Advances in Social Science, Education and Humanities Research, volume 124
International Conference on Contemporary Education, Social Sciences and Humanities (ICCESSH 2017)

The Visualization Analysis of Second Language Acquisition's Research Evolution

Yancheng Zhang
College of Chinese Language and Literature
Wuhan University
Wuhan, China

Wan Sun
College of Chinese Language and Literature
Wuhan University
Wuhan, China

***Abstract*—Using a visual analytical approach to explore the evolution of second language acquisition research, this paper was based on 3,940 papers including the theme of "second language acquisition" in the SCI, SSCI, and A&HCI databases from 1964 to 2014. It uses methods of bibliography modeling analysis to draw second language acquisition's mapping knowledge domains. Mapping knowledge domains show the interdisciplinary characteristic of second language acquisition research and the central role of language education and neuroscience. It also shows the evolving process of research hotspots changing from the brain, foreign language acquisition, and fossilization to bilingualism, language level, language cognition, and so on.**

Keywords—second language acquisition; mapping knowledge domain; knowledge visualization; bibliography

I. INTRODUCTION

Since the 1960s, second language acquisition (SLA) research has developed quickly with the help of the evolution of language technology. Early SLA research focuses on language education to direct teaching practices for second language education. Since the research technique has advanced, SLA research has maintained a close relationship with other areas of research through research findings. Currently, the independence of SLA research has been improved, which gives SLA the characteristics of being an interdisciplinary, cross-cutting, cross-time study. More and more researchers and institutions involved in SLA research. The total number of SLA research institutions are now increasing quite fast "Fig. 1".

This work is supported in part by funding from the Department of Language Information Management, Ministry of Education of the People's Republic of China, grant #YB125-146, as well as from the Language and culture research center at Wuhan, grant #2015A02..

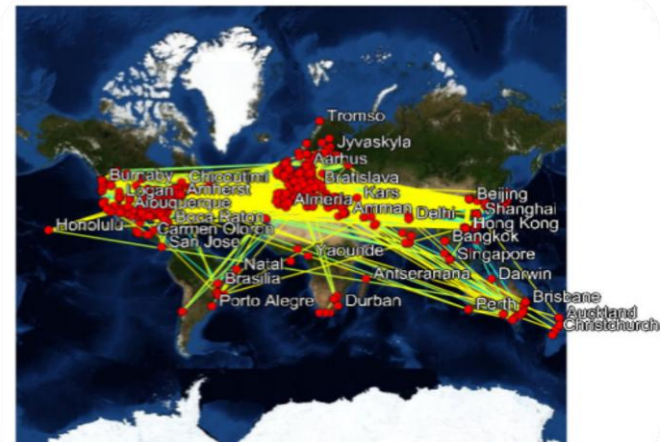

Fig. 1. Important institutions for SLA research (2004-2008).

The number of research results is growing yearly that represent the fact that SLA research has become a research hotspot in the area of language and linguistics research "Fig. 2".

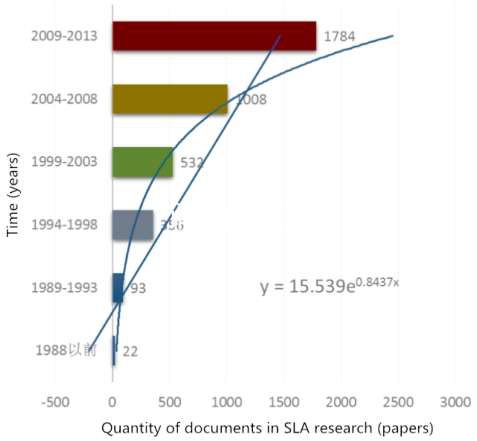

Fig. 2. Document numbers of SLA research in Web of Science database.

A number of studies pay attention to SLA's research status. Because it is limited by research techniques, the traditional study of SLA has the following limitations: first, the analysis may be guided by subjective views because scholars hold different points of view. Second, the summaries of SLA research are mostly not comprehensive due to the neglect of

interdisciplinary features in the study of SLA. Third, inaccurate mastering of research hotspots affects the prediction of future research trends. In this paper, the authors use bibliometric methods to describe the current research status of SLA, including most important papers in SLA research, research hotspots of SLA, and interdisciplinary nature of SLA.

In the Web of Science (SCI, SSCI, and A&HCI) database, we found 3,940 bibliographic records on the theme of SLA. We downloaded all of these bibliographic records, which formed a dataset of SLA research to map knowledge domains and show the cross-subject characteristic of SLA research and the central role of language education and neuroscience.

II. Analysis of Interdisciplinary Subject Areas

To explore the essence and rules of SLA and help second language learners to improve their target language level, linguistic scholars and researchers of other subjects have thoroughly researched a wild range of fields and research directions with regard to SLA. We have drawn the mapping knowledge domain of SLA categories formed by the main subject areas "Fig. 3".

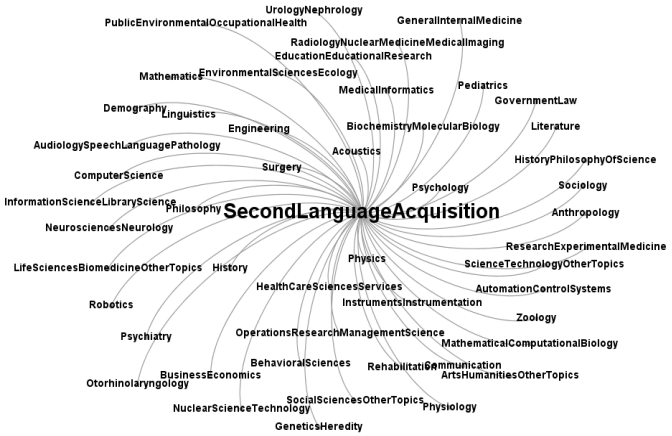

Fig. 3. Map of the area of direct concern regarding SLA.

Many disciplines are shown in the figure, including linguistics, medical science, mathematics, philosophy, physiology, zoology, pedagogy, sociology, physics, culturology, anthropology, demography, cognitive science, behavioral science, computing science, and so on. The cross relationship between SLA and different subjects varies by time and degree. To find the main categories of SLA research, we analyzed the bibliographic records of the top 10% of the most-cited papers published from 1963 to 2014. We then drew the mapping knowledge domain of the categories formed by the main subject areas "Fig. 4".

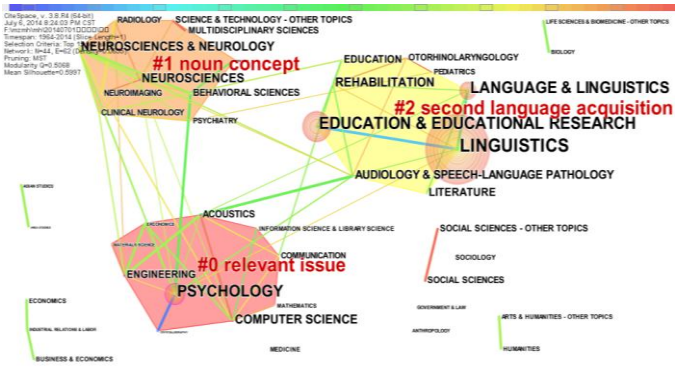

Fig. 4. Mapping knowledge domain of the main subject areas and categories in SLA research.

"Fig. 5" shows the size of the concentric circles representing the quantity of citations. The color of the citation ring signifies the date of the corresponding citations. The thickness of the ring is proportional to the number of citations in a given period of time. The small number next to the center of a node is the citation throughout the entire time interval. The categories of language technology formed in the subject areas are shown using polygons of different colors. The size of the polygons represents the quantity of citations in specific categories (Chen, 2006).

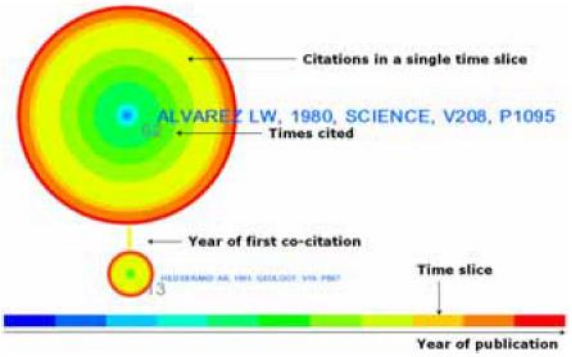

Fig. 5. Legends of the Mapping Knowledge Domain (Chen C, 2006).

As shown in "Fig. 6", the purple category labels indicate the most frequently used index terms in the related articles. The corners of each new polygon are the main subject areas and are marked in black. For example, the category "#3hearing" is formed by the following 3 main subject areas: "rehabilitation computer science," "rehabilitation" and "education."

Fig. 6. Example of a cluster.

Bibliographic records of the most-cited papers are shown in "Fig. 4", which reflects the universal focus of all scholars and the focal point of the debate to a certain extent. This mapping knowledge domain contains several main subject areas. In

addition, some independent hot-topic subjects are shown on the edge of "Fig. 4", such as “social sciences,” “social sciences-other topics,” and so on.

To present the interdisciplinary properties of SLA, we calculated the size and silhouette of each category and then listed the main categories, which are united by interdisciplinary subjects "Table I". "Table I" is organized in the order of category size. The name of a category showed in Figure 3 is linked to its silhouette value: the higher, the cleaner, and vice versa. The closer the silhouette value is to 1, the more certainty the precise category has. Particular years appear in the row of means, representing the time when a category first appears.

TABLE I. SLA RESEARCH CLUSTERS AND SUBJECT AREAS

Cluster ID & Cluster	Size	Silhouette	Mean (Year)	Subject areas
#0 relevant issue	10	0.715	1990	Computer Science Acoustics Communication Information Science & Library Science Mathematics Imaging Science & Photographic Technology
#1 noun concept	10	0.802	1994	Linguistics Education & Educational Research Language & Linguistics Rehabilitation Audiology & Speech-Language Pathology Literature Education Otorhinolaryngology Pediatrics
#2 second language acquisition	9	0.679	1987	Neurosciences & Neurology Neurosciences Behavioral Sciences Science & Technology - Other Topics Multidisciplinary Sciences Clinical Neurology Neuroimaging Radiology Psychiatry Zoology

According to "Table I" and "Fig. 4", the biggest cluster is “#0 relevant issue” in SLA research, which includes a number of subjects, such as psychology, computer science, communication, engineering, acoustics, mathematics, and so on. The “#0 relevant issue” cluster focuses on experimental research methods and the applicability of theoretical research. For example, Vargas and Delais-Roussarie (2012) propose a case study of the final contours and the prosodic structure observed in yes-no questions in French as an L2. The research was produced by 15 Mexican-Spanish learners of French (L2), 10 French speakers, and 10 Mexican speakers. The results prove that the acquisition of phrasing is more important than the acquisition of tonal patterns in French as an L2.

The cluster “#1 noun concept” explores SLA theoretically and practically from various aspects, including linguistics, education, literature, medicine, and otorhinolaryngology.

The cluster “#2 SLA” studies SLA using experimental methods from neurosciences, behavioral sciences, and science and technology.

According to the figure and table, the interdisciplinary nature of SLA is quite complex. The cross relationship exists in the internal cluster and between different clusters. Such a complex nexus leads to the development of a new research subject.

III. ANALYSIS OF IMPORTANT CATEGORIES

To show the research hotspots, we listed the 14 most popular subjects related to SLA "Table II".

TABLE II. 10 HOT SUBJECTS OF SLA

Centrality	Subject Areas & Time	Cluster #
0.17	Psychology, 1967	0
0.15	Computer Science, 1992	0
0.14	Neurosciences, 1991	1
0.11	Engineering, 1995	0
0.10	Neurosciences & Neurology, 1991	1
0.10	Audiology & Speech-Language Pathology, 1985	2
0.10	Acoustics, 1991	0
0.09	Rehabilitation, 1991	2
0.09	Clinical Neurology, 1994	1
0.06	Language & Linguistics, 1973	2

In the column “Subject Areas & Time,” a particular year represents when the certain subject developed an interdisciplinary relationship with SLA research. From "Table II", we find that philosophy, linguistics, and SLA have progressed to cross-discipline subjects very early, but some later subjects have a higher centrality rate.

Centrality value refers to the number of the shortest paths across a certain discipline, which measures a subject’s interconnect function in the area of SLA. The subjects with higher centrality may likely develop into important research nodes. The node subjects of SLA are mostly in the area of technology, for instance, psychology, computer science, neuroscience, engineering, audiology, and speech-language pathology. This is due to the technology feature in the Web of Science database. The papers recorded in the Web of Science are bias with regard to empirical, experimental and calculational research.

SLA and computer science are areas of interdisciplinary research especially with reference to new methods. For example, Fan (2011) reviews the principles of teaching, including cognitive, affective, and linguistic principles. The author thinks these principles are not isolated and should be combined. Teachers should apply teaching principles to teaching practices according to realistic needs.

De Vries (2011) believes corrective feedback (CF) has an important role in SLA. The tests on CF’s effectiveness have produced mixed results. Additionally, the tests on CF do not take into account individual differences. The author used a CALL system, which employed automatic speech recognition

(ASR) and was designed to adapt to individual learner differences.

Ng, Chow, and Chu (2011) consider the use of online games in the implementation of Japanese language education at the elementary level, which is both desirable and challenging.

IV. Analysis of Keywords Changing Trends

We drew the mapping knowledge domain of the evolution of the most frequently used keywords in the dataset "Fig. 6".

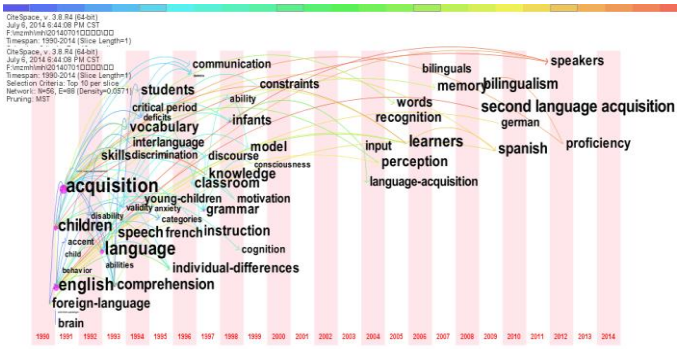

Fig. 7. Keyword Evolution in Language Acquisition.

"Fig. 6" displays the 10 most-used keywords for each year from 1990 to 2014. The size of a certain character represents the number of articles using this keyword. For example, in "Fig. 6", "language," "acquisition," and "SLA" have the highest frequency of occurrence because they are the search words.

According to the horizontal time axis, over the past 24 years, important keywords in SLA began to appear in the first 5 years. Meanwhile, some popular fields in SLA are continuing to develop. For example, "brain" became a popular keyword in 1991, showing that SLA research focuses on brain mechanisms of attention early on. In the database of SLA bibliographic records, half of most-cited papers used the keyword "brain."

The article Distinct Cortical Areas Associated with Native and Second Languages used functional magnetic resonance imaging (fMRI) to determine the spatial relationship between native and second languages in the human cortex. The results shown are located within Broca's area, whereas second languages acquired in adulthood ("late" bilingual subjects) are spatially separated from native languages. In the early language acquisition stage of development, native and second languages tend to be represented in common frontal cortical areas. This discovery of language-specific regions in Broca's area advances our understanding of the cortical representation underlying multiple language functions.

One of most-cited articles, The Bilingual Brain-Proficiency and Age of Acquisition of the Second Language, used functional imaging methods in a PET investigation on bilingualism to show differences in the pattern of cerebral activation associated with the subject's native language (L1) compared with a second language (L2). Several brain areas, similar to those observed for L1 in low-proficiency bilinguals, were activated by L2. These findings suggest that, at least for pairs of L1 and L2 languages that are fairly close, attained proficiency is more important than age of acquisition as a determinant of the cortical representation of L2.

In addition, the keywords "foreign language," "English," "French," "German," and "Spanish" became popular keywords in 1990, 1991, 1996, 2009, and 2010, respectively. This phenomenon shows that in the background of globalization, the widespread languages are paid more attentions.

Keywords of language elements appear in sequence, such as "accent," "vocabulary," "grammar," and "discourse." The sequence of these words generally remain in line with the shifts of research focuses in linguistics field.

Interlanguage has been a popular topic since 1993. The hotspots related to interlanguage are as follows: studies on fossilization, variability, developmental patterns, linguistic universals, and so on. The natural development process and representation features of the interlanguage are the latest hotspots to appear.

The discussion of the ability of language communication has triggered other popular keywords, such as "ability," "skill," "communication," and so on.

Krashen's input hypothesis was proposed quite early, but until 2004, the research value of this topic was not fully recognized. The hypothesis of interaction was proposed earlier than the input hypothesis but became a popular topic soon after it was published. The discussion and criticism of these two hypotheses made them longstanding research focuses.

By integrating keywords into related research groups, four research groups formed automatically, called the following: "#0 sensitive period," "#1 age," "#2 morphology," "#3 teacher education," and "#4 L1 transfer." The aforementioned four groups show the four central problems in SLA research. "Fig. 8".

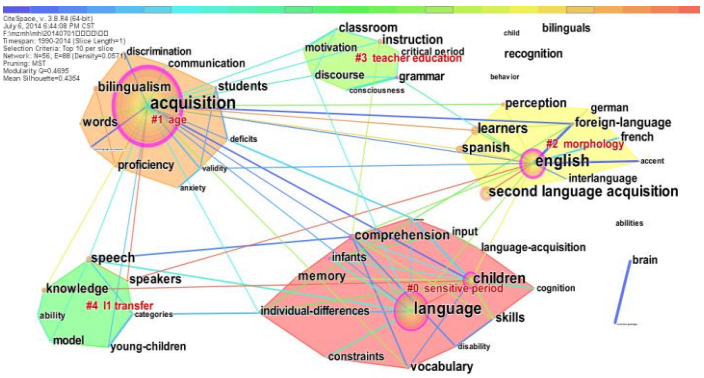

Fig. 8. Clusters of Most-cited Keywords of SLA.

It is apparent that there are numerous ligatures between the clusters in Fig.8 , especially those between cluster "#0 sensitive period" and cluster "#1 age," which indicates their close relationship. Both the clusters names associate with time factors influenced in L2 acquisition while the former focus more on children and the latter on adults. The main research content of these two clusters contains the study of bilingualism and mental lexicon organization under bilingual conditions. Due to the complex links between two language systems, when

bilinguals use one language, the other is likely to be activated spontaneously. This will produce two languages in activation and inhibition.

The effect of age factors in acquiring a second language is an important aspect of cluster "#1 age." Due to the openness of the problem itself, many issues are still debatable at present. This problem has both theoretical and practical significance. Furthermore, researchers have investigated and analyzed the topic of critical age from different perspectives. For example, in discussing the relationship between SLA and first language acquisition from a neurophysiology perspective, three main variables exist: the age of SLA, level of SLA, and performance of subjects in completing language tasks.

From the angle of social psychology, adults and children, especially young children, performed quite different in language acquisition. Cognitive perspective research tends to focus on the problem of age. As for the study materials, most researchers use lexicological studies, and others use phonetic and grammar studies.

Cluster "#2 morphology" concerns more about L2 Learners' acquisition of inflected forms, suffixes and morphosyntax of their interlanguages. Orders and sequences in the acquisition of L2 morphemes is a general topic. Perceptive approaches gain extensive attention. This hotspot includes several wide-ranging languages, such as English, Spanish, French, and German.

Cluster "#3 teacher education" pays close attention to the important role of teachers in language education. Quantitative and qualitative methods are commonly used in the study of L2 classroom arrangement, L2 learners' motivation and L2 grammar instruction.

Cluster "#4 L1 transfer" is very controversial. At present, language transfer is the most complex concept in SLA research. With the exception of the debate over the definition of L1 transfer, the function of first language in SLA is the most significant hotspot. In contrast with the key words "perception" "comprehension" in cluster "#2 morphology" and "#0 sensitive period", the key words "speech" "speaker" in this cluster shows a different approach which focus on speech production. Different models adopted in this cluster, such as unified model, typological primacy model, Full Transfer/Full Access model, parameter-setting model, etc., is a significant phenomenon.

V. CONCLUSION

Through visual analysis, we have discovered some features of SLA research development from 1963 on macro level.

Firstly, SLA is a cross discipline and a comprehensive research field. This feature is a similarity shared by all contemporary disciplines. Through the integration, intercrossing, and expansion of modern disciplines, different disciplines will gradually form.

Secondly, SLA has many hotspots, and these hotspots have been researched for a long time. The upcoming trend of SLA may strengthen its scientific nature and develop the intelligence surrounding it. Overall, SLA concentrates on realistic practices and has been widely used in communication, network, computer science, education, rehabilitation, and management. With the progress and innovation of technology, the research on SLA will extend further in realistic practice.

Thirdly, SLA research asks scholars to be sensitive to new technologies and methods which constitute the key factors for promoting the subjects.

REFERENCES

[1] Chen, C. M. 2006. "CiteSpace Π Detecting and Visualizing Emerging Trend and Transient Patterns in Scientific Literature". *Journal of the American Society for Information Science and Technology* 57:359–377.

[2] de Vries, B. P., C. Cucchiarini, H. Strik et al. 2011. Adaptive Corrective Feedback in Second Language Learning. Berlin: Springer-Verlag Berlin.

[3] Fan, L. P. 2011. *Research on Methods of Improving L2 Language Learning Based on Principles of Language Teaching with Web Resources*. Berlin: Springer-Verlag Berlin.

[4] Kim, K. H., N. R. Relkin, K. M. Lee et al. 1997. Distinct Cortical Areas Associated with Native and Second LSanguages. *Nature* 388(6638): 171–174.

[5] Ng, S., C. Chow, and D. Chu. 2011. *The Enhancement of Students' Interests and Efficiency in Elementary Japanese Learning as a Second Language through Online Games with Special Reference to Their Learning Styles*. Berlin: Springer-Verlag Berlin.

[6] Perani, D., E. Paulesu, N. S. Galles et al. 1998. "The Bilingual Brain - Proficiency and Age of Acquisition of the Second Language". *Brain* 121(10): 1841–1852.

[7] Vargas, F. S., and E. Delais-Roussarie. 2012. Acquiring Phrasing and Intonation in French as Second Language: The Case of Yes-No Questions Produced by Mexican Spanish Learners. *Proceedings of the 6th International Conference on Speech Prosody Vols I and II* 338–341.