Advances in Engineering and Applied Science Research, ISSN 2617-1023
2021 International Conference on Intelligent Manufacturing Technology and Information Technology
Vol. 11, 2021, pp. 27-32

Systematic Review of Artificial Intelligence in Language Learning

Yunfei Du[1,a]

[1]School of Foreign Language, Chongqing University of Technology, Chongqing, China
[a] feeba@qq.com

ABSTRACT. Artificial Intelligence is widely developed and adopted in language learning over the last decade. In order to track and evacuate the hot spots for AI in language learning, CiteSpace, a bibliometric software, was launched to accomplish the co-citation analysis based on the datebase retrieved from Web of Science. The results firstly indicate that neural network is a dominant method in language learning, training machine to learn, read, write, listen, speak, and assess, while other technologies including user modeling, intelligent language tutoring, automated scoring, and data mining made AI applicable in this chosen field. Secondly, the influences that AI exerted mainly concentrate on such prominent scenarios as the transformation of personalized and adapted mobile learning and data-driven learning, the construction of authentic and motivated virtual worlds, and the reinforcement of intelligence aided reading and writing. Finally, the main target language that AI applied with is English, especially learning English as second language.

KEYWORDS: Artificial intelligence, Language learning, Bibliometric analysis, Citespace

1. Introduction

With the explosion of data resources, the improvement of computing capacities and the generational shift of data transmission, Artificial Intelligence (AI) has unleashed in the third boom since John McCarthy firstly proposed the idea of "intelligence" in 1950s. Along with the comprehensive and explosive development, AI creates numerous possibilities and approaches changed and changing the world. The most tangible feature of this change is that such AI-based technologies as machine vision, data mining, text analysis, and speech recognition are applied in the realistic application scenarios, and serve for various contexts. In light of that, the evolution involving machine learning, neural network and deep learning is taking place in the context of language learning, as the same as what it exerts influence on other fields.

To understand the adaptability and accessibility of AI in language pedagogy and learning, and excavate the hot spots for AI in language learning, it is vitally important to synthesize previous studies and study the emerging and interdisciplinary subject from the comprehensive and bibliometric perspectives. At present, the studies pertaining to this issue are confined with three factors: unlike the technological and theoretical success achieved in the field of artificial intelligence in education (AIEd), relatively few attention has been paid to AI in language learning; several synthetic studies emphasize on specific aspects such as intelligent tutoring system (ITS) and automated essay scoring (AES) for language learning; no attempt has been implemented to use the bibliometric method to measure the increasingly growing literature [1-3]. Therefore, the study initiates a bibliometric analysis to synthesize relevant and important studies of AI in language learning, for the sake of visualizing the hot topics in this chosen field.

2. Methodology

2.1 Data Collection

When collecting data for bibliometric analysis in CiteSpace, the bibliographic records retrieved from Web of Science Core Collection were selected as the primary data source. After consulting interdisciplinary experts in AI and language learning, 7 terms which are highly concerned and closely associated with AI including "artificial intelligence", "natural language processing", "machine learning", "deep learning", "speech recognition", "intelligent tutoring system", and "data mining", are used as representatives together with "language learning" and "language teaching" for retrieval in titles, abstracts, or indexing terms.

After data collection, clean-up and pre-processing, the database of this research has acquired and gathered 1014

  DOI: 10.23977/IMTIT2021007

publications including 491 articles, 498 proceeding papers, 7 book chapters, 11 reviews, 5 editorial materials and 2 meeting abstracts, published between the period from 2010 to 2019. The main reason of confining the publications within the period from 2010 to 2019 is that the research and application of AI have transformed from burst to boom alongside with the outbreak of deep learning since 2010 [4].

2.2 Visualization and Analysis

CiteSpace is a typical bibliometric software for progressive knowledge domain visualization initiated by Chaomei Chen [5]. The visualization graphs sketches a number of hot topics in a certain field, constructing with different types of nodes and links. The nodes refer to countries, authors, keywords, subject categories, and cited references while the links signify the relationships among these nodes.

After a descriptive analysis of annual outcomes, a latest version of CiteSpace (5.6.R2) is applied in this study to conduct co-citation analysis. It is defined that the co-citation relation occurs between two authors when the literature of the two authors is simultaneously cited by a later literature of the third author. Based on co-citation relations, nodes in the network are aggregated into different clusters in terms of the expectation maximization algorithm. These major clusters would be regarded as the embodiment of highly concerned topics in this field. Two indexes, modularity Q and silhouette, are used to measure the significance and validity of clustering. The modularity Q value over 0.3 indicates that the clusters in the network are significant while the silhouette value above 0.7 refers to high efficiency and validity of clusters. If the silhouette value were less than 0.5, it means the cluster is undesirable and would not be generated in the network.

3. Results and Discussion

3.1 Annual Outcomes

Figure 1 presents the progression of annually published articles, proceeding papers, reviews, and book chapters in the field of AI in language learning during the period from 2010 to 2019. The annual number of publications prior to 2012 was below 90 and since then gradually increased to 150 in 2016. After reached to the peak, the research outputs in this given field have been in decline for two years but remained over the number of 100. Although a sharp decrease is observed in 2019, the number of 58 may not be the overall outputs of this year due to many articles would be accessed in 2020. It proves that this subject is still emerging and insufficiently studied over the last decade.

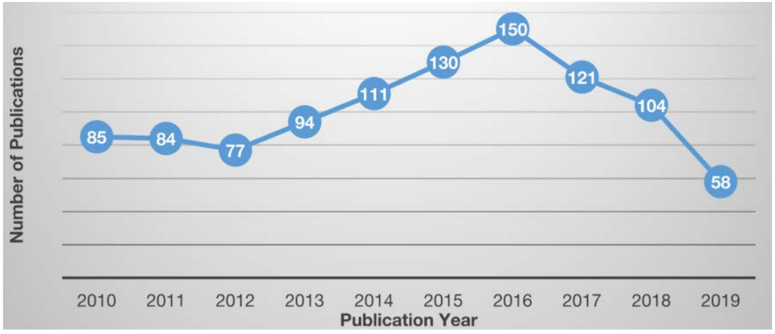

Fig.1 Number of Publications in the Field of Ai in Language Learning (2010–2019)

3.2 Co-Citation Analysis

Co-citation analysis is used to construct the clusters which comprises hot topics for AI in language learning. After the generation of co-citation network, CiteSpace aggregated 560 nodes in the matrix into a group of clusters which comprise corresponding research outcomes (Figure 2). The 15 clusters are filled with different colors in terms of time established and labeled by noun phrases from keywords, using a log-likelihood ratio (LLR) weighting algorithm. LLR is an algorithm to compare the fit of two models, has proven effective for discriminating and extracting pairs of features that have interesting degrees of co-occurrence. It is evidently indicated that the network is significant and valid due to the modularity Q of 0.8851 and the mean silhouette is over 0.5. Since there are 3 clusters with lower silhouette scores ranged from 0.5 to 0.7 including #10 research methodology, #12 psycholinguistics, and #13 synchrony evaluations, they are not so desirable that would not be taken into consideration in the further discussion.

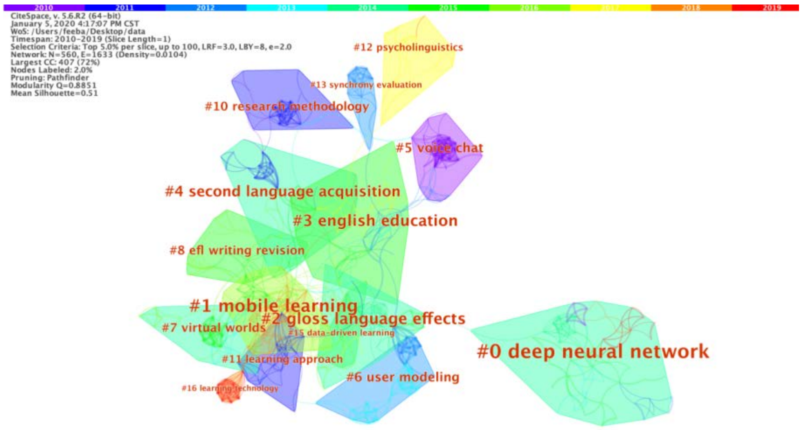

Fig.2 A Visualization of Co-Citation Clusters Labeled by Noun Phrases from Keywords

3.2.1 Dominant Ai Technologies Applied in Language Learning

The first part of the co-citation network concerns about dominant AI technologies applied in language learning. The largest cluster, #0 deep neural network, has 56 members which is relatively bigger than the rest, representing the wide application of neural network in language learning since 2010. Unlike the biology of human being's brain, this kind of artificial neural networks contain discrete layers, connections, and directions of data propagation[6]. A variety of types of neural network such as deep neural network (DNN), conventional neural network (CNN), and recurrent neural networks (RNNs) have been involved and become the common and fundamental methods to implement speech evaluation and writing assessment. It is typically exemplified that various perspectives of speech evaluation such as pronunciation, fluency, accent and syntactic complexity for language learners received a great number of concerns [7-8]. Other similar but more applicable models have been constructed to address more specific problems. For example, Karbasi et al. proposed a multi-modal-specifically audio-visual-speech recognition for detecting mispronunciations in acoustically noisy or otherwise challenging environments [9]. In addition, neural network also contributes to the evaluation of students' learning performance. A perceptron neural network program running with PyPhox mobile application and the Deductor Platform was developed by Prom et al., in which students' task can be fairly assessed [10].

Apart from neural network, cluster #6 and #16, the two clusters concentrate on other AI technologies applicable in language learning. Maria Virvou's lab has consistently worked on user modeling in the context of cooperative language learning and online learning since 2012. For example, a multiple language tutoring system called "Comulang" integrated with user modeling was tested and used in the creation of student groups and clusters based on the accumulation of user data and characteristics [11]. In the context of online learning, Virvou et al. embed a student modeling component with users' Facebook profile to provide advice in accordance to users' knowledge level [12]. Besides, other AI technologies including intelligent language tutoring, automated scoring, and data mining are derived from cluster #16 learning technology. It is exemplified that language tutoring systems have potential to achieve multidimensional functions more than just deal with simple grammar, vocabulary or sentence pattern in language learning. For instance, an ITS can be interacted with affective computing, which could recognize students' emotion in Japanese learning, explore a variety of learning styles and sequentially assist the improvement of learning efficiency [13]. Whitehill and Movellan implemented another approach to produce teaching policies by using ITS in concept learning tasks [14]. This approach, as a reinforcement hierarchical control technique, is capable to track learning path in depth, presenting a more satisfactory performance than two other manually designed teaching policies. In terms of data mining, it can be used to dynamically evaluate students' learning performance and automatically assist students working in collaboration in a multi-mode interactive foreign language learning community [15].

3.2.2 Prominent Scenarios

The second part of the co-citation network illustrates some prominent application scenarios of AI in language learning, which is composed of a large number of clusters, such as #1 mobile learning, #2 gloss language effects, #5 voice chat, #7 virtual worlds, #8 efl writing revision, #11 learning approach, #15 data-driven learning. It embodies many new features and changes that AI brought to language pedagogy and learning. The first change is that mobile learning and data-driven learning have been transformed more personalized and adapted that are capable to increase classification accuracy of learning strategies, cluster multiple user characteristics, provide learning spaces, recommend learning materials, and offer instructors the possibility of creating and editing the learning materials to meet learners' needs and learning preferences [16-18]. The second feature is that the AI-enabled virtual worlds has allowed language learning more authentic, motivated, interactive, engaging, and synchronously or asynchronously communicated [19]. The third change mainly lies in two such specific aspects as reading and writing, which are derived from cluster #2 gloss language effects and # 8 efl writing revision, respectively. In terms of reading enhancement, Chang and Hsu embed two modes with the functions of instant translation and instant translation annotation in a CALL system for the intensive reading course [20]. Gao built a data acquisition and feature pre-processing model in the integration of virtual learning environment and vocabulary retrieval, which embodied the language authenticity and enhanced the performance of vocabulary classification [21].

In accordance with the third change, from another perspective of EFL writing revision, the major adoption is to train the machine to assess language errors. In order to automatically extract and predict word order errors in learners' writing, Brendan Flanagan and his teammates identified error characteristics of different native languages with support vector machine (SVM), created an error corpus by training SVM classifiers, and applied the classifiers to automatically extract writing errors in a sequential manner from 2014 to 2018 [22-24]. In purpose of learning the relation between a text and its given grade, a RNNs method developed by Taghipour and Ng is introduced in the laboratory [25]. In another automated text scoring (ATS) system designed by Alikaniotis, Yannakoudakis, and Rei, a RNNs network manages essay representations and a DNN model accompanied is used to acquire both local contextual and usage information from compressed automated text assessment [26]. The two studies, coincidentally, employ a single-layer LSTM over the word sequence in modeling. In addition to and in accordance with the prior EFL writing researches, some Automated Essay Scoring (AES) systems have been already commercially available, including Intelligent Essay Assessor (IEA), e-rater, Criterion, IntelliMetric, MY Access!, Bayesian Essay Test Scoring System (BETSY), iWrite, and Pigai.com, which could automatically analyze the quality of the composition and give corresponding marks to the writing [27].

3.2.3 Target Languages

The third part of the co-citation network gathers the target languages that AI applied with. Cluster #3 with 37 members and cluster #4 with 33 members are labeled as english education and second language acquisition. It represents that the primary research coverage in this chosen field is to apply AI to facilitate, assist, and integrate with second language acquisition, especially learning English as second language while less attention has been paid to native language, sign language and minority language learning. For instance, a multi-language deep neural network (DNN) system adapted with transfer learning has been adopted to enhance non-native young students' speech, shared phonetic lexicon by Italian, English and German [28]. Also, a relatively small coverage lies in learning other languages including Dutch, Japanese, Arabic, and Spanish as second languages.

4. Conclusion

The bibliometric study based on CiteSpace has outlined a distinct and interdisciplinary sketch for AI in language learning over the last decade. In accordance with the descriptive analysis and co-citation analysis, it is indicated that AI in language learning, as an emerging and subcategory of computer-assisted language learning, comprises three significant groups of hot topics such as the dominant AI technologies applied in language learning, prominent scenarios and target languages. Firstly, neural network becomes a commonly promising method in this chosen field, training machine to learn, read, write, listen, speak, and assess, while other technologies including user modeling, intelligent language tutoring, automated scoring, and data mining made AI applicable in language learning. Secondly, with regard to prominent scenarios, the influences that AI exerted mainly concentrate on the transformation of personalized and adapted mobile learning and data-driven learning, the construction of authentic and motivated virtual worlds, and the reinforcement of intelligence aided reading and writing. In the third place, English is the main target language that AI applied with, especially learning English as second language, rather than that with native or sign language learning.

After tracking the hot spots of AI in language learning, the study is promising to encourage more researchers, teachers, schools and enterprises rationally and sequentially study, use, adapt and stimulate AI in this chosen field, so as to promote the breadth and scope of language learning. In spite of that, there is room to improve. The present research, which has been accomplished with influential bibliographic records, is restricted by virtue of the choice of publications. Along with the growing research outputs year after year, it will further require the extension to other data sources that include less influential but more emerging publications.

5. Acknowledgments

This work is funded by the [Chongqing Federation of Social Science Circles] under grant number of

[2019WYZX01], and Chongqing Municipal Education Commission under grant number of [20SKGH178], respectively.

References

[1] Luckin, R. & Holmes, W. Intelligence Unleashed: An argument for AI in Education. Person, 2016,[online]Available:https://www.pearson.com/corporate/about-pearson/innovation/smarter-digital-tools/intelligence-unleashed.html

[2] V. Slavuj, B. Kovačić, & I. Jugo, "Intelligent tutoring systems for language learning". IEEE 38th International Convention on Information and Communication Technology, Electronics and Microelectronics, pp.814-819, May 2015.

[3] Ettlinger, M., Morgan-Short, K., Faretta-Stutenberg, M., & Wong, P.C. "The relationship between artificial and second language learning". Cognitive Science, Vol.40, No.4, pp.822-847, 2016.

[4] [6]Copeland, M. What's the difference between artificial intelligence, machine learning, and deep learning? 2016. [online] Available: https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai

[5] Chen, C. "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature". Journal of the American Society for information Science and Technology, Vol.57, No.3, pp.359-377, 2006.

[6] A. Al Hindi, M. Alsulaiman, G. Muhammad, & S. Al-Kahtani, "Automatic pronunciation error detection of nonnative Arabic Speech". IEEE/ACS 11th International Conference on Computer Systems and Applications, pp.190-197, November 2014.

[7] S. Mao, Z. Wu, J. Jiang, P. Liu, & F.K.Soong, "NN-based ordinal regression for assessing fluency of ESL speech". IEEE International Conference on Acoustics, Speech and Signal Processing, pp.7420-7424, May 2019.

[8] M. Karbasi, S. Zeiler, J. Freiwald, & D. Kolossa, "Toward robust mispronunciation detection via audio-visual speech recognition". International Work-Conference on Artificial Neural Networks", pp.655-666, June 2019.

[9] N.A. Prom, E.A. Litvinova, A.V. Shokhnekh, T.G. Yovanovich, & N.I. Lomakin, "Perseptron for assessing the students' task performance in learning a foreign language according to the competencies using bi data". IOP Conference Series: Materials Science and Engineering, Vol.483, No.1, pp.012048, March 2019.

[10] Troussas, C., Virvou, M., & Alepis, E. "Comulang: towards a collaborative e-learning system that supports student group modeling". SpringerPlus, Vol.2, No.1, pp.387, 2013.

[11] M. Virvou, C. Troussas, J. Caro, & K.J. Espinosa, "User modeling for language learning in Facebook". International Conference on Text, Speech and Dialogue, pp.345-352. September 2012.

[12] Y.C. Ma, "The development of an affective tutoring system for Japanese language learners". In International Symposium on Emerging Technologies for Education, pp.363-371, September 2017.

[13] Whitehill, J., & Movellan, J. "Approximately optimal teaching of approximately optimal learners". IEEE Transactions on Learning Technologies, Vol.11, No.2, pp.152-164, 2017.

[14] C. Lv, "Application study on data mining technology of English learning virtual community". IEEE International Conference on Intelligent Transportation, Big Data & Smart City, pp.271-274, January 2018.

[15] S. Hu, & Y. Cheng. "Classification of college students' mobile learning strategies based on principal component analysis and probabilistic neural network". IEEE 4th International Conference on Computer Science and Network Technology, No.1, pp. 58-61, December 2015.

[16] Sandberg, J., Maris, M., & Hoogendoorn, P. "The added value of a gaming context and intelligent adaptation for a mobile learning application for vocabulary learning". Computers & Education, No.76, pp.119-130, 2014.

[17] C. Troussas, K. Chrysafiadi, & M. Virvou, "Machine learning and fuzzy logic techniques for personalized tutoring of foreign languages". International Conference on Artificial Intelligence in Education, pp.358-362, June 2018.

[18] Ibáñez, M.B., Rueda, J.J.G., Galán, S., Maroto, D., Morillo, D., & Kloos, C.D. "Design and implementation of a 3D multi-user virtual world for language learning". Educational Technology & Society, Vol.14, No.4, pp.2-10, 2011.

[19] Chang, C.K., & Hsu, C.K. "A mobile-assisted synchronously collaborative translation–annotation system for English as a foreign language (EFL) reading comprehension". Computer assisted language learning, Vol.24, No.2, pp.155-180, 2011.

[20] L. Gao, “Research on application of data mining in virtual community of foreign language learning”. IEEE International Conference on Intelligent Transportation, Big Data & Smart City, pp.331-334, January 2018.

[21] Flanagan, B., Yin, C., Suzuki, T., & Hirokawa, S. “Classification of English language learner writing errors using a parallel corpus with SVM”. International Journal of Knowledge and Web Intelligence, Vol.5, No.1, pp.21-35, 2014.

[22] B. Flanagan, & S. Hirokawa, “Automatic extraction and prediction of word order errors from language learning SNS”. IEEE 5th IIAI International Congress on Advanced Applied Informatics, pp.292-295, July 2016.

[23] Flanagan, B., & Hirokawa, S. “An automatic method to extract online foreign language learner writing error characteristics”. International Journal of Distance Education Technologies, Vol.16, No.4, pp.15-30, 2018.

[24] K. Taghipour, & H.T. Ng, “A neural approach to automated essay scoring”. Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing, pp.1882-1891, November 2016.

[25] Alikaniotis, D., Yannakoudakis, H., & Rei, M. “Automatic Text Scoring Using Neural Networks”. meeting of the association for computational linguistics, 2016.

[26] Hussein, M.A., Hassan, H.A., & Nassef, M. “Automated language essay scoring systems: A literature review”. PeerJ Preprints, 2019, [online] Available: https://peerj.com/articles/cs-208/

[27] M. Matassoni, R. Gretter, D. Falavigna, & D. Giuliani, “Non-native children speech recognition through transfer learning”. IEEE International Conference on Acoustics, Speech and Signal Processing, pp.6229-6233, April 2018.