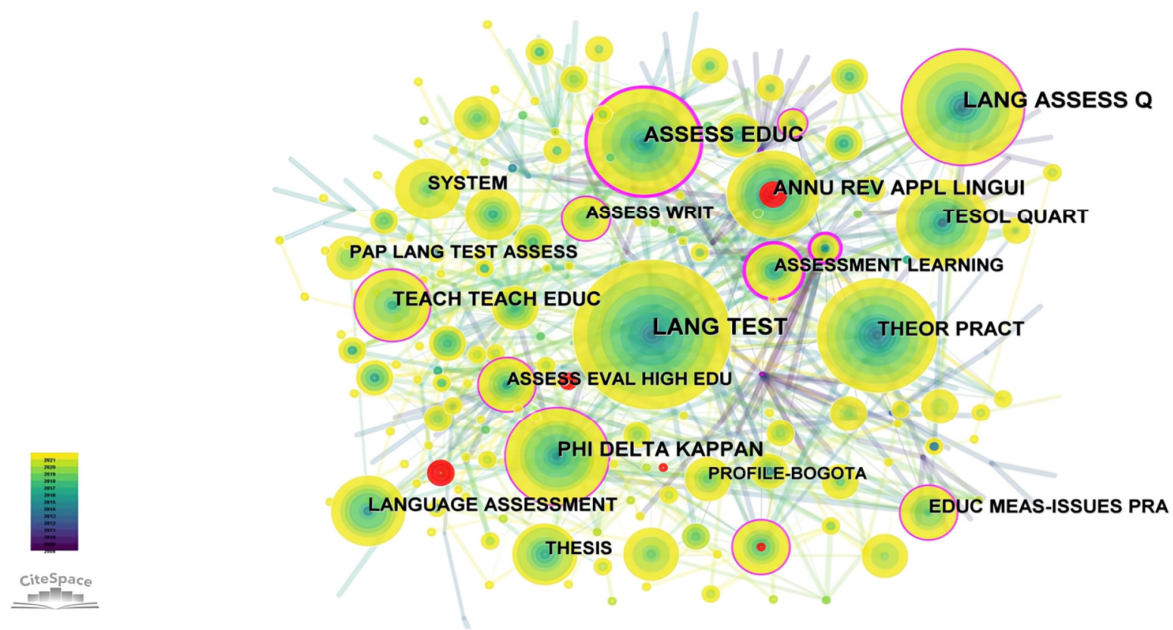
Figure 13. Journal co-citation network.

Publication sources were analyzed and the most representative cited journals in the LAL field were detected. Papers related to the LAL field were published in 432 journals, with ten most productive being listed in Table 7. Language Testing is the most productive journal, publishing 127 LAL-related papers, followed by Language Assessment Quarterly (106), Educational Assessment Evaluation and Accountability (87), Phi Delta Kappan (71), Annual Review of Applied Linguistics (62), TESOL Quarterly (58), Theory into Practice (56), Teaching and Teacher Education (54), Educational Measurement: Issues and Practice (44), and Language Assessment (43). The study used the Journal Citation Report (JCR) as the standard because a higher impact factor does not equate to higher journal quality. Based on the index of JCR, each subject category of journals is graded into four quartiles: Q1, Q2, Q3, and Q4. Q1 is the highest level and indicates a journal's high quality and significant influence. Among the listed journals, Q1 and Q2 accounted for 70%. Language Testing, Annual Review of Applied Linguistics, TESOL Quarterly, and Teaching and Teacher Education are rated as Q1. This means that most works cited by LAL research are of high quality. In addition, Language Assessment Quarterly, the second largest source of cited documents, directly contributes to the development of LAL research. That Language Assessment Quarterly produces many materials supporting LAL research proves LAL's success as an independent concept. Figure 13 displays that the network of journals the cited documents belong to is complex because their co-citation relation is sophisticated. The tangled lines in the figure indicate the nature of LAL as a multidisciplinary concept.

Table 7. Top 10 LAL journals.

Count	Centrality	Year	Journals
127	0.08	2010	Language Testing
106	0.04	2013	Language Assessment Quarterly
87	0.12	2008	Educational Assessment Evaluation and Accountability
71	0.07	2010	Phi Delta Kappan
62	0.07	2008	Annual Review of Applied Linguistics
58	0.01	2013	TESOL Quarterly
56	0.04	2013	Theory into Practice
54	0.05	2014	Teaching and Teacher Education
44	0.14	2008	Educational Measurement: Issues and Practice
43	0.02	2016	System

4. Discussion

A bibliometric analysis was conducted, and 163 English articles and reviews published over the past 13 years were chosen from the SSCI, A&HCI, and ESCI databases to review LAL's emerging trends, the intellectual landscape, critical changes, and thematic patterns.

4.1. Emerging Trends and Critical Changes

In 2001, seven years before LAL was first indexed in the WoS core collection, the CEFR introduced assessment into the language assessment field and challenged language teachers' assessment literacy. Brindley [17] proposed that language teachers' assessment literacy should be independent of the general education field in 2001, but the idea of LAL was not explicitly mentioned in their outline. 2001 to 2008 can be considered as a pre-stage of LAL research, during which researchers noticed the importance of differentiating LAL from AL but were still in the exploration stage. Through the scientometric analysis of LAL literature between 2008, when LAL was first indexed, and 2021 (20 years after the first time it was deemed necessary to separate LAL from AL), it is clear that LAL research has shown a fluctuating upward trend since the initial stagnation, thereby demonstrating a broad development prospect. This study divides the entire time range into 2008 to 2012, 2013 to 2017, and 2018 to 2021. Inbar-Lourie's [19] article marked the coming of the first stage of LAL research, where LAL was clearly defined for the first time. Inbar-Lourie further proposed that LAL needs to address three key questions: the "how-to" (skills), the "what" (knowledge), and the "why" (principles) of assessment. In the same year, ALTE discussed the research and development of foreign language teachers' LAL, and members of the association began to focus on foreign language teachers' LAL in their own countries. During the first stage, researchers put significant effort into identifying the concept of LAL and constructing the theoretical framework. However, since LAL is a multilayered entity and defining it is a significant challenge [70], relevant studies have progressed slowly and delivered few results. The second stage of LAL research has had the most considerable impact. In 2013, Taylor [20] discussed LAL stakeholder groups and introduced theoretical frameworks for each group. This article was published in Language Testing's featured special issue on LAL in 2013, a key outcome of the 33rd LTRC. The conference brought more attention to LAL, and field research moved into the second stage. From 2008 to 2017, related studies showed a fluctuating upward trend. However, at the beginning of 2018, the associated studies dramatically increased. This indicates that LAL is increasingly recognized as an independent concept, and that more attention is being paid to language discipline characteristics. During the third stage of LAL research, the theoretical framework of LAL has been developed further, and many researchers have conducted empirical studies on the status of teachers' LAL and their training needs [23,29,45]. Professional conferences in language assessment and testing have significantly contributed to the development of LAL and are closely related to emerging and critical milestones.

4.2. Intellectual Landscape of Authors

Authors are those who conduct and publish research. Past work in a particular field can be tracked and understood through their work. Focusing on important scholars in an area is a way to learn about the field. In this context, the condition of importance in the LAL research is complex. Sections 3.4.1 and 3.6.2 discuss LAL scholars from the categories of publication, citation, and co-citation. Four lists referring to authors are included in this paper: authors who are most cited (Table 1, e.g., Yueting Xu and Gavin Brown), authors who published the most articles (Table 2, e.g., Frank Giraldo), authors who are most cited together with other scholars (Figure 11, e.g., Glenn Fulcher), and cited authors with the most significant citation bursts (Figure 12, e.g., Karin Vogt and Dina Tsagari). The overlap between the four lists is exceptionally low, which means "most valued authors by other scholars ≠ authors having most articles ≠ authors having most links with others ≠ authors leading research upsurge." One possible explanation for this situation is that research work by LAL scholars' is relatively independent (Section 3.4.1), indicating that scholars usually

have their own research emphasis and citation principles and resulting in a diverse research context. Hence, when authors and author linkage networks are used in building the LAL research network, the work of scholars who are important for various reasons should be considered to guarantee the completeness of the network. These are especially the case as visualization research is increasingly employed, and scholars may focus on one dimension, like author linkage, to carry out discussions. When exploring influential authors or other intellectual landscapes, scientometric reviews contribute significantly as they can be used to analyze various indicators.

4.3. Thematic Patterns and Research Focus on Teachers

LAL is closely related to testing and most LAL research subjects have been language teachers. "Teacher" appears as one of the most frequently used words in abstract sections and as a keyword in many related studies. As essential stakeholders in teaching activities, language teachers directly participate in the assessment process and take direct responsibility for assessment results. In the context of school- and class-based assessments, teachers' LAL is highly valued by academia [8,28,70]. There are numerous factors leading to such bias. One is that teachers have versatile talents. Teachers carry out teaching, design courses, and create standards for judging and assessing students and their work. Teachers take on multiple roles in the LAL stakeholder structure. The complexity of this group endows them with higher research and practical values and more study perspectives that can be further explored. However, LAL research is still developing, and research on other stakeholders has increased [4,22,24]. The appearance of "assessment reform" in the keywords shows that attention has also been paid to student initiatives during the worldwide curriculum reform process. In other words, the evaluation of LAL has gradually shifted from only the teacher to the teacher and the student. "Peer assessment" is also used in many language learning practices, such as writing and speaking, making students' LAL a new trend in LAL research.

4.4. Interdisciplinary Nature of LAL

Through a co-occurrence analysis among terms, keywords, and categories, results show that LAL research is not only focused on language and education but also psychology, computer science, environmental sciences, and green and sustainable science; fields not directly related to LAL. As LAL research continues to advance, researchers' concerns are becoming more diverse and research topics continue to be enriched. Moreover, the journal co-citation network shows that the co-citation relation between documents is complicated. These findings indicate that LAL is a complex concept with multiple dimensions, and the current LAL research and development trends have interdisciplinary and multi-perspective characteristics. In addition, the dual-map overlay suggests that LAL is underdeveloped in various fields, and the number of multidisciplinary studies is limited. For example, testing is still the most closely related sub-discipline. Language Testing, Language Assessment Quarterly, and Educational Assessment Evaluation and Accountability have become the most recognized journals in this field. As such, the interdisciplinary nature of LAL research still leaves room for exploration.

5. Conclusions

Through the scientometric visual analysis of 163 articles and review articles from 77 journals across 10 WoS categories, the current study identifies the stage changes of LAL research by analyzing LAL-related articles published from 2008 to 2021. This study suggests that the theoretical framework of LAL is currently being constructed. LAL stakeholders, who directly influence the components and needs of LAL, are significant perspectives in exploring and developing the framework. "Teacher" appears as one of the most frequently researched stakeholder groups and research on other stakeholders has increased, making students' LAL a new trend in LAL research. Hence, future studies should encourage the analysis of different stakeholders.

It is worth noting that cross-disciplinary cooperation requires paying attention to research differences between various fields as different disciplines have different research backgrounds and methods. Issues related to LAL are encouraged to be more thoroughly and comprehensively studied and reviewed. Furthermore, precise measurement instruments, criteria, and channels must be utilized for efficient assessment. Finally, scientometric visual analyses should be encouraged in language assessment and testing. This study demonstrates the benefits and suitability of using a visualization method for clearly and systematically sorting complex disciplinary concepts and their development processes.

The research method used in this study does have certain limitations. First, the data for this study was collected from SSCI, A&HCI, and ESCI in WoS. While this study incorporated some of the most influential journals, prestigious publications from other databases may have been excluded. In addition, this study only focused on English language journal articles in the database, excluding non-English articles and different types of research such as dissertations and conference papers. Moreover, to ensure the accuracy of search results, namely collecting and analyzing articles highly focused on LAL, this study performed a more precise search, which may have resulted in fewer search results. As more data and documents are available for a more comprehensive understanding of developments in the field of LAL, future studies should consider expanding the scope of the database search and target language to create a richer visualization of LAL by rigorously screening relevant data.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: All data generated or analyzed in this review are included in this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Alderson, J.C.; Brunfaut, T.; Harding, L. Bridging Assessment and Learning: A View from Second and Foreign Language Assessment. *Assess. Educ.* **2017**, *24*, 379–387. [CrossRef]
2. Banitz, B. Language Assessment in Mexico: Exploring University Language Teachers' Backgrounds, Practices, and Opinions. *System* **2022**, *110*, 1–12. [CrossRef]
3. Jasinska, K.; Akpe, Y.H.; Debayeseri, B.A.; Zinszer, B.; Agui-Kouadio, R.Y.; Mulford, K.; Curran, E.; Ball, M.C.; Tanoh, F. Evaluating Bilingual Children's Native Language Abilities in Cote d'Ivoire: Introducing the Ivorian Children's Language Assessment Toolkit for Attie, Abidji, and Baoule. *Appl. Lingusit.* **2022**, *28*, 1116–1142. [CrossRef]
4. Butler, Y.G.; Peng, X.L.; Lee, J.Y. Young Learners' Voices: Towards a Learner-Centered Approach to Understanding Language Assessment Literacy. *Lang. Test.* **2021**, *38*, 429–455. [CrossRef]
5. Jeong, H. Defining Assessment Literacy: Is It Different for Language Testers and Non-Language Testers? *Lang. Test.* **2013**, *30*, 345–362. [CrossRef]
6. Xu, Y.; Brown, G.T. Teacher Assessment Literacy in Practice: A Reconceptualization. *Teach. Teach. Educ.* **2016**, *58*, 149–162. [CrossRef]
7. Clapham, C. Assessment and Testing. *ARAL* **2000**, *20*, 147–161. [CrossRef]
8. Mansouri, B.; Molana, K.; Nazari, M. The Interconnection between Second Language Teachers' Language Assessment Literacy and Professional Agency: The Mediating Role of Institutional Policies. *System* **2021**, *103*, 102674. [CrossRef]
9. Crusan, D.; Plakans, L.; Gebril, A. Writing Assessment Literacy: Surveying Second Language Teachers' Knowledge, Beliefs, and Practices. *Assess. Writ.* **2016**, *28*, 43–56. [CrossRef]
10. Leenknecht, M.; Wijnia, L.; Köhlen, M.; Fryer, L.; Rikers, R.; Loyens, S. Formative Assessment as Practice: The Role of Students' Motivation. *Assess. Eval. High. Educ.* **2021**, *46*, 236–255. [CrossRef]

11. Chan, K.T. Embedding Formative Assessment in Blended Learning Environment: The Case of Secondary Chinese Language Teaching in Singapore. *Educ. Sci.* **2021**, *11*, 360. [CrossRef]

12. Xu, Y.; Brown, G.T.L. University English Teacher Assessment Literacy: A Survey-Test Report from China. *Pap. Lang. Test. Assess.* **2017**, *6*, 133–158. Available online: http://hdl.handle.net/2292/34245 (accessed on 8 February 2023).

13. Inbar-Lourie, O. Evaluating Language Assessments. *Lang. Assess. Q.* **2021**, *18*, 199–203. [CrossRef]

14. Barber, M. The Virtue of Accountability: System Redesign, Inspection, and Incentives in the Era of Informed Professionalism. *J. Educ.* **2005**, *185*, 7–38. [CrossRef]

15. Stiggins, R.J. Assessment Literacy. *Phi Delta Kappan* **1991**, *72*, 534–539.

16. Davison, C.; Leung, C. Current Issues in English Language Teacher-Based Assessment. *TESOL Q.* **2009**, *43*, 393–415. [CrossRef]

17. Brindley, G. Language Assessment and Professional Development. In *Experimenting with Uncertainty: Essays in Honour of Alan Davies*; Elder, C., Brown, A., Hill, K., Iwashita, N., Lumley, T., McNamara, T., O'Loughlin, K., Eds.; Cambridge University Press: Cambridge, UK, 2001; pp. 126–136. ISBN 978-0-5217-7576-2.

18. Davies, A. Textbook Trends in Teaching Language Testing. *Lang. Test.* **2008**, *25*, 327–347. [CrossRef]

19. Inbar-Lourie, O. Constructing a Language Assessment Knowledge Base: A Focus on Language Assessment Courses. *Lang. Test.* **2008**, *25*, 385–402. [CrossRef]

20. Taylor, L. Communicating the Theory, Practice and Principles of Language Testing to Test Stakeholders: Some Reflections. *Lang. Test.* **2013**, *30*, 403–412. [CrossRef]

21. Baker, B.A.; Riches, C. The Development of EFL Examinations in Haiti: Collaboration and Language Assessment Literacy Development. *Lang. Test.* **2018**, *35*, 557–581. [CrossRef]

22. Kremmel, B.; Harding, L. Towards a Comprehensive, Empirical Model of Language Assessment Literacy across Stakeholder Groups: Developing the Language Assessment Literacy Survey. *Lang. Assess. Q.* **2020**, *17*, 100–120. [CrossRef]

23. Yan, X.; Zhang, C.; Fan, J.J. "Assessment Knowledge is Important, but . . . ": How Contextual and Experiential Factors Mediate Assessment Practice and Training Needs of Language Teachers. *System* **2018**, *74*, 158–168. [CrossRef]

24. Vogt, K.; Tsagari, D.; Csépes, I.; Green, A.; Sifakis, N. Linking Learners' Perspectives on Language Assessment Practices to Teachers' Assessment Literacy Enhancement (TALE): Insights from Four European Countries. *Lang. Assess. Q.* **2020**, *17*, 410–433. [CrossRef]

25. Scarino, A. Language Assessment Literacy as Self-Awareness: Understanding the Role of Interpretation in Assessment and in Teacher Learning. *Lang. Test.* **2013**, *30*, 309–327. [CrossRef]

26. Berry, V.; Sheehan, S.; Munro, S. What does Language Assessment Literacy Mean to Teachers? *ELT J.* **2019**, *73*, 113–123. [CrossRef]

27. Lam, R. Language Assessment Training in Hong Kong: Implications for Language Assessment Literacy. *Lang. Test.* **2015**, *32*, 169–197. [CrossRef]

28. Vogt, K.; Tsagari, D. Assessment Literacy of Foreign Language Teachers: Findings of a European Study. *Lang. Assess. Q.* **2014**, *11*, 374–402. [CrossRef]

29. Lam, R. Teacher Assessment Literacy: Surveying Knowledge, Conceptions and Practices of Classroom-Based Writing Assessment in Hong Kong. *System* **2019**, *81*, 78–89. [CrossRef]

30. Gan, L.; Lam, R. A Review on Language Assessment Literacy: Trends, Foci and Contributions. *Lang. Assess. Q.* **2022**, *19*, 503–525. [CrossRef]

31. Coombs, A.; DeLuca, C. Mapping the Constellation of Assessment Discourses: A Scoping Review Study on Assessment Competence, Literacy, Capability, and Identity. *Educ. Assess. Eval. Acc.* **2022**, *34*, 279–301. [CrossRef]

32. DeLuca, C.; LaPointe-McEwan, D.; Luhanga, U. Teacher Assessment Literacy: A Review of International Standards and Measures. *Educ. Assess. Eval. Acc.* **2016**, *28*, 251–272. [CrossRef]

33. Gotch, C.M.; French, B.F. A Systematic Review of Assessment Literacy Measures. *Educ. Meas.* **2014**, *33*, 14–18. [CrossRef]

34. Dere, İ.; Ateş, Y. Studies on Literacy Skills in Social Studies Education: A Systematic Literature Review (1996–2020). *Scand. J. Educ. Res.* **2021**, *1*, 1–17. [CrossRef]

35. Harding, L.; Kremmel, B. Teacher Assessment Literacy and Professional Development. In *Handbook of Second Language Assessment*; Tsagari, D., Banerjee, J., Eds.; De Gruyter Mouton: Boston, MA, USA, 2016; pp. 413–428. ISBN 978-1-614-51542-5.

36. Weng, F.; Shen, B. Language Assessment Literacy of Teachers. *Front. Psychol.* **2022**, *13*, 1–7. [CrossRef]

37. Chen, C. *How to Use CiteSpace*; Lean Publishing: Düsseldorf, Germany, 2022; Available online: https://leanpub.com/howtousecitespace (accessed on 1 December 2022).

38. Roemer, R.C.; Borchardt, R. *Meaningful Metrics: A 21st Century Librarian's Guide to Bibliometrics, Altmetrics, and Research Impact*; Association of College and Research Libraries: Chicago, IL, USA, 2015; ISBN 978-0-8389-8755-1.

39. Chen, C. Searching for Intellectual Turning Points: Progressive Knowledge Domain Visualization. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5303–5310. [CrossRef]

40. Demir, Y.; Kartal, G. Mapping Research on L2 Pronunciation: A Bibliometric Analysis. *Stud. Second Lang. Acquis.* **2022**, *44*, 1211–1239. [CrossRef]

41. Guo, X. A Bibliometric Analysis of Child Language During 1900–2021. *Front. Psychol.* **2022**, *13*, 1–17. [CrossRef]

42. Taylor, L. Developing assessment literacy. *Annu. Rev. Appl. Linguist.* **2009**, *29*, 21–36. [CrossRef]

43. Xu, Y.; Liu, Y. Teacher Assessment Knowledge and Practice: A Narrative Inquiry of a Chinese College EFL Teacher's Experience. *TESOL Q.* **2009**, *43*, 492–513. [CrossRef]

44. Giraldo, F. Language Assessment Literacy: Implications for Language Teachers. *Profile Issues Teach. Prof. Dev.* **2018**, *20*, 179–195. [CrossRef]

45. Levi, T.; Inbar-Lourie, O. Assessment Literacy or Language Assessment Literacy: Learning from the Teachers. *Lang. Assess. Q.* **2020**, *17*, 168–182. [CrossRef]

46. Chen, C.; Leydesdorff, L. Patterns of Connections and Movements in Dual-Map Overlays: A New Method of Publication Portfolio Analysis. *J. Am. Soc. Inf. Sci. Technol.* **2014**, *65*, 334–351. [CrossRef]

47. Aksnes, D.W.; Langfeldt, L.; Wouters, P. Citations, Citation Indicators, and Research Quality: An Overview of Basic Concepts and Theories. *SAGE Open* **2019**, *9*, 1–17. [CrossRef]

48. Fulcher, G. Assessment Literacy for the Language Classroom. *Lang. Assess. Q.* **2012**, *9*, 113–132. [CrossRef]

49. Malone, M.E. The Essentials of Assessment Literacy: Contrasts between Testers and Users. *Lang. Test.* **2013**, *30*, 329–344. [CrossRef]

50. Hurtado-Marín, V.A.; Agudelo-Giraldo, J.D.; Robledo, S.; Restrepo-Parra, E. Analysis of Dynamic Networks Based on The Ising Model for The Case of Study of Co-Authorship of Scientific Articles. *Sci. Rep.* **2021**, *11*, 5721. [CrossRef]

51. Price, D. *Little Science, Big Science*; Columbia University Press: New York, NY, USA, 1963; ISBN 978-0-231-88575-1.

52. van Raan, A.F. Fatal Attraction: Conceptual and Methodological Problems in the Ranking of Universities by Bibliometric Methods. *Scientometrics* **2005**, *62*, 133–143. [CrossRef]

53. Liu, W.; Hu, G.; Tang, L.; Wang, Y. China's Global Growth in Social Science Research: Uncovering Evidence from Bibliometric Analyses of SSCI publications (1978–2013). *J. Informetr.* **2015**, *9*, 555–569. [CrossRef]

54. Sun, Y.; Lan, G. Research Trends in 'Trans-' Studies on Writing: A Bibliometric Analysis. *System* **2021**, *103*, 102640. [CrossRef]

55. van Nunen, K.; Li, J.; Reniers, G.; Ponnet, K. Bibliometric Analysis of Safety Culture Research. *Saf. Sci.* **2018**, *108*, 248–258. [CrossRef]

56. Coombe, C.; Vafadar, H.; Mohebbi, H. Language Assessment Literacy: What Do We Need to Learn, Unlearn, and Relearn? *Lang. Test. Asia* **2020**, *10*, 1–16. [CrossRef]

57. Wu, X.M.; Dixon, H.R.; Zhang, L.J. Sustainable Development of Students' Learning Capabilities: The Case of University Students' Attitudes towards Teachers, Peers, and Themselves as Oral Feedback Sources in Learning English. *Sustainability* **2021**, *13*, 5211. [CrossRef]

58. Wu, X.M.; Zhang, L.J.; Liu, Q. Using Assessment for Learning: Multi-Case Studies of Three Chinese University English as a Foreign Language Teachers Engaging Students in Learning and Assessment. *Front. Psychol.* **2021**, *12*, 1–16. [CrossRef]

59. Small, H. Co-Citation in The Scientific Literature: A New Measure of The Relationship between Two Documents. *JASIST* **1973**, *24*, 265–269. [CrossRef]

60. Chen, C. Science Mapping: A Systematic Review of the Literature. *J. Data Inf. Sci.* **2017**, *2*, 1–40. [CrossRef]

61. Tsagari, D.; Vogt, K. Assessment Literacy of Foreign Language Teachers around Europe: Research, Challenges and Future Prospects. *Pap. Lang. Test. Assess.* **2017**, *6*, 41–63.

62. Inbar-Lourie, O. Language Assessment Literacy. In *Language Testing and Assessment*; Shohamy, E., May, S., Eds.; Springer International Publishing: Cham, Switzerland, 2017; pp. 257–270. ISBN 978-3-319-02260-4.

63. Lantolf, J.P.; Poehner, M.E. Dynamic Assessment. In *Encyclopedia of Language and Education: Language Testing and Assessment*; Shohamy, E., Hornberger, N.H., Eds.; Springer Science and Business Media: New York, NY, USA, 2008; pp. 273–285. ISBN 978-0-387-32875-1.

64. Malone, M.E. Training in Language Assessment. In *Encyclopedia of Language and Education: Language Testing and Assessment*; Shohamy, E., Hornberger, N.H., Eds.; Springer Science and Business Media: New York, NY, USA, 2008; pp. 273–285. ISBN 978-0-387-32875-1.

65. Council of Europe. *Common European Framework of Reference for Languages: Learning, Teaching, Assessment–Companion Volume*; Council of Europe Publishing: Strasbourg, France, 2020; ISBN 978-92-871-8621-8.

66. Klinger, D.A.; McDivitt, P.R.; Howard, B.B.; Munoz, M.A.; Rogers, W.T.; Wylie, E.C. *The Classroom Assessment Standards for PreK-12 Teachers*; Kindle Direct Press: Washington, DC, USA, 2015; ASIN B00V6C9RVO.

67. Brown, J.D.; Bailey, K.M. Language Testing Courses: What are They on 2007? *Lang. Test.* **2008**, *25*, 349–383. [CrossRef]

68. Borg, S. English Language Teachers' Conceptions of Research. *Appl. Lingusit.* **2009**, *30*, 358–388. [CrossRef]

69. Chen, C.; Ibekwe-SanJuan, F.; Hou, J. The Structure and Dynamics of Co-Citation Clusters: A Multiple-Perspective Co-Citation Analysis. *J. Am. Soc. Inf. Sci. Technol.* **2010**, *61*, 1386–1409. [CrossRef]

70. Inbar-Lourie, O. Guest Editorial to the Special Issue on Language Assessment Literacy. *Lang. Test.* **2013**, *30*, 301–307. [CrossRef]