CiteSpace II: Idiom Studies Development Trends

Wenyu Liu Ph.D.
Haibo Shen
School of Foreign Languages
Dalian University of Technology, China

ABSTRACT

Idioms, frequently used in daily language, are a typical metaphorical language and may be a cue to uncover the universal language processing mechanism. For the purpose of better mastery of the trends and front of idioms studies, CiteSpace II, an application designed to detect and visualize the development process within a scientific field, is adopted for comprehensive literature review. It is found that (1) idioms studies have thrived since 1990s with American scholars contributing the most, especially those from University of California; (2) suppositions on idiom comprehension mechanism have been inspired by different scholars including Lakoff, Swinney and Gibbs; (3) the exploration of the neurological bases for idiom comprehension has become the pursuit of researchers across different domains.

Key Words: idiom studies, CiteSpace II, citation analysis, development trends

1. Introduction

Idioms are a common language phenomenon and a typical metaphorical language as well. Compared to metaphors and irony, both the construction and non-literal meanings of idioms are more conventionalized. Idioms are also frequently used in every language, thus may be a cue to uncover the universal language processing mechanism. Therefore, scholars across the world have conducted numerous studies on idioms, and proposed different theories concerning idiom comprehension mechanism. Furthermore, with the adoption of modern technologies like corpora, functional magnetic resonance imaging (fMRI), repetitive transcranial magnetic stimulation (rTMS), event-related potentials (ERP) and eye-tracking, research into idiom has been greatly broadened. It seems necessary to have a better understanding of the development and trends of idiom studies, in order to thoroughly investigate the idiom comprehension mechanism that is endowed with human.

Currently the achievements of scientific research are represented by publications, mainly academic journals, books, meeting periodicals, etc. Thus development process of idiom studies could be reflected by academic publications during a certain time span. Scientific literature generally falls into two groups according to the citation half-lives of the articles, namely the classic articles and transient ones. Classic articles include those that have been frequently cited since their publications while transient articles refer to those that are highly cited.
within a short period of time. (Price, 1965; Chen, 2006) The term research front was first proposed by Price (1965), normally referring to the “actively cited” articles within a specific scientific domain (Chen, 2004). The research front, to some extent, represents the latest research directions since scholars always tend to cite the most recent publications in their papers. In contrast, intellectual base is comprised by the articles cited in research front, reflecting the classical theories and knowledge foundations. In other words, the intellectual base may be defined as “the citation and co-citation footprint in scientific literature of a research front”, as Chaomei Chen (2004) puts it. In addition, according to Small and Griffith (1974), the currently activated scientific research may be reflected and discerned by clusters of co-cited papers, namely, the co-citation cluster. Therefore, through partitioning co-citation clusters and identifying pivotal points, the intellectual turning points and burst terms can be detected, further demonstrating the dynamic changes and developments in idiom studies.

2. Methodology
2.1 Instrument
Considering the vast volume of experiments and papers on idiom, the use of scientometrical tool like CiteSpace would be of great help. CiteSpace, a freely available Java application, is designed to help scholars better understand the evolution, development and trends within a certain scientific area by identifying the intellectual turning points and dynamically visualizing the citation networks. CiteSpace, developed by Chaomei Chen from Drexel University, is of great help considering the increasing difficulty of filtering valuable information and materials in an age of explosive data everywhere. Thus using it, the co-citation networks in the domain of idiom studies will be depicted, the intellectual base and research front will be discerned, and the critical points will be analyzed, thus presenting a general picture of idiom studies.

CiteSpace II version 3.1 R3 (64 bit), the most up-to-date one, is used in the current study to answer the following questions: (1) What is the latest research front in idiom comprehension? (2) What are the hottest research front terms? (3) Which articles are associated with these terms?

2.2 Data Sources and Processing
Adopting the method of CiteSpace analysis, it is essential to clarify the source of literature in the first place, in order to gain a comprehensive and thorough understanding of the current situation and trends within the idiom related research domain. For the purpose of encompassing all valuable literature, all the papers for co-citation analysis in the current paper were accessed and downloaded from three databases, namely Social Sciences Citation Index (SSCI), Science Citation Index Expanded (SCI-E) and Conference Proceedings Citation Index-Science (CPCI-S) through Web of Science which is regarded as the largest online academic citation index providing access to six major indexing databases. Via key words searching, bibliographic records including authors, titles, abstracts, references and sources could be retrieved for further study. More specifically, searching with idiom or idiomatic as the topical terms, confining the publication years to the period from 1960 to the present, excluding the document types of note, discussion, database review and software review, 1855 pieces of records were obtained and then imported into the software CiteSpace II for further analysis with the following procedures. (1) Create a
new project based on the bibliographic records. (2) Select ten years as the length of a single time slice, dividing the whole range of time interval into 6 slices. (3) Select top 30 most cited items from each slice as the threshold level. (4) Display the standard layout of cluster view and timezone view. (5) Inspect the display of visual attributes by adjusting different parameters.

3. Data Analysis and Discussion

3.1 Temporal and Geographical Distribution of Publication Records

Concerning the number of idiom related research, it is easy to draw from figure 1 that the published items within this academic area during the past five decades constantly increased in general, especially after the year 1990. Furthermore, over 100 items were published each year from 2007 to 2011, suggesting that scholars have paid increasing attention to the issue of idiom.

Figure 2 displays the top ten nations with the most productive achievements in the domain of idiom study from the year 1960 to 2012. Scholars from the United States published 794 academic items, accounting for 42.87% of all publications records. Next to USA, England and Canada claim 171 and 103 respectively. Altogether 1068 published works were generated in these three nations, occupying nearly 60%, which is in accordance with the fact that English idiom related studies are more thorough and comprehensive compared to other languages.

In addition, the top ten organizations that have contributed the most to idiom related studies are depicted in figure 3. Similar to the geographical distribution, universities and institutions in United States conducted the most research, with University of California (Santa Cruz), Harvard University and University of California (Los Angeles) laying top of the list.
3.2 Citation analysis

CiteSpace II provides different display modes, including cluster view and time-zone view. Time-zone view highlights the co-citation network changes with time; while cluster view emphasizes on the division of co-citation clusters within a time span. In either time-zone view or cluster view of co-citation network provided by CiteSpace II, several critical attributes are represented by specified rings, lines and colors. A ring depicts the citation history of a cited reference, with its thickness denoting its amount of citations within a time range. The larger the ring, the more cited the reference. A line between two rings reflects the co-citation link of two cited references, with its thickness showing the strength of co-citation and its color showing the time of the first co-occurrence. The color bar on top indicates different time slices of publication years. Rings and lines with the certain colors pertain
to the corresponding time range. In addition, red is usually used to label the citation burst and purple is added to a ring to demonstrate a high betweenness centrality. According to Chen et al. (2010), betweenness centrality, measured according to the number of links passing a node in a network, implies the degree of significance of a node. Thus a node with high betweenness centrality value and citation frequency usually signifies a revolutionary scientific work that proposes new theories or innovations.

3.4.1 Time-zone View of Co-citation Network
Figure 4 below is the time-zone view of the co-citation network of idiom related publications in which nodes with high betweenness centrality (greater than 0.05) are labeled. It is evident that idiom related research have thrived since the end of 1980s with the publication of the three scholars’ works which are highlighted with labels.

The two books by George Lakoff, *Metaphors We Live By* (1980) and *Women Fire Dangerous* (1987), with the betweenness centrality values of 0.1 and 0.07 respectively, laid the foundation of idiom comprehension studies. Lakoff, an American linguist, proposed the profound idea of conceptual metaphor, arguing that metaphor is fundamentally conceptual rather than linguistic. In contrast to the traditional view that metaphor is one form of rhetoric, metaphor, according to this conceptual metaphor theory, is the main mechanism through which people comprehend abstract concepts and perform abstract reasoning and metaphorical language is merely a surface manifestation of conceptual metaphor. Inspired by the completely different perspective on metaphor, gradually more linguists turned their attention to comprehension and cognition mechanisms of metaphors, even language, which is a significant turning point in the scientific domain of figurative language study. Since then, figurative language research has been deeply connected to human cognition rather than a mere rhetoric tool of language or literature. In virtue of this theoretical breakthrough, idioms, a typical figurative language with both literal meanings and figurative meanings, have become a hot topic for linguists and cognition researchers. In addition, Lakoff’s view of conceptual metaphor that languages are based on concept metaphors exerted profound influence on Raymond Gibbs and inspired his far-reaching decomposition hypothesis (more detailed illustration of Gibbs and his ideas on idiom can be found below).

From figure 4, it is easy to distinguish out the article *Lexical Access During Sentence Comprehension: (Re) Consideration of Context Effects* by David Swinney (1979) in which the online method of Cross-Modal Priming Task (CMPT) was first developed, making great contributions to make up the deficiency of offline measures in lexical and syntactic processing investigations and improve psycholinguistic research techniques. Thus he investigated the comprehension process of sentences with ambiguous words in hope of identifying the retrieval order of words’ meanings, i.e. the order that people access to different meanings of a polysemic word. Experiment results suggested that people tend to retrieve multiple meanings of ambiguous words, regardless of the contexts orientation (Swinney, 1979). Considering the duality of idiom meanings (literal meaning and non-literal meaning), the above mentioned conclusion was greatly provoking for idiom comprehension, leading to the Lexical Representation Theory (or Simultaneous Processing Hypothesis) which was influential and accelerated idiom
studies. To be more specific, according to Swinney and Cutler (1979), idioms are stored in mental lexicon as long words which have no difference from ordinary phrases. Both the literal and non-literal meanings will be activated upon hearing or reading an idiom, and then single out the correct meaning that fits into the context. Thus the processing time of idioms and normal phrases are almost the same, which is consistent with the experiment results and refutes the Idiom List Hypothesis (or Literal-First-Hypothesis) by Bobrow and Bell (1973) who claim that literal meanings will be accessed first in idiom comprehension.

According to figure 4, Raymond W. Gibbs Jr. with his paper How to Kick the Bucket and not Decompose: Analyzability and Idiom Processing published on Journal of Memory and Language in the year 1989 claims a betweenness centrality value of 0.06 and citation frequency of 76, implying a significant influence on idiom study. Gibbs is an experimental psycholinguist and cognitive scientist from University of California, Santa Cruz, mainly focusing on people’s use and comprehension of figurative language including metaphor, irony and idioms. He once proposed the widely accepted Direct Access Hypothesis, suggesting that only the non-literal meanings are activated first without reference to the literal meaning because of the strong conventionality feature of idioms. Unable to explain why the literal meanings would also be activated upon idiom comprehension in some cases, he challenged the previously widely accepted idea of indecomposable idioms by proposing the Decomposition Hypothesis. By this theory, he believes that the literal meanings of each component of idioms contribute to the real whole meaning, or in other words the figurative meaning of idioms to some extent. According to Gibbs, all
Idioms fall into three categories. The first type is non-decomposable idioms whose real meanings cannot be directly comprehended via literal meanings of each word, like *kick the bucket* and *spill the beans*. The second type is normally decomposable idioms whose real meanings can be roughly grasped merely through literal meanings, like *add fuel to the flames* and *zip your lips*. The third type is abnormally decomposable idioms whose real meanings involve deep cognition and concepts systems like the use and comprehension of metonymy and metaphor. Idioms like *carry a torch* belong to the third type. (Gibbs, 1989) With his classification of idioms, Gibbs provides a deeper perception on the composition of idioms which has been widely accepted by other researchers, leading to more considerations about the feature of decompositionality or compositionality of idioms in the recent studies including empirical experiments on idiom comprehension. Gibbs and his insights into idioms are also inspiring and prompting for the emergence of the latter hybrid perspectives on idiom comprehension like Giora’s Graded Salience Hypothesis (Giora, 1999) and Sprenger’s Superlemma Theory of Idiom Production (Sprenger et al. 2006) that are more integrated and plausible to address the issue of idiom comprehension.

### 3.4.2 Cluster View of Co-citation Network

The cluster view of co-citation network of idiom studies is depicted as figure 5 with labels for different clusters. From this chart, it is apparent that idiom publications during the period from 1960 to 2012 can be classified into seven clusters with four core ones (cluster 4, 5, 6, 7). The four core clusters primarily focused on idiom processing mechanisms, either children acquisition or adult understanding, either the health or patients, thus forming the nucleus of idiom research (as is reflected by key terms and publications listed in Table 1). From Table 1, it is clear that scholars like Gibbs, Cacciari, Lakoff, Glucksberg, Rapp and their colleagues have contributed a lot to idiom studies involving various aspects including children acquisition, adult comprehension, structure composition, semantic processing, psychological and neural foundations. Among these works, Superlemma Theory of Idiom Production by Sprenger and the Graded Salience Hypothesis by Giora have been always quoted as the theoretical foundation for recent experimental studies.
The Superlemma Theory of Idiom Production holds that every idiom could be considered as a composition of several small lemma or words which could be activated by a superlemma. Corresponds to the idiom’s real meaning, the superlemma are activated by a specific concept related to the idiom. For example, the concept of *dying* may activate the superlemma *kick the bucket*, then further activate every lemma *kick*, *the* and *bucket*. Meanwhile the other superlemma like *bite the dust* can be activated simultaneously by the concept of *dying*. That is the reason why idiom production and comprehension may take longer time than simple words and phrases (Sprenger et al., 2006). This theory involves not only idiom processing but also idiom production, which covers the weaknesses in idiom production research. While the Graded Salience Hypothesis focuses on the feature of saliency, avoiding the literal and non-literal divide (Giora & Fein, 1999). It argues that the salient meanings rather than the literal or non-literal meanings are activated first regardless of the contextual constraints. More specifically, the salient meanings of idiomatic phrases refer to the highly conventionalized meanings stored in people’s mental lexicon, either literal or non-literal meanings. In other words, the salient meanings are those that first occur to people upon reading or hearing idioms, with marginal relations to contexts or component words. *Salience* here is not a question of either/or, rather a continuum graded by several factors like conventionality, familiarity, frequency and prototypicality. The comparatively more salient meanings would be accessed faster and retrieved earlier than the less salient ones. Only after the activation of the salient meanings, the contextual effects would take control. That is, if the salient meanings are not fitted into the contexts, the non-salient ones would then be activated by language users (Giora, 1999, 2003). This new perspective on idiom comprehension gradually attracted more attention in recent years and has been tested via different experiments with the adoption of rTMS, ERP and fMRI (Turner & Katz, 1997; Oliveri et al., 2004; Laurent et al., 2006; Cieslicka & Heredia, 2011).

From the above cluster view of co-citation network, references in cluster 7 are generally highlighted in yellow or orange which represent the time range from 2000 to 2012, suggesting that the potential research front, namely, exploring the idiom processing mechanism through studies on schizophrenic patients or healthy adults, lies in this area. In addition, concerning the major journals where the major idiom works published, the interaction between psychology, neurology and linguistics has continuously boosted with the formation of inter-disciplines of psycholinguistics, neuro-linguistics and cognitive linguistics. This reflects the current growing trend of querying about the neurological bases for idiom comprehension with more advanced technologies including fMRI, rTMS, ERP and eye-tracking. However, controversies involving the biological foundations of idiom comprehension still remain in spite of progress and advancements. The traditional view that the right hemisphere is responsible for figurative language comprehension (Kempler et al., 1999; Eviatar & Just, 2006) has been gradually challenged with more evidence on the left hemisphere involvement (Oliveri et al, 2004; Rapp et al., 2004; Lee & Dapretto, 2006; Stringaris et al., 2007; Romero Lauro et al, 2008). Elaborately designed studies on both right and left hemisphere damaged patients have proved the significant role of both hemispheres in idiom comprehension, in consistent with studies on health participants using rTMS (Papagno et al., 2002; Oliveri et al., 2004). To resolve the disputes over idiom comprehension, more close and active collaborations between linguists and neurologists should be initiated to work out more comprehensive experiment designs taking all essential linguistic factors like
contexts, conventionality and society into consideration.

Table 1. Key Publications in Clusters 4, 5, 6, 7 During 1960-2012

<table>
<thead>
<tr>
<th>Clusters</th>
<th>Key Terms</th>
<th>Titles of Key Publications</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Compositional idiom; Dissociation; Semantic processing; Psycholinguistic study</td>
<td>Metaphors We Live By (Lakoff &amp; Johnson, 1980) Syntactic Frozenness in Processing and Remembering Idioms (Gibbs &amp; Gonzales, 1985) On the Process of Understanding Idioms (Gibbs, 1985)</td>
</tr>
</tbody>
</table>
4. Conclusion

Through the aforementioned co-citation analysis with the help of CiteSpace, it is easy to perceive the shifts of idiom research cores as well as the improvement of methodology adopted by researchers. More specially, three major development features and trends can be detected.

First, idioms, as one typical figurative language, share numerous commons with metaphors. New theories on metaphor comprehension have shed lights upon idiom studies, encouraging different perspectives. Therefore, idiomatic research has gradually emphasized more on the cognition essence rather than the semantic origins with the emergence and prevalence of the idea that metaphor is a way of thinking.

Second, taking research focuses into consideration, the popularity and advancement of inter-disciplines would be reasonable. Beneficial to the whole society and human development, exploration into figurative languages including metaphor, idiom and irony is essential to clarify the nature of language comprehension, requiring frequent cooperation and mutual learning. Knowledge and methods across different domains are thus shared and synthesized for an integrated and thorough understanding. This also applies to idiom studies, making psycho-linguistic and neuro-linguistic research even more promising and critical.

Third, in order to identify the psychological and neurological bases for idiom acquisition and comprehension, experiment and research methods and technologies (interviews, questionnaires, computer sciences, fMRI, ERP, etc.) widely accepted in other fields like psychology, neurology and sociology have been introduced into linguistics and adopted in those inter-disciplinary investigations. Cooperation and co-prosperity further promote innovations and breakthroughs.

Acknowledgement: This research is supported by the Fundamental Research Founds for National University, Dalian University of Technology (DUT13RW303).

References


