A Bibliometric Analysis and Future Research Directions on Language and Literacy

Hazlina Abdullah[1,2], Syadiyah Abdul Shukor[3], Humaira Raslie[4], Ezihaslinda Ngah[5], Noryani Neni Ahmad Jamain[2,6]

[1]Faculty of Major Language Studies, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia, [2]School of Education, University of New South Wales, Sydney NSW, 2052 Australia, [3]Faculty of Economics and Muamalat, Universiti Sains Islam Malaysia, Bandar Baru Nilai, 71800 Nilai, Negeri Sembilan, Malaysia, [4]Faculty of Language and Communication, Universiti Malaysia Sarawak, 94300 Kota Samarahan, Sarawak, Malaysia, [5]Centre for Modern Languages, Universiti Malaysia Pahang, 26600 Kuala Pahang, Pekan, Pahang, Malaysia, [6]Centre for Language Studies, Universiti Tun Hussein Onn Malaysia, 86400 Batu Pahat, Johor, Malaysia.
Corresponding Author Email: hazlina@usim.edu.my/ z3538500@ad.unsw.edu.au

Abstract
There has been considerable research conducted in the area of language and literacy focussing on themes such as reasons and factors, performance and achievements, ways, and methods for improvement, as well as risks and problems. However, an overview of a worldwide depiction of the research conducted in this area is hard to find. Thus, this research aimed to ascertain the global developments and trends in language and literacy over time by analysing three main levels of analysis: sources, authors, and documents. A bibliometric analysis was utilised to examine a final sample of 1,463 articles published in approximately the last 22 years (2000-2023). In carrying out the analysis, the bibliometrix R-Tool was used and the metadata from the SCOPUS database was retrieved. The results demonstrated an upward trend of publications with a high number of single-country publications and a fair number of multiple-country publications. The primary conclusion reveals that language and literacy can be considered a substantial field of research, and the most important themes identified are connected to five thematic clusters: (1) General development in language and literacy (2) Language and literacy in Early Childhood Education (ECE) (3) Language and literacy in English as a Second Language (ESL), (4) Issues in language and literacy and (5) Language and literacy for bilinguals. The field is also marked by constant growth; as indicated by the international collaborations, new trends and other related fields that conveyed great interest in the areas of language and literacy. These findings are crucial in informing researchers, educators and policy makers on the current needs related to language and literacy

implementations, as well as potential research endeavours that are critical to the development of the field.

Keywords: Language, Literacy, Bibliometric, Literature Review

Introduction
In this contemporary world, the conventional and long-established notion of literacy (or illiteracy) as the capability (or incapability) of reading and writing has now begun to take on a new practical aspect. This aspect is conceptualised within different domains as possessing knowledge, skills, and competence in particular fields for specific purposes. One is expected to comprehend the gist of certain areas and to connect with it accordingly. With the current growing number of fields and fast-paced advances, multiple literacies emerge as contemporary trends—computer literacy, media literacy, academic literacy—among others. Nevertheless, amidst all the new concepts of literacy, the authors are interested to explore the directions of "language and literacy" because they remain relevant and important for academic success (e.g., Glew et al., 2019; Brooks, 2019) and is a primary influence in the general education literature (e.g., Kim & Johnson, 2021; Lysenko et al., 2019). This study delved into the research directions and trends of language and literacy by looking at related studies in this area, examining the impact of publications and visualising some bibliometric networks (Ahmi, 2022; Ahmi & Nasir, 2019; Ahmi & Mohamad, 2019) of language and literacy.

From an empirical standpoint, the field of language and literacy is a well-established, and flourishing research area. However, it appears that a comprehensive bibliometric study that captures the growing state of the field is scarce. A search for bibliometrics analysis articles within the social sciences and education spheres generated minimal results; and those that identified with bibliometric analysis were limited to the following trends: foreign language teaching (e.g., Yilmaz et al., 2022), gamification (e.g., Marti-Parreno et al., 2016) and technology in Higher Education (e.g., Jiménez et al., 2019). According to Ahmi (2022), a bibliometric analysis is relevant across disciplines; be it in science, technology, or social sciences. As a powerful tool to examine scientific outputs, bibliometric analysis is able to zoom into a specific area of research and identify its structure, research trends, as well as themes (Liu, 2022). However, this technique of analysis is more popular in science, and technology-related research. As a field that is exponentially expanding both in terms of theories and practices, bibliometric investigations on subject matters such as third language acquisition, (Liu, 2022) and reading literacy (Lan & Yu, 2022) would capture the field's potentials; especially in providing empirical and practical resources (Lan & Yu, 2022; Liu, 2022). This study therefore addressed the current paucity in bibliometric analysis and future research directions on language and literacy, specifically aiming to:

1. quantify the research field and describe its main outputs and evolution,
2. identify the most productive contributors of the language and literacy field,
3. identify the most investigated themes of language and literacy.
4.

In light of the aims, this paper sought to answer the following research questions:

a) What is the current publication trend for studies in language and literacy?
b) Who are the most productive contributors in the field of language and literacy?
c) Which themes are the most popular among the scholars of language and literacy?

Literature Review

In the 21st century, the ability to read, write, and understand information is still very much relevant and influential to one's academic achievement (Mohd-Asraf et al., 2016) as well as employability (Department of Employment and Workplace Relations, 2022). The interpretation of what it means to be literate has also changed to include and reflect the literacy demands of today's world, as well as the society that we live in. From this standpoint, literacy is defined as the "understanding, evaluating, using, and engaging with written texts to participate in society, to achieve personal goals and ambitions and to develop knowledge and potential" (Kangan Institute, 2021).

The broaden understanding of literacy as the ability to read, write, speak, and listen in ways that permit one to communicate effectively to various types of audiences will allow one to make sense of the world. In addition, it is important to realise that reading and writing, when combined with speaking, listening, viewing and critical thinking will establish beneficial aspects of literacy in modern life. Undoubtedly, literacy is a vital element in ensuring one, especially a student, to have the best prospect in his or her studies and daily life.

The literature on language and literacy has experienced a consistent expansion in the last decades. However, even with the rising interest in language and literacy, not much has been achieved in terms of systematically categorising the articles that contributed to the area of language and literacy, especially in the more recent period of time. The existing studies paid attention to the following themes: the evolution of language and literacy within child language (Guo, 2022), bilingualism (Kayadibi, 2022), technology-based learning (Muhammad Fuad et. al., 2022), teacher's feedback (Xie, 2022), third language (L3) acquisition (Liu, 2022) and academic literacy (Meza, 2021). Yet the more universal bearing and current trends in the language and literacy field is lacking and needs to be highlighted. This research therefore attempted to gain a general and systematic overview of the extant research on language and literacy on the basis of three main levels of bibliometric analysis, namely the sources, authors, and documents.

Materials and Methods

This study used the Scopus database in bibliometric analysis and mapping which aimed to ascertain the global developments and trends in language and literacy over a specific period of time, encompassing the main analysis of sources, authors, and documents. The data were extracted on January 21, 2023, from the Scopus database, as depicted in Figure 1. For the purpose of this study, the keyword search in Scopus was set between year 2000 and 2023 and article titles that contained "language" AND "literacy" were included. Only English-language article publications were chosen, and all selected documents were checked to ensure there were no duplicates in the sample. This led to the extraction of a total of 1,463 records, or article journals. All records were included in the bibliometric study.
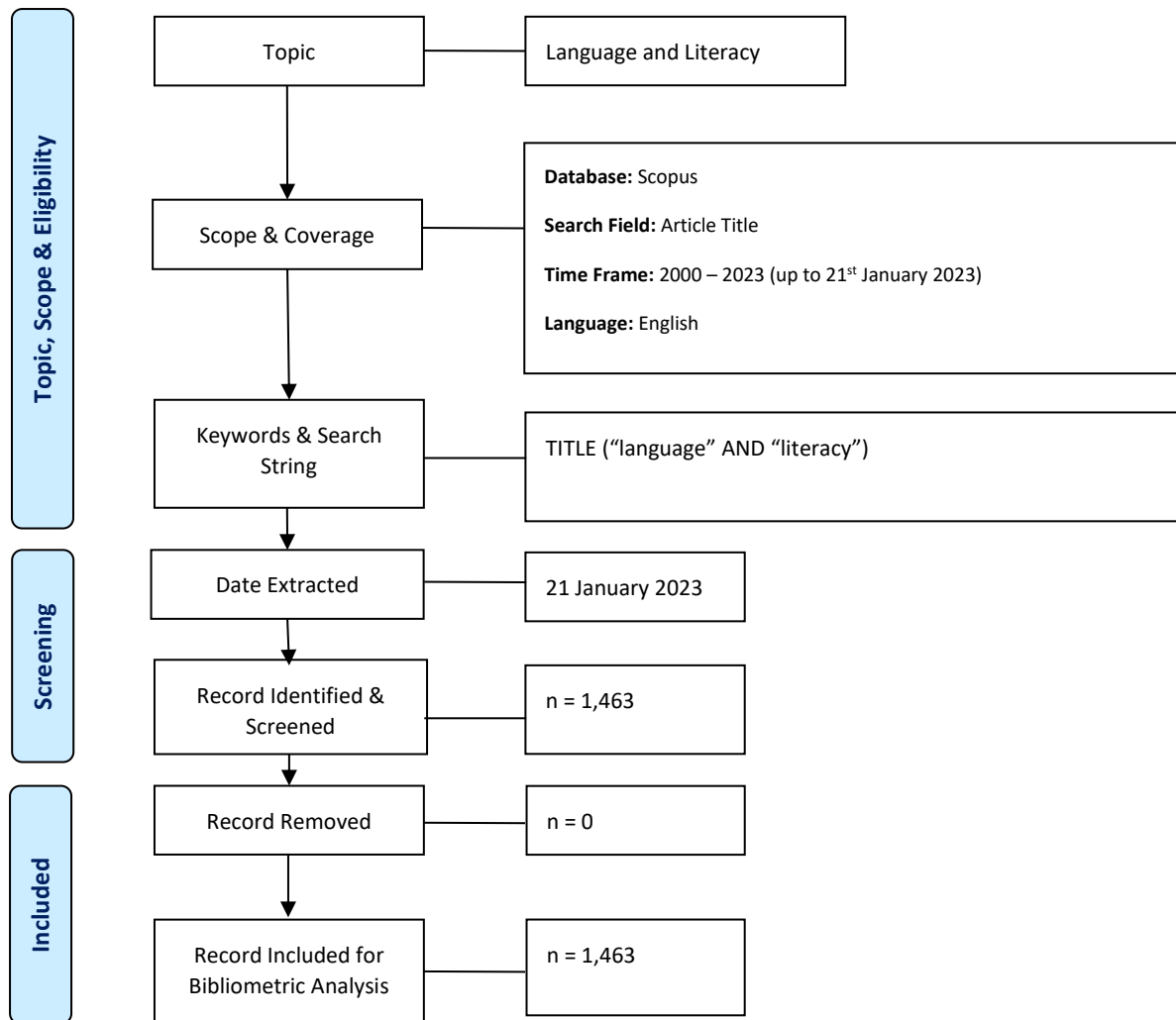
Figure 1. Flow diagram of the search strategy
Source: Zakaria et al(2020)

Results and Discussion

Main information

Table 1 provides an overview of the dataset run for this study. A total of 1,463 articles from 542 journals were identified from the Scopus database from 2000 to 2023. The average citations per document is 20.22 while average citations per year per document is 1.833. In terms of authors, 3,348 authors wrote these articles; 392 documents were single authored while the remaining 2,956 documents were multi-authored documents. The collaboration index was 2.82, which means on average, two to three authors have written an article. The next part of the paper will present the findings based on the research questions listed earlier.

Table 1

Main information regarding selected articles.

Description	Results
Main Information About Data	
Period	2000:2023
namSources (Journals)	543
Documents	1,463
Average citations per documents	20.22
Average citations per year per doc	1.833
Authors	
Authors	3,348
Author Appearances	4,187
Authors of single-authored documents	392
Authors of multi-authored documents	2,956
Authors Collaboration	
Single-authored documents	415
Documents per Author	0.437
Authors per Document	2.29
Co-Authors per Documents	2.86
Collaboration Index	2.82

Current publication trend for studies in language and literacy

In terms of publication trend for studies in language and literacy, Figure 1 shows the number of articles published according to year, beginning with 2000. It is found that within the period of 22 years (2000-2023), there is an increasing number of papers published in the area of language and literacy, recording manifold of increase throughout the years. This indicates that the area of language and literacy is still relevant and has increased in popularity in the recent years.
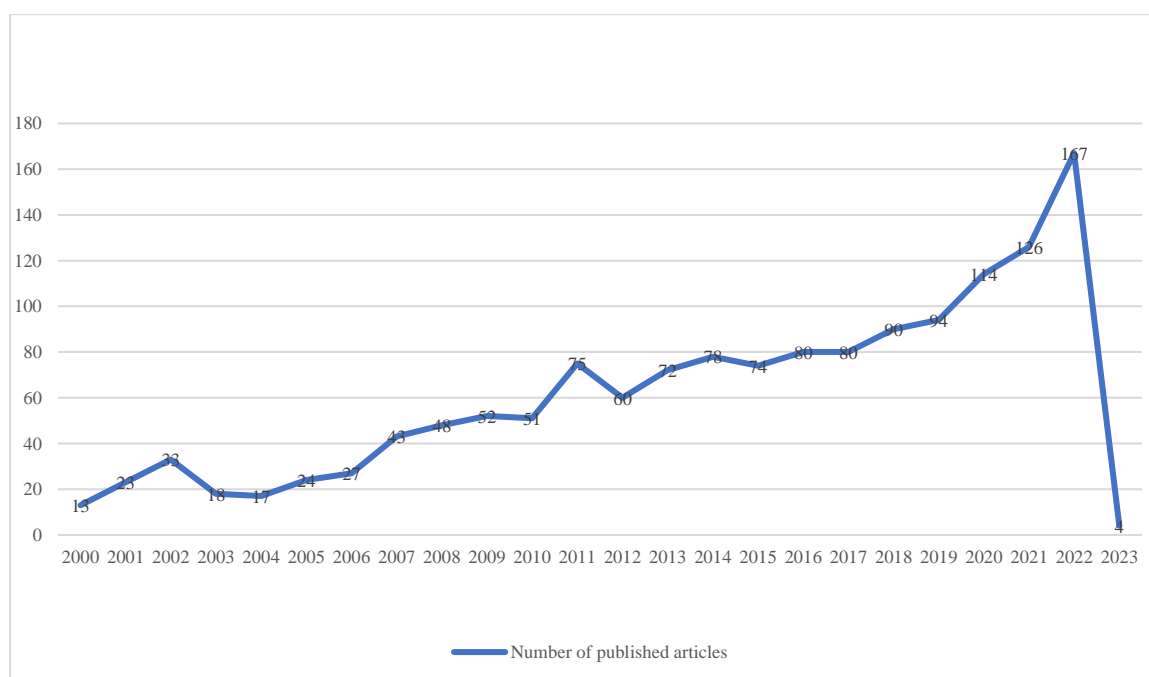

Figure 1. Annual publication trends

The most productive contributors in the field of language and literacy
To address the most productive contributors in language and literacy research, this study analysed the following data: a) most relevant sources, b) most relevant affiliations, c) most productive authors, and d) most cited papers.

Most relevant sources
"The most relevant sources" is a reference to "analysis that identifies each source's total number of publications" (Ahmi, 2022, p. 106). It demonstrates the productivity of the source title. For language and literacy, the *Early Childhood Research Quarterly* is the highest contributor of publications with 44 articles. The second highest is the *International Journal of Bilingual Education and Bilingualism* with 31 articles, followed by the *Journal of Early Childhood Literacy* with 24 articles. Next come the *Journal of Adolescent and Adult Literacy*, and the *Reading and Writing* journal. Each journal published 23 articles. Interestingly, the next three top journals in the field of language and literacy are *Early Childhood Education Journal* and *Early Education and Development* with 22 articles each, and *Early Child Development and Care* with 19 articles—all of which are related to early childhood education.

Table 2
Most relevant sources

Sources	Articles
Early Childhood Research Quarterly	44
International Journal of Bilingual Education and Bilingualism	31
Journal of Early Childhood Literacy	24
Journal of Adolescent and Adult Literacy	23
Reading and Writing	23
Early Childhood Education Journal	22
Early Education and Development	22
Early Child Development and Care	19
Linguistics and Education	18
Reading Teacher	16
Topics in Language Disorders	16
Bilingual Research Journal	15
Language Testing	15
Language Speech and Hearing Services in Schools	14
Reading Research Quarterly	14
Child Language Teaching and Therapy	13
International Journal of Language and Communication Disorders	13
TESOL Quarterly	13
Journal of Educational Psychology	12
L1 Educational Studies in Language and Literature	12

Most relevant affiliations
The findings have also shown the most relevant affiliations in the field of language and literacy, which refers to the productivity of publications based on the institutions that the authors affiliated themselves with at the time of publication. Based on Figure 2, the Ohio State University produced the highest number of publications in the field (64 publications),

followed by the University of California (57 publications), University of Virginia (35 publications) and Florida State University (34 publications). It is observed that in the top ten list, nine universities are from the United States and one institution is from Denmark (AARHUS University). The other universities in the list are from Australia (La Trobe University, University of Melbourne), the United Kingdom (University of London, University of Oxford) and Canada (University of Toronto).
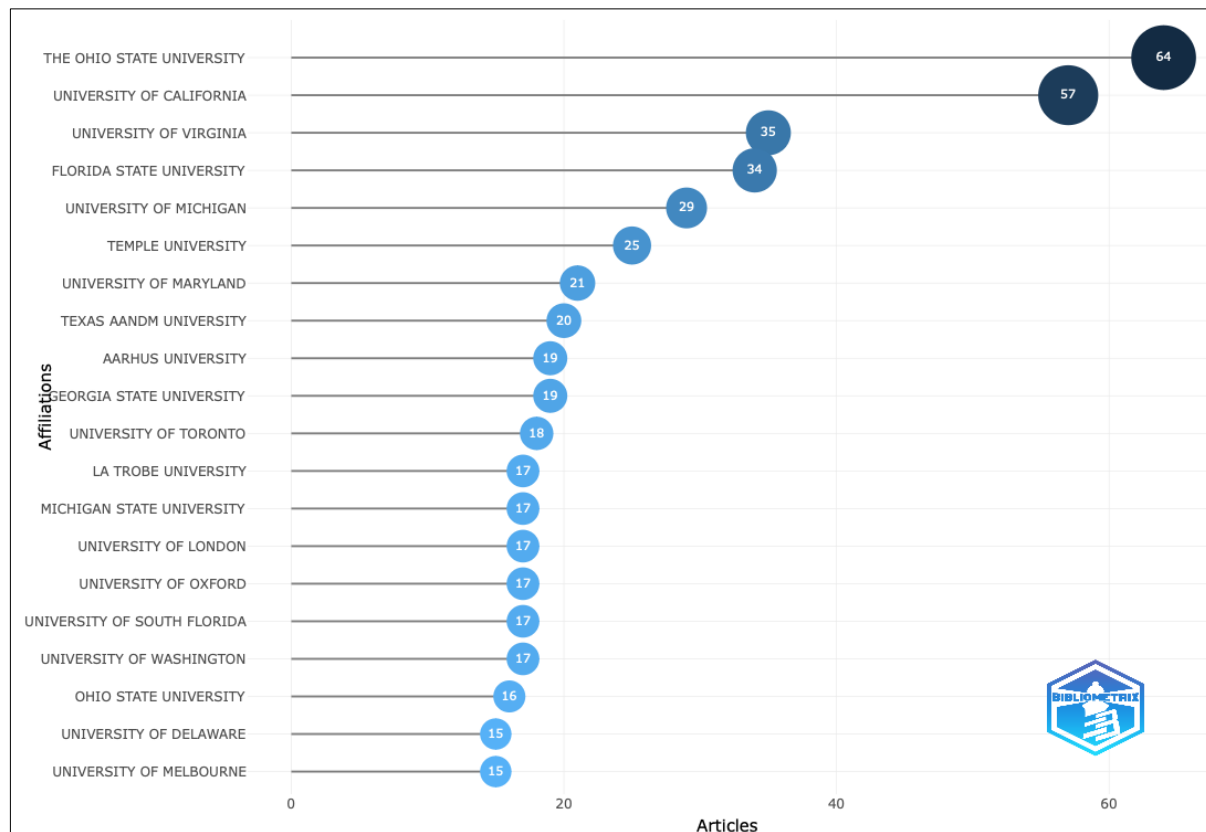

Figure 2. Most relevant affiliations

Most Productive Authors
As shown in Figure 3, Justice, L.M. from the Ohio State University is the most productive author in the field of language and literacy with 46 publications. The author's h-index of 96 and the high number of citations she has received for many of her publications (https://scholar.google.com/citations?user=92E5f5IAAAAJ&hl=en) reflects her widely regarded expertise in the area, of early childhood education (ECE). A correlation is established between an author's research productivity with relevance an affiliation; Justice, L.M. is affiliated with the Ohio State University at the time of this study, so as the other highly productive authors, namely Piasta, S.B., Lonigan, C.J, and Hindman, A.H. These authors contributed to the overall high relevance of the university.
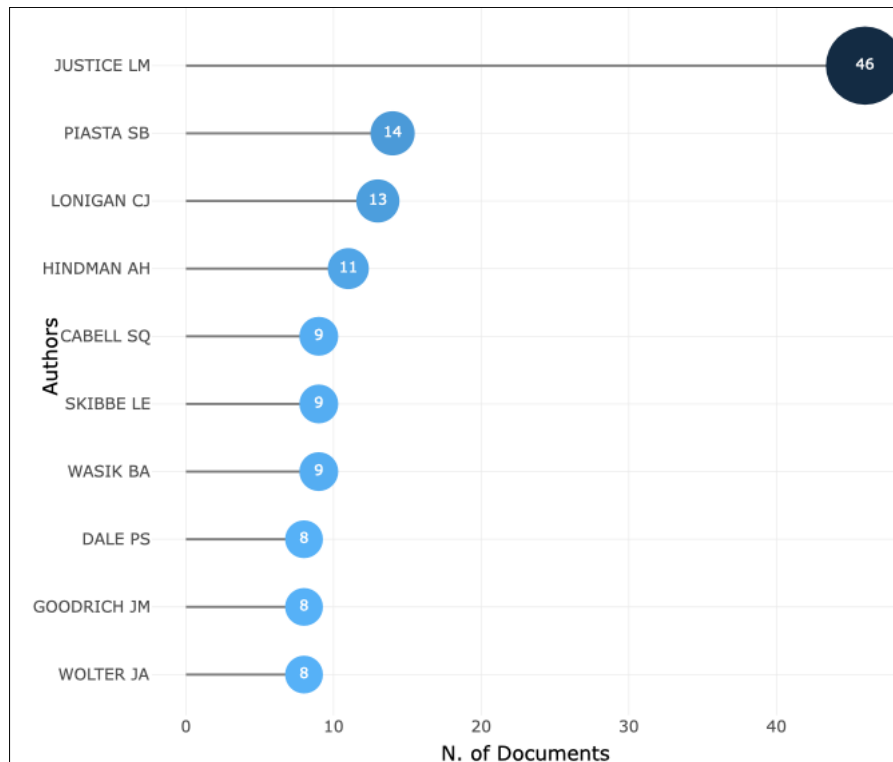
Figure 3. Most productive authors

Table 3 depicts the number of research articles published by writers from the same country (single-country publications [SCP]) versus publications by authors from other countries (multiple-country publications [MCP]). The former represents intra-country collaboration whereby all authors belong to the same country while the latter denotes inter-country or international collaboration in which the authors belong to different countries. Congruent to the earlier results on authors and affiliations, the United States produces the most publications with 493 SCP and 35 MCP. This is followed by the United Kingdom (71 SCP and 17 MCP), Canada (60 SCP and 14 MCP), followed by Australia (59 SCP and 9 MCP). Other corresponding authors' countries are detailed in Table 3 below.

Table 3
Corresponding author's countries

Country	Articles	Freq	SCP	MCP	MCPRatio
USA	528	0.48044	493	35	0.0663
United Kingdom	88	0.08007	71	17	0.1932
Canada	74	0.06733	60	14	0.1892
Australia	68	0.06187	59	9	0.1324
China	28	0.02548	17	11	0.3929
South Africa	22	0.02002	21	1	0.0455
Germany	19	0.01729	10	9	0.4737
Hong Kong	18	0.01638	16	2	0.1111
New Zealand	18	0.01638	13	5	0.2778
Sweden	15	0.01365	12	3	0.2
Norway	13	0.01183	9	4	0.3077
Netherlands	12	0.01092	8	4	0.3333
Finland	11	0.01001	9	2	0.1818
Israel	11	0.01001	10	1	0.0909
Spain	11	0.01001	8	3	0.2727
Denmark	10	0.0091	3	7	0.7
Greece	10	0.0091	9	1	0.1
Korea	9	0.00819	5	4	0.4444
Belgium	8	0.00728	5	3	0.375
Brazil	8	0.00728	5	3	0.375

Most cited papers
Hyland's (2007) study on "Genre pedagogy: Language, literacy and L2 writing instruction" obtained the highest number of citations. In this empirical investigation he introduced the principles of genre-based language instruction and drafted broad classroom models, specifically on English for Specific Purposes (ESP) and Systemic Functional Linguistics (SFL) approaches. Hyland (2007) also explored what it means to implement genre teaching in more practical terms, setting out some keyways in which teachers can plan, sequence, support, and assess learning. The second highest citation is Weiland and Yoshikawa's (2013) article entitled "Impacts of a Prekindergarten Programme on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills". Weiland and Yoshikawa (2013) examined the impact of a prekindergarten programme that implemented a coaching system and consistent literacy, language, and mathematics curricula on these and other non-targeted, essential components of school readiness, such as executive functioning.

Like Weiland and Yoshikawa, the third and fourth most cited publications are oriented on the ECE. Justice's work on, "Quality of language and literacy instruction in preschool classrooms serving at-risk pupils" characterised the quality of language and literacy instruction in public pre-schools that serve at-risk children. It is found that the quality of language and literacy instruction in classrooms is at the low level, although some teachers convey high quality instruction. The paper entitled "The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children" by Dickinson (2003) discussed the relationship between oral language and literacy skills, which included 533 preschool children on their

receptive vocabulary, phonological awareness, and print knowledge. Interestingly, in the top ten list of most cited papers, only a small number deals with literacy and science as highlighted in two studies; the first one by Yore et al.'s (2003) "Examining the literacy component of science literacy: 25 years of language arts and science research" and Moje et al.'s (2001) "Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science". The other highly cited papers are listed in Table 4 in which each article received above 250 citations.

Table 4
Most cited papers

Author (Year)	Title of Paper (DOI)	Journal	Total Citations	TC Per Year	Normalized TC
Hyland (2007)	Genre Pedagogy: Language, Literacy and L2 Writing Instruction 10.1016/J.Jslw.2007.07.005	*Journal of Second Language Writing*	441	25.9412	7.2963
Weiland & Yoshikawa (2013)	Impacts of a Prekindergarten Programme on Children's Mathematics, Language, Literacy, Executive Function, and Emotional Skills 10.1111/Cdev.12099	*Child Development*	394	35.8182	13.2375
Justice (2008)	Quality of language and literacy instruction in preschool classrooms serving at-risk pupils 10.1016/J.Ecresq.2007.09.004	*Early Childhood Research Quarterly*	377	23.5625	10.394
Dickinson (2003)	The comprehensive language approach to early literacy: The interrelationships among vocabulary, phonological sensitivity, and print knowledge among preschool-aged children. 10.1037/0022-0663.95.3.465	*Journal of Educational Psychology*	348	16.5714	4.4807
Yore et al (2003).	Examining the literacy component of science literacy: 25 years of language arts and science research	*International Journal of Science Education*	315	15	4.0558

Author (Year)	Title of Paper (DOI)	Journal	Total Citations	TC Per Year	Norma lized TC
	10.1080/09500690305018				
Briscoe (2001)	Phonological Processing, Language, and Literacy: A Comparison of Children with Mild-to-moderate Sensorineural Hearing Loss and Those with Specific Language Impairment 10.1017/S0021963001007041	*The Journal of Child Psychology and Psychiatry and Allied Disciplines*	291	12.6522	4.2441
Neuman (2009)	The Impact of Professional Development and Coaching on Early Language and Literacy Instructional Practices 10.3102/0002831208328088	*American Educational Research Journal*	280	18.6667	7.3498
Roberts (2005)	The Role of Home Literacy Practices in Preschool Children's Language and Emergent Literacy Skills 10.1044/1092-4388(2005/024)	*Journal of Speech, Language, and Hearing Research*	263	13.8421	7.0446
Oller (2007)	Profile effects in early bilingual language and literacy 10.1017/S0142716407070117	*Applied Psycholinguistics*	260	15.2941	4.3017
Moje et al (2001)	"Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science, 10.1002/Tea.1014	*Journal of Research in Science Teaching*	254	11.0435	3.7045

Most popular themes
To determine most popular themes for studies in language and literacy, this study analysed the following data: a) most frequent keywords, b) trend topics, c) co-occurrence network, and d) thematic map.

Most frequent keywords
It is found that the top ten most frequent words in the field of language and literacy are "literacy" (242), "language" (108), followed by "early literacy" (49), "emergent literacy" (49), "English language learners" (46), "reading" (46), "language development" (37), "preschool" (37),"health literacy" (33), and "oral language" (33), as indicated in Figure 4. The appearance

of many words related to ECE also corresponds with the most productive journals reported earlier.
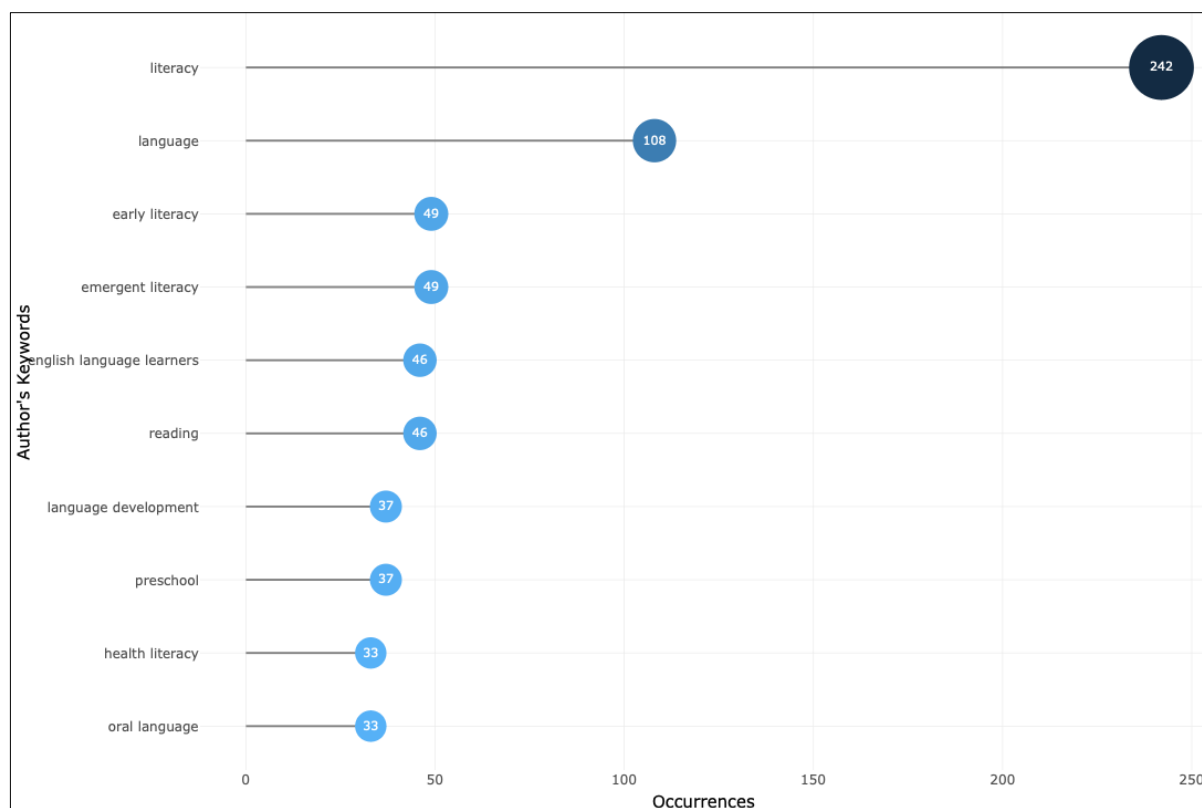

Figure 4. Most frequent keywords

Trend topics

The article proceeds with the trend topics analysis which looks at the evolution of terms based on "the start year of the terms' first appearance, the last year they were shown, and the year's median" (Ahmi, 2022, p. 165). Figure 5 displays the trend topics in a graph form containing the author's keywords from 2000-2023 in the field of language and literacy. The topmost five topics since 2020 are "translanguaging", "socioeconomic status" and "shared book reading", "higher education" and "COVID-19". It is anticipated that COVID-19 and higher education topics have gained popularity of late due to the global pandemic that hit the world. It is also interesting to witness the topic of translanguaging gaining significance. Translanguaging—the use of different languages together—is a powerful tool. It is "the act performed by bilinguals of accessing different linguistic features or various modes of what are described as autonomous languages, in order to maximise communicative potential" (Garcia, 2009, p. 140). In other words, translanguaging is about communication, and not just about language itself.
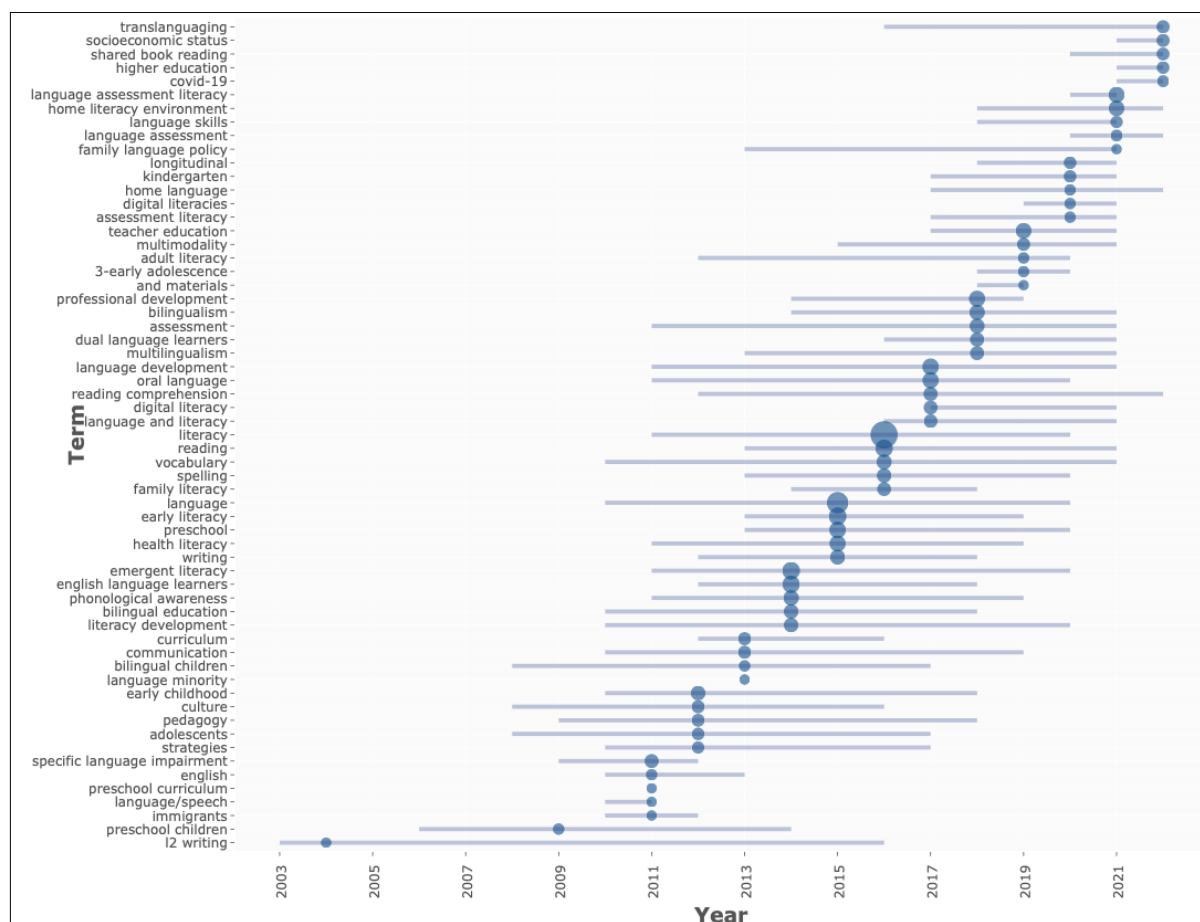
Figure 5. Trend topics

Co-occurrence network

Co-occurrence network is when "two or more things occur together or simultaneously" (Ahmi, 2022) which uncovers the conceptual structure and associations between concepts (Aria & Cuccurullo, 2017). This section explains the co-occurrence network in the language and literacy field. Based on Figure 6, there is a thick line between the nodes of "literacy" and "language" with "reading", which indicates the stronger relationship between these two keywords. The bubble size of the following keywords namely "language", "literacy" and "reading" are also larger. This shows that these words are used more by the academics. The findings have also identified five thematic clusters in terms of language and literacy research. The themes were coded into and differentiated by the distinct colours (refer to Figure 4). Table 5 contains the details of the categorisation:

Red : Cluster 1 – General development in language and literacy
Blue : Cluster 2 – Language and literacy in Early Childhood Education (ECE)
Green : Cluster 3 – Language and literacy in English as a Second Language (ESL)
Purple : Cluster 4 – Issues in language and literacy
Orange: Cluster 5 – Language and literacy for bilinguals

Cluster 1 is composed of six keywords: "language development", "home literacy environment", "literacy development", "dual language learners", "bilingual" and "head start"; signalling the broad-spectrum of topics in language and literacy. Cluster 2 involves thirteen keywords, which are oriented on the ECE part of language and literacy. ECE is considered as one of the greatest investments a country can make to ensure the continuity

of progress because it prepares the children for learning and future growth. Cluster 3 contains ten keywords (e.g., "English language learners", "language learners", "English learners") related to ESL. In 2023, English is still the most widely spoken language in the world (Statistics & Data, 2023), which concurrently explains the emergence of this cluster. Cluster 4 is the largest cluster with 16 keywords such as "language policy", and "academic literacy" reflects the various issues or subjects in the area of language and literacy. Lastly, Cluster 5, with five keywords ("bilingualism", "bilingual education", "spelling", "reading comprehension", "biliteracy") highlights the research of language and literacy in bilingualism. As the world becomes more borderless, many people develop bilingualism to meaningfully engage with members of different linguistic groups.
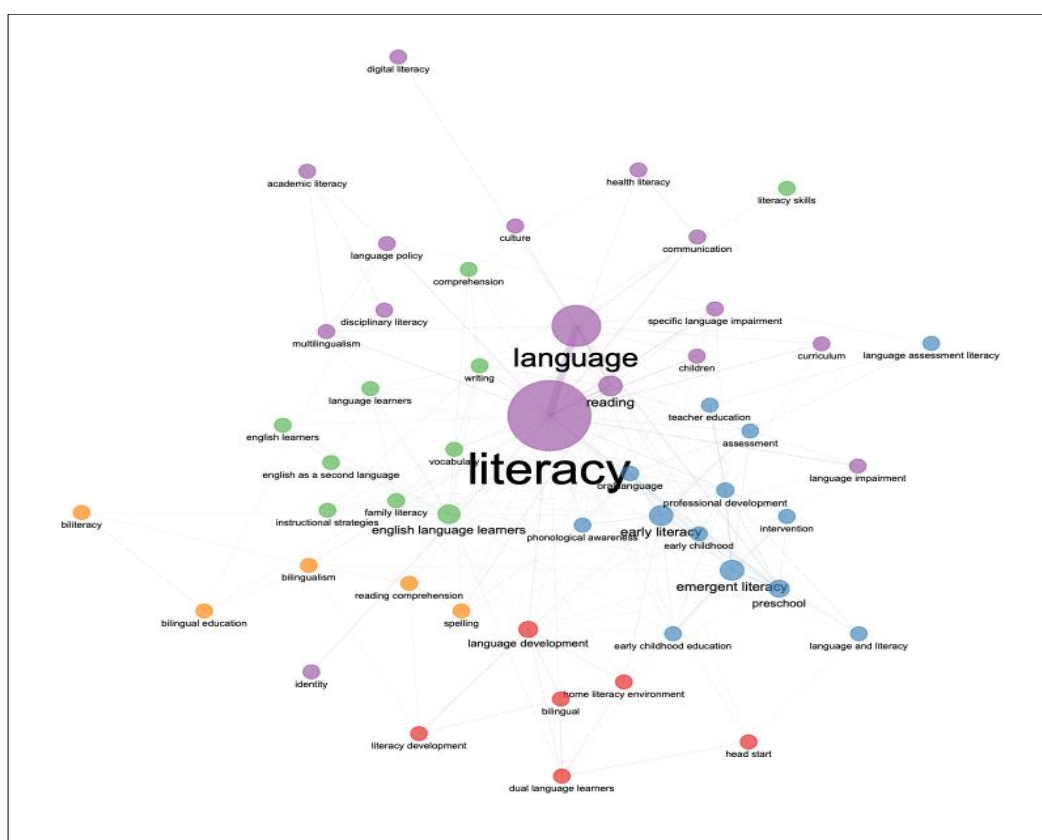

Figure 6. Co-occurrence network

Table 5 presents the details about the nodes or cluster where they belong, together with their betweenness centrality (showing which nodes are "bridges" between nodes in a network), closeness centrality (measuring the closeness to other nodes) and page rank (measuring the nodes whose influence extends beyond their direct connections into the wider network) (Disney, 2020 in Ahmi, 2022, p. 181).

Table 5
Details of the 5 clusters

Node	Cluster	Betweenness	Closeness	PageRank
language development	1	73.2559354	0.011235955	0.033069075
home literacy environment	1	2.380906627	0.009174312	0.013892484
literacy development	1	2.809400574	0.008064516	0.010587893
dual language learners	1	11.91539349	0.01010101	0.016435205
bilingual	1	1.060198514	0.008264463	0.010409907
head start	1	3.303696679	0.009433962	0.008887535
early literacy	2	30.08465444	0.010869565	0.032834558
emergent literacy	2	21.37421792	0.011235955	0.035385209
preschool	2	34.23178941	0.010752688	0.03492428
oral language	2	21.22267619	0.011111111	0.028181117
professional development	2	69.74083449	0.011494253	0.03336944
language assessment literacy	2	4.41224382	0.0078125	0.007615526
teacher education	2	6.530749221	0.008264463	0.011747722
phonological awareness	2	3.961582698	0.010204082	0.019549069
assessment	2	27.95043808	0.010416667	0.017319476
early childhood	2	10.58058456	0.010638298	0.019123158
intervention	2	3.060533565	0.010309278	0.015228808
early childhood education	2	0.936018848	0.010526316	0.014841186
language and literacy	2	4.498193385	0.008196721	0.009747182
English language learners	3	86.18892689	0.011494253	0.034085371
vocabulary	3	31.91758026	0.011235955	0.029068743
writing	3	67.24184647	0.010869565	0.028135045
family literacy	3	5.204701315	0.009433962	0.013696307
language learners	3	12.23751829	0.009090909	0.016516717
comprehension	3	1.576596652	0.009615385	0.012063915
English learners	3	10.89779587	0.009615385	0.012701502
literacy skills	3	0	0.007142857	0.004195739
English as a second language	3	3.980524562	0.008064516	0.01247585
instructional strategies	3	4.175690406	0.009009009	0.010971147
literacy	4	547.3509894	0.014925373	0.137400132
language	4	136.7349066	0.011904762	0.083820561
reading	4	18.21925098	0.010989011	0.039199548
health literacy	4	0.285714286	0.008849558	0.008574524
multilingualism	4	53.73822755	0.01	0.013018267
specific language impairment	4	0.874494912	0.009803922	0.014139961
digital literacy	4	0	0.007575758	0.004003486
language impairment	4	0.382286814	0.009174312	0.010306875

Node	Cluster	Betweenness	Closeness	PageRank
language policy	4	49.9995489	0.00952381	0.014541348
academic literacy	4	48	0.007092199	0.012057892
children	4	2.164776546	0.009615385	0.011618709
disciplinary literacy	4	0	0.005291005	0.006416403
communication	4	0.355959054	0.009090909	0.011320448
culture	4	0	0.007575758	0.005508714
curriculum	4	0	0.009174312	0.008198711
identity	4	0	0.008695652	0.005572469
bilingualism	5	13.54926511	0.009345794	0.012204676
bilingual education	5	12.94600444	0.009708738	0.011946036
spelling	5	0.889018541	0.00952381	0.017458507
reading comprehension	5	12.2527359	0.010204082	0.016683108
biliteracy	5	0.525592357	0.007874016	0.00895046

Thematic Map

The next section will present the thematic map which allows an assessment of topics based on the quadrants. The upper right quadrant shows the motor themes that are characterised by high centrality and density. The motor theme in this quadrant is "early literacy", and it is connected to other concepts such as "English language learners" and "oral language". They are considered as the most developed theme in the literature, and the key concerns within the language and literacy field.

The upper left quadrant represents high density themes but with insignificant external links-which render them as themes with low centrality and have limited importance for the field. Figure 7 shows that there are no actual themes located in the upper left quadrant but "reading" and other related concepts like "phonological awareness" and "vocabulary" are situated in the middle of both the upper quadrants, which signals that although they are quite specialised, they are also gaining prominence in the field of language and literacy.

The lower left quadrant, which conveys themes that are marginal and inadequately developed, represents either emerging or disappearing themes. In Figure 7, it can be seen that "language assessment literacy", "bilingual education" and "digital literacy" are among the emerging but weakly developed themes. Apart from that, there are also themes that are becoming increasingly unpopular, such as "emergent literacy", "language development" and "home literacy environment".

Finally, the lower right quadrant displays transversal, general, and basic themes which are vital, but not developed. From Figure 7, there are no themes found in this quadrant, but "literacy", "language" and "preschools" are the themes that belong to the shared quadrants of upper right and lower right; thus, indicating that the themes are gaining importance in the language and literacy field.
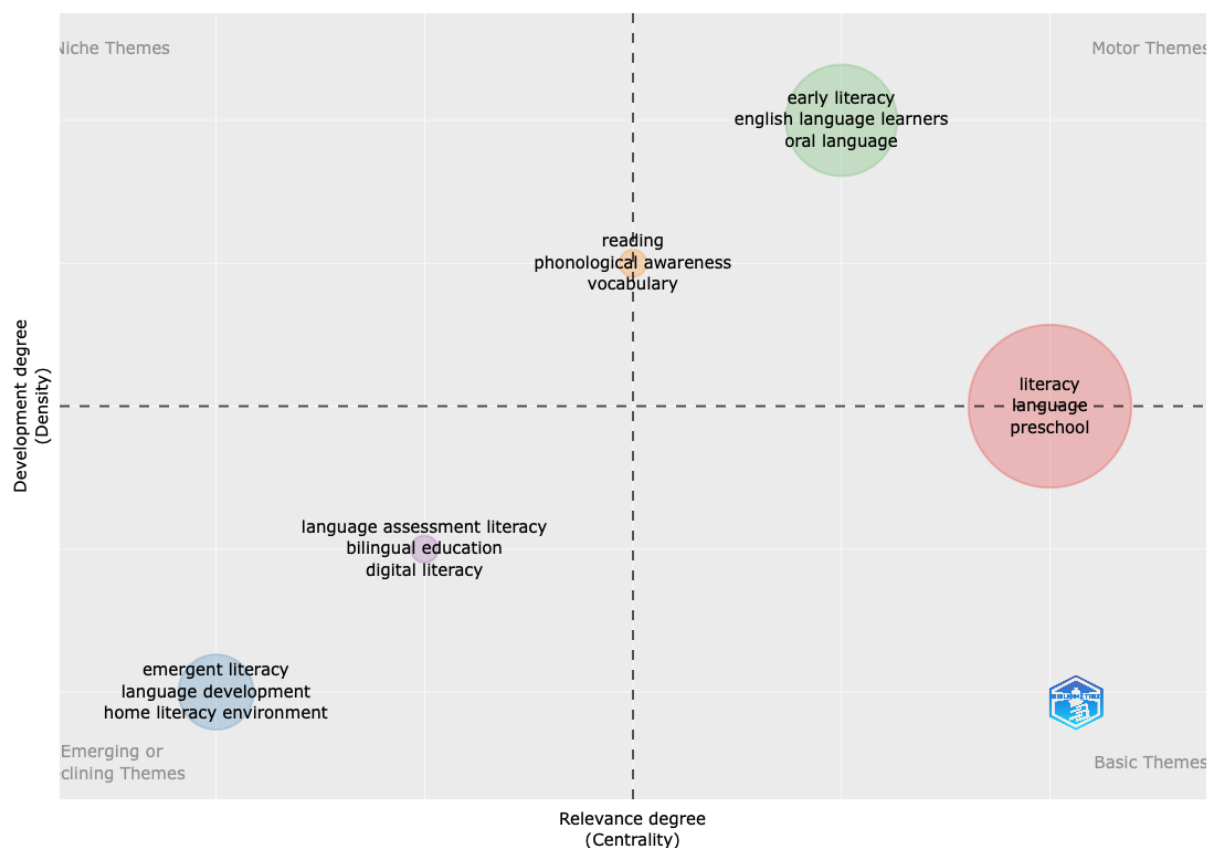
Figure 7. Thematic map

Conclusion

In the present study, the authors set out to examine the literature on language and literacy acquisition over the period of two decades (2000–2023) through analysing the changes in influential sources, authors, institutions, and countries. This bibliometric analysis offers an overview of the state-of-art for language and literacy. It has presented a review of the academic literature dealing with the evolution and progress of language and literacy in the Scopus databases. It is found that the area of language and literacy has undergone positive expansion as evident from the increase in the number of publications in this field. The most popular and widespread issues discussed within this field are related to ECE, English language and ESL which; can be observed from the kinds of most popular journals, most cited authors, most occurred keywords and the co-occurrence network or clusters.

Consequently, the thematic map has also shown that ECE and ESL are still predominantly important scopes within the field; shown through phrases like "early literacy", "early language learners", "literacy", "language" and "preschool". Nevertheless, terms like "emergent literacy", "language development" and "home literacy environment" are found to be disappearing despite its close relation to ECE and ESL. This may be due to the "over-popularity" of these terms during the time when they were in trend, which was in the early years of the millennium.   The more prevalent keywords nowadays include COVID-19 (refer to Figure 5), in line with the COVID-19 pandemic that struck the whole world and left a significant impact on all aspects of life including education as well as research and publication.

Inferring from the findings of this research, a future vision of language and literacy will still be oriented to the five clusters identified. It is foreseen that language and literacy will also expand into other areas like digital literacy. Digital literacy is an area of growing importance;

particularly due to the emphasis placed on enhancing international communication skills to facilitate 21st century learners' success in global interaction and communication.

Finally, the limitations of this paper should also be addressed. Firstly, the data were only extracted from the SCOPUS database. Even though SCOPUS, which was introduced by Elsevier in 2004 is the largest abstract and citation database of peer-reviewed research literature, it is much younger than Web of Science (WoS) (Murudka, 2022). Therefore, future study can be expanded by extracting data from other databases, especially WoS. The next limitation is related to the types of documents looked at, which covers solely journal articles. In future study, the inclusion of other types of documents such as books, and indexed conference proceedings would provide complementary insights on the subject matter. Given that the area of language and literacy seems perennial, further bibliometric analysis could also be extended by running the three fields plot, co-citations, and collaborations analyses to provide a more holistic picture.

Acknowledgements

References

Ahmi, A. (2022). Bibliometric analysis using R for non-coders: A practical handbook in conducting bibliometric analysis studies using Biblioshiny for Bibliometrix R package. Malaysia.

Ahmi, A., & Mohamad, R. (2019). Bibliometric analysis of global scientific literature on web accessibility. International Journal of Recent Technology and Engineering, 7(6), 250–258.

Ahmi, A., & Nasir, M. H. (2019). Examining the trend of the research on extensible business reporting language (XBRL): A bibliometric review. *International Journal of Innovation, Creativity and Change*, *5*(2), 1145–1167.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959-975.

Brooks, M. D. (2019). *Transforming literacy education for long-term English learners: Recognising brilliance in the undervalued*. Routledge.

Department of Employment and Workplace Relations. (2022). *Foundation Skills*. Australian Government. https://www.dewr.gov.au/foundation-skills

García, O. (2009). Education, multilingualism and translanguaging in the 21st century. In: Mohanty, A., Panda, M., Phillipson, R., & Skutnabb-Kangas, T. (eds). *Multilingual Education for Social Justice*: *Globalising the local*. New Delhi: Orient Blackswan, pp. 128-145.

Glew, P. J., Ramjan, L. M., Salas, M., Raper, K., Creed, H., & Salamonson, Y. (2019). Relationships between academic literacy support, student retention and academic performance. *Nurse Education in Practice*, *39*, 61-66.

Guo, X. (2022). A Bibliometric Analysis of Child Language During 1900–2021. *Frontiers Psychology*, pp. 1-17. Doi: 10.3389/fpsyg.2022.862042.

Hyland, K. (2007). Genre pedagogy: Language, literacy and L2 writing instruction. *Journal of second language writing*, *16*(3), 148-164.Rodríguez Jiménez, C., Sanz Prieto, M., & Alonso García, S. (2019). Technology and higher education: A bibliometric analysis. *Education Sciences*, 9(3), 169. MDPI AG. Retrieved from

http://dx.doi.org/10.3390/educsci9030169.

Kangan Institute. (2021). *The importance of literacy and numeracy skills*. TAFE Victoria. https://www.kangan.edu.au/study-with-us/our-blog/the-importance-of-literacy-and-numeracy-skills

Kayadibi, N. (2022). Bibliometric analysis of educational research on bilingualism. *International Journal of Education and Literacy Studies*, 10 (3), pp. 80-90.

Kim, G. M., & Johnson, L. L. (2021). Playful practices: Reimagining literacy teacher education through game-based curriculum design. *Research in the Teaching of English*, *55*(3), pp. 241-264.

Lan, X., & Yu, Z. (2022). A Bibliometric Review Study on Reading Literacy over Fourteen Years. *Education Sciences*, 13(1), 27. http://dx.doi.org/10.3390/educsci1301002

Liu Z. (2022). Mapping the research trends of third language acquisition: A bibliometric analysis based on Scopus. *Frontiers in Psychology*, 13, pp. 1-25. https://doi.org/10.3389/fpsyg.2022.1021517

Lysenko, L., Abrami, P. C., Wade, C. A., Marsh, J. P., WaGioko, M., & Kiforo, E. (2019). Promoting young Kenyans' growth in literacy with educational technology: A tale of two years of implementation. *International Journal of Educational Research*, *95*, pp. 176-189.

Marti-Parreno, J., Mendez-Ibanez, E., & Alonso-Arroyo, A. (2016). The use of gamification in education: A bibliometric and text mining analysis. *Journal of Computer Assisted Learning*, 32 (6). pp. 663-676.

Meza, C. S. R. (2021). A bibliometric analysis of academic literacy: A review of the state of the art, from the past to the future. *Revista de Educacion*, 394, pp. 63-94. DOI: 10.4438/1988-592X-RE-2021-394-501.

Mohd-Asraf, R., Abdullah, H., & Zamin, M. A. A. (2016). Literacy among Malaysian Primary Schoolers: How Do Boys Perform Relative to Girls? *International Electronic Journal of Elementary Education*, 9(1), pp. 225-238.

Moje, E. B., Collazo, T., Carrillo, R., & Marx, R. W. (2001). "Maestro, what is 'quality'?": Language, literacy, and discourse in project-based science. Journal of Research in Science Teaching: *The Official Journal of the National Association for Research in Science Teaching*, 38(4), pp. 469-498.

Murudkar, S. (2022). *ISI, SCOPUS or PubMed: How do I choose the best journal based on its indexing?* https://www.enago.com/academy/isi-scopus-or-pubmed/

Weiland, C., & Yoshikawa, H. (2013). Impacts of a prekindergarten program on children's mathematics, language, literacy, executive function, and emotional skills. *Child development, 84*(6), 2112-2130.

Xie, W. (2022). A bibliometric analysis on teacher's feedback research based on citeSoace through data visualisation. *Scientific Programming for Industry 5.0: Theory, applications and technological development*. https://doi.org/10.1155/2022/3789810

Yilmaz, R. M., Topu, F. B., & Tulgar, A. T. (2022) An examination of the studies on foreign language teaching in pre-school education: A bibliometric mapping analysis, *Computer Assisted Language Learning,* 35(3), 270-293, DOI: 10.1080/09588221.2019.1681465

Yore, L., Bisanz, G. L., & Hand, B. M. (2003). Examining the literacy component of science literacy: 25 years of language arts and science research. *International journal of science education*, 25(6), 689-725.

Statistics & Data. (2023). *The most spoken languages 2023*. statisticsanddata.org. https://statisticsanddata.org/data/the-most-spoken-languages-2023/