

Description of the Lexical Meaning Structure of Geological Chinese Vocabulary and its Synset Construction

Shan Xiao ¹, Jinzhu Zhang ², Yuzhen Zhang ³

¹ International Education College
China University of Geosciences
Wuhan, Hubei 43007, China
wdxshan@aliyun.com

²The PLA University of Foreign Languages
Kunshan, Jiangsu, 215300, China
95339041@qq.com

³ School of Foreign Language
China University of Geosciences
Wuhan, Hubei 43007, China
572396767@qq.com

Selected Paper from Chinese Lexical Semantic Workshop 2014

ABSTRACT. Vocabulary teaching plays an important role in Teaching Chinese as a Foreign Language (TCFL). The purpose of scientific Chinese vocabulary teaching is to strengthen the professional and technical Chinese capacity of foreign students doing majors or for a specific business use. The domestic and foreign research of scientific Chinese vocabulary has several problems: less stringent structure, rough semantic particle size and limited range of application, which does not include professional fields. This paper considers lexical semantic networks in Geology based on ‘Synet-Lexeme Anamorphosis’. This method is an effective solution to the problems. It is also a valuable and useful method to analyze the derivations and differences between geological Chinese vocabulary and basic Chinese vocabulary.

Keywords: Conceptual Semantic Features; Geological Chinese Vocabulary; Teaching Chinese as a Foreign Language

1. Introduction. International promotion of the Chinese language needs the major support of Teaching Chinese as a Foreign Language (TCFL). TCFL can be divided into two types which are ‘Chinese for General Purposes (CGP)’ and ‘Chinese for Specific Purpose (CSP)’ according to function and purpose.

According to the resources allocation and investment proportion in the current market, we may find out that although the studies of CGP and CSP both started from the 1980s, the development of CSP cannot be compared with CGP because of its greater difficulties and challenges. That is why there is a lack of appropriate teachers, textbooks and scientific research. The relevant research and the Monographs which we can find number less than one hundred. The published CSP textbooks are not only few in number, but also more concentrated in the fields of Business Chinese and Chinese Medicine, which don't cover all the teaching areas. This is not proportional to the social needs. Along with the development of China's economy, science and technology, our country has attracted a large number of overseas degree students. The non-liberal arts colleges or specialized institutions such as Schools of Science and Engineering, Military academies, Financial and Economic Colleges and Geological Colleges need to focus on teaching and researching scientific Chinese vocabulary teaching and research combined with professional theoretical contents. To take our China University of Geosciences (Wuhan) as an example, more than 70% to 80% of all the 700 foreign students major in geology or a related discipline.

Unlike the language students, those students can only obtain China's diplomas and degrees after undertaking the formal academic education in our country. So they will face all the difficulties in major learning immediately after they have completed the one-year Chinese language foundation training. Because of lack of transition from daily Chinese language to expertise, many foreign students had bad academic performance or even dropped out. We should take these students as the main objects of CSP and find a more effective way to help them clear the language barriers or decrease their anxieties and uncomfortable feelings. In this paper, we use Geological Chinese Vocabulary (GCV) as one typical case to do the research and hope we can provide a useful reference for other Scientific Chinese Vocabulary researches.

2. Semantic Features and Structures Analysis of GCV. Vocabulary Teaching is long-term work, which runs through the whole TCFL. It will affect the ability of listening, speaking, reading and writing of Chinese learners directly. Therefore, our study should first start from the professional vocabulary teaching, make clear the intension and extension of GCV and find out the connections or differences between GCV and the ordinary Chinese vocabulary. GCV has its specific semantic features and lexical meaning structures which can be used to explore a new way of improving the Chinese professional language ability for foreign students.

2.1. Definition and category of GCV. GCV belongs to Scientific Chinese Vocabulary which has the features of scientific style but carries the unique Geological knowledge and concepts as well. This can be divided into three categories.

The first type is the word which only has a single specialized meaning and fixed usage, such as ‘石英 [shiyin](quartz)’ ‘闪煌岩 [shanhuangyan](camptonite)’ ‘砾岩

[liyan](conglomerate) ‘刚玉[gangyu](corundum)’ ‘角峰[jiaofeng](horn peak)’ ‘绿柱石[lüzhushi](beryl)’ and etc. The second type is distributed widely in geological textbooks or references. It has a complete semantic interpretation but also has the ability to form other related words. For example, ‘作用[zuoyong](action/function)’ can make the words such as ‘风化作用[fenghua zuoyong](weathering)’ ‘冰蚀作用[bingshi zuoyong](glacial erosion)’; ‘运动[yundong](movement)’ can form the words like ‘地壳运动[diqiao yundong](crustal movement)’ ‘造山运动[zaoshan yundong](orogenic movement)’; ‘圈[quan](sphere)’ can be used to build the words like ‘水圈[shuiquan](hydrosphere)’ ‘岩石圈[yanshiquan](lithosphere)’ ‘大气圈[daqiquan](atmosphere)’; ‘构造[gouzhao](texture/structure)’ can form the words such as ‘砂质结构[shazhi jiegou](arenaceous texture)’ ‘他形结构[taxing jiegou](xenomorphic texture)’ ‘伴生结构[bansheng jiegou](companion structure)’ and so on.

The third type is the words that are written in Classical Chinese language style such as ‘读作/称为[duzuo/chengwei](be called)’ ‘故/则[gu/ze](therefore/thus)’ ‘即[ji](namely)’ ‘与之[yuzhi](with)’ ‘将[jiang](by/with)’ etc.

From the description comparisons we may find that foreign students can learn the first type of word by various dictionaries and reference books such as English-Chinese Dictionary of Geology, because of their simple meaning and lack of ambiguity. Therefore this type does not belong to the vocabulary scope of this paper.

As for the third type of word, these are also likely to occur in other science and engineering learning areas such as mathematics, chemistry and physics. For instance, ‘There are two variables x and y , for each argument x , the corresponding unique y in the codomain is called the function value at x or the image of x under f . Therefore, it is written as $f(x)$.’ [2] We may see that although this kind of words has the features of universality, this is not a property of geological vocabulary exclusively and thus should be excluded from our research scope.

Only the second type, which can be called ‘Semi-Technological words’ are consistent with our research purpose and teaching standards. They are the basis of GCV. We should take this type as our study objects and analyze their particular features.

2.2. Case Analysis on Semantic Feature and Structures. The Semi-Technological words of GCV have three features which are ‘Universality’, ‘Relative Stability of Meaning’ and ‘Usage Flexibility’.

‘Universality’ means these words are distributed widely and the repetition rate is very high in geological textbooks and references. In order to ensure comprehensiveness and universality when choosing the research words, we use the authoritative geological textbook ‘Introduction to Earth Science (Second Edition)’ [8] as the main corpus resource and use other relatively long time used references such as ‘General Geology’ [3] ‘Geology (Third Edition)’ [9] and so on for frequency count. ‘Relative Stability of Meaning’ is contrasted with ‘Absolute Stability’ of GCV. The words may have a clear geological semantic meaning, or even have the same or similar word form as the ordinary Chinese, but

the specific connotation and extension differ from ordinary Chinese. ‘Usage Flexibility’ indicates that these words can maintain their own stability and also have a strong ability for word-formation combination as well to form different other related geological words. These three features unified together can be used to judge or distinguish whether one word is a Semi-Technological word of GCV or not. We called this ‘Three Principles’.

Let us take ‘作用 [zuoyong] (action/function)’ as a case to analyze.

[Language Dictionary](The subject) has an impact or effect on things (the objects).

[Pharmacology Field] Primary response or pharmacological effect caused by the drugs.

[Engineering Structure Field] A set of forces on the structure or the reasons which cause the structural deformation.

[Geology Field] A kind of Geological works, because of the influence of some energy lead to the continuous changes and formations of the earth's crust material, crustal structure or surface configuration and etc.

From the above, we can see that this word ‘作用 [zuoyong](action/function)’ may express different meanings in different fields. Its feature of ‘Universality’ is quite obvious. At the same time, the teachers in TCFL just teach its basic language meaning ‘has an impact or effect on things’ to foreign students during their one-year Chinese language foundation training. This kind of word meaning is definitely a ‘Relative stability’ explanation. However, whether the ‘action/function’ is a ‘response’ or a ‘force’ or ‘a change or formation’ will not be distinguished or highlighted during teaching. We called this ‘Words-meaning Extension and Expansion’. Hence, after the foreign students enter the professional field, they have to redefine and relearn the connotations and extensions of these kinds of words. It increases the students’ learning task and gives them a lot of headaches.

In addition, these kinds of words also have a strong capacity for word-formation. Like the example we mentioned above, dozens of geological words (and phrases) are derived or associated with ‘作用 [zuoyong](action/function)’. It can be written as a set (G) as followed.

G(作用): { 火山作用 [huoshan zuoyong](volcanic action), 风化作用 [fenghua zuoyong](weathering), 氧化作用 [yanghua zuoyong](oxidation), 搬运作用 [banyun zuoyong](transportation), 成层作用 [chengceng zuoyong](layering), 重结晶作用 [chongjiejin zuoyong](recrystallization), 胶结作用 [jiaojie zuoyong](cementation), 水动力作用 [shuidongli zuoyong](hydrodynamism), 冰劈作用 [bingpi zuoyong](frost bursting), 根劈作用 [genpi zuoyong](root wedging), 水解作用 [shuijie zuoyong](hydrolytic dissociation), 刨蚀作用 [paoshi zuoyong](glacial ploughing), 侧蚀作用 [ce shi zuoyong](lateral erosion), 风蚀作用 [fengshi zuoyong](wind abrasion), 下蚀作用 [xiashizuoyong](incising erosion), 吹蚀作用 [chuishi zuoyong](deflation)... }

In summary, the word ‘作用 (zuoyong, action/function)’ accords with the ‘Three Principles’ which we mentioned above. So it should belong to the Semi-Technological words of GCV.

3. Conception of Synset Construction in Geology Field. The first step is find out the Semi-Technological words of GCV is the first step of our whole research. Except for knowing and mastering the vocabulary correctly, the extent of vocabulary will directly affect the students' language skills development. Not only that, in this century of comprehensive development of Informationization, the traditional vocabulary teaching method of 'Distributed Teaching ' is out of accord with the times. In order to help and improve the efficiency of vocabulary acquisition for foreign students, we must explore the new way that uses the network features of vocabulary to build up the lexical semantic relations among words. That is our conception of synset construction in Geology Field.

3.1. Theories Basis of Building the synset. There are two theoretical bases for building a Lexical Semantic Network. One is the 'Spreading Activation Model of Vocabulary Memory' (Collins and Loftus, 1975) [1] and another is 'Synet-Lexeme Anamorphosis' Method [11] which was advanced by Guozheng Xiao in 2007.

The former is one of the most important theories in cognitive psychology. The model is built on a complex semantic network in which specific memories are distributed in conceptual space with related concepts that are linked by connections. The nodes correspond to concepts and the links correspond to various types of relations between concepts such as same, similar, different or generic relationship. Thus, under the spreading activation model, the semantic relation between words can be further explored. A semantic network can be constructed in terms of the semantic similarities to aid the students to store and memorize words. The conclusion is that the spreading activation can enhance the memory effect of vocabulary and decrease the forgetting rate.

The latter is the further extension and reflection of the network model in language system. This theory was built on the basic cognitive law of peoples' understanding the objects. As we known, cognitive prototypes always exist when people begin to understand objects and they are reflected in the vocabulary system that constitutes the primary words. We called the position of these words 'Basic Lexemes' or 'Word-ontology', and the other words which called 'Non-basic Lexemes' are the variant relied on them. They are formed by different values of concepts in different attributes. Basic Lexemes and Non-basic Lexemes together constitute the Synonym Synset. So the key idea of the 'Synset-Lexeme Anamorphosis' method is that the semantic system of a language is a synonym synset made up of basic lexemes and their variants. Therefore, as long as we are able to identify all basic lexemes and their variants well, find out the differences among them, describe each variant from many aspects by a limited conceptual features set, then we may construct a Lexical Semantic Network like the human brain.

3.2. Basic Lexeme Defined and its Semantic Structure Expression. The meaning of a word is generally described as basic meaning (conceptual meaning) plus the additional meanings. These two kinds of meanings are both composed of one or several semantic primitives. The expression of the primitives can be the single word or the combination or other forms. Such as the word '作用 [zuoyong](action/function)' mentioned in the

previous section, it can be used as the concept of superordination to guide the whole set(G). And in this set, the basic lexeme is ‘has an impact or effect on things’, the other different related ‘作用 (action/function)’ which are derived or associated have different additional meanings. The basic lexeme and its variants together constitute one ‘Generic-specific’ synonymous set (synset).

Confirming the synset is the first step, then the reanalysis and reinterpretation of each member of the semantic group becomes important. The meaning not only contains the basic concept, but also extends the original meaning from a professional point of view, thus it will be more in accord with the field of specialization than before. Finally, we may use a semantic structure to express this as follows.

‘Semantic Structure of a Word = [Basic Lexeme (Word-ontology)] + [Distinctive Semantic Features]’

As shown in Figure 1 below.

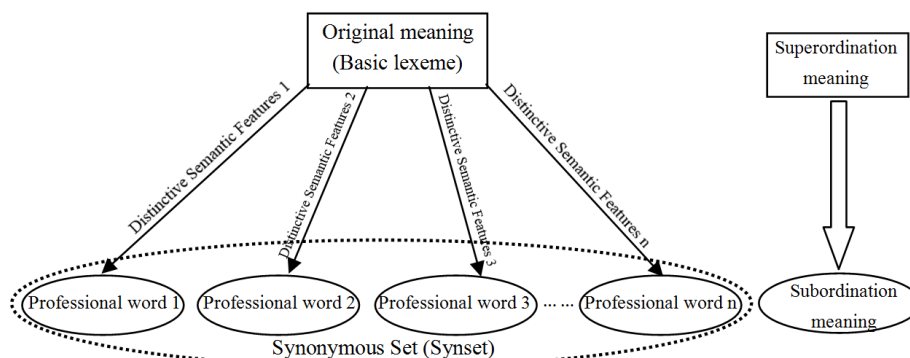


FIGURE 1. SEMANTIC STRUCTURE EXPRESSION OF THE WORD IN SYNSET

For instance, ‘风化作用 [fenghua zuoyong](weathering) belongs to the set (G) of ‘作用 [zuoyong] (action/function)’. It is defined in the professional textbook of ‘Introduction to Earth Science (Second Edition)’ as follows.

‘Because of the temperature changes or some components of atmosphere such as oxygen, water interacts with the minerals in the rocks which, often accompanied by the participation of living things, leads to composition changes and structural damage of the rocks. This effect is called weathering.’

The above is the original definition. However, we may interpret it again after filtering and identifying all the semantic elements in the following way.

‘The influence/effect that composition changes and structural damage of the rocks which given by the action or function of the atmosphere temperature changes or some components interaction is called weathering.’

Comparing with the basic lexeme of ‘作用 [zuoyong](action/function)’, we can find out that the main semantic elements are still the[+the subject],[+the things (objects)] and [+have an impact or effect] in this interpretation. But the difference is that [+the subject]

indicates [+atmosphere temperature or some components], [+the things (objects)] indicates [+the rocks] and [+have an impact or effect] indicates [+composition changes and structure damages] clearly and definitely. These are the most obvious distinctive features linking ‘风化作用 [fenghua zuoyong] (weathering)’ and ‘作用 [zuoyong] (action/function)’.

In this way, it can help students establish more effective contact channels between the new words and the old ones, form a vocabulary memory network with a wide activation mechanism so that each word may find its own position in this network. This way can also help to change the present situation that the vocabulary teaching is mixed and disordered in TCFL. Many Chinese teachers just explain the word which they meet one by one without any systematization.

Certainly, the Lexical Semantic Network in the Geology Field which we want to construct is not just a word connection network but a system composed of concepts and their variants. This system should not only have higher ability in differentiating one semantic feature from another, but also possess more accurate knowledge particles, thus recreate the true appearance of ‘object-concept-word’ in the human brain and help people to enlarge their entire vocabulary efficiently and realize intelligent learning by computer. Therefore, the form of the concept can be the word including the morpheme or the phrase. The following set (H) will help to clarify what this means as an example.

H(蚀):{剥蚀[boshi](denudation), 熔蚀[rongshi](resorption), 磨蚀[moshi](ablation), 溶蚀[rongshi] (dissolution), 腐蚀[fushi] (corrosion), 顶蚀[dingshi] (stoping), 冲蚀[chongshi] (erosion), 海蚀[haishi] (marine erosion), 风蚀[fengshi] (wind erosion), 浪蚀[langshi](wave erosion), 冰蚀[bingshi](ice erosion), 河蚀[heshi](river erosion)...}

Here the set (H) is a synonymous set (synset). In this set, morpheme ‘蚀 [shi] (erode/erosion)’ is the concept of superordination and its semantic structure can be described as ‘[Action/function+ (Other) surface materials+Destroy/damage+ External force (itself)]. Semantic structure is read from right to the left, namely the synset meaning of this set (H) is ‘the action/function which external force destroys or do damage to other surface materials.’

After confirming the synset meaning, we may discover that this synset can be divided into at least two or more sub-synset according to the distinctive features differences among the members as show below.

H1:{冲蚀[chongshi](erosion), 磨蚀[moshi](ablation), 溶蚀[rongshi](dissolution)...}

H2:{海蚀[haishi](marine erosion), 风蚀[fengshi](wind erosion), 冰蚀[bingshi](ice erosion)...}

Hn:

The set (H1) is the synset of synonymy relations while the set (H2) is the generic relation. By vertical comparison of H1, H2 and H and horizontal contrast between H1 and H2, we found that although H1 and H2 have the same synset meaning, H1 should add one new basic lexeme[+Be dissociate from/peel off] and H2 need to add another new basic lexeme[+Sender of power /force] as well. Consequently, the sub-synset meaning of H1 and H2 can be described by the semantic structures as follows.

H1 {chongshi,moshi,rongshi}:[Action/function+Be dissociated from/peel off+(Other)

surficial materials + Destroy/damage+External force(itself)]

H2{haishi,bingshi,fengshi}:[Action/function+(Other)surface materials+Destory/damage +Sender of power/force

Based on the sub-synset meaning, we only need to take related pragmatic meaning (primitives) values in order to distinguish the different distinctive features of each word in sub-synset, thus the semantic structure expression of each word can also be completed in this way. Take several words of H1 and H2 as examples as follows.

H1 :{ 冲蚀 (erosion), 磨蚀 (ablation), 溶蚀 (dissolution)...}

{ 冲蚀 [chongshi](erosion)}:[Action/function+Be dissociated from/peel off+(Other) surficial materials + Destroy/damage+External force(itself)+shock way]

{ 磨蚀 [moshi](ablation)}:[Action/function+Be dissociated from/peel off+(Other) surficial materials + Destroy/damage+External force(itself)+rub way]

{ 溶蚀[rongshi](dissolution)}:[Action/function+Be dissociated from/peel off+ (Other) surficial materials + Destroy/damage+External force (itself) +chemical dissolution way]

H2 :{海蚀 (marine erosion), 风蚀 (wind erosion), 冰蚀 (ice erosion)...}

{ 海蚀 [haishi](marine erosion)}:[Action/function+(Other)surface materials + Destory/damage+Sender of power (Sea)]

{ 风蚀 [fengshi] (wind erosion)}: [Action/function+ (Other) surface materials + Destory/damage+Sender of power (wind)]

{ 冰蚀 [bingshi] (ice erosion)}:[Action/function+(Other) surface materials + Destory/damage+Sender of power (ice)]

In this way, the original complicated vocabulary of geological textbooks and references can establish simple and clear semantic relations between each, and words can be distinguished from each other through the distinctive semantic features, especially the synonym and semantic-overlapped words. So students can easily continue to learn the basic scientific Chinese vocabulary and extend to other new related vocabularies more quickly and efficiently on the basis of this lexical semantic network.

At present, one of the author's scientific projects 'Knowledge Database Construction of GCV Teaching for Foreign Students' is quantitative research. This project started from 2013 and there will be 400 to 600 vocabularies included and researched in this database by 2015. It has currently reached nearly 200 and the vocabulary capacity may be gradually expanded in the subsequent development.

4. Conclusion. There is scope for growth in GCV research. This paper suggested that we must pay more attention to CSP than to CGP and the research achievements can develop and enrich the theories of vocabulary teaching in TCFL. We discussed and explored the description of the lexical semantic structure and its synset construction rules according to the 'Synset-Lexeme Anamorphosis' method. This method may change the lexical semantic network in students' mind into a visible mode of teaching or learning such as the vocabulary courseware or database designed using the network organization of information. This will not only help foreign students to study the characteristics and rules of each Chinese vocabulary, but also help the students to learn the related or similar professional

vocabulary by analogy .It is a better way to improve their cognitive academic language proficiency compared with the traditional ‘Distributed Teaching’ method.

Acknowledgment. This work is partially supported by a grant from Teaching Research Project of China University of Geosciences (Grant No.2013B30), Educational Commission of Hubei Province (Grant No.13g048) and the National Social Science Fund(Grant No.14BYY121).

REFERENCES

- [1] Collins, A.M. and E.F.Loftus, A spreading-activation theory of semantic processing, *Psychological Review*, Washington.D.C: American Psychological Association, pp407-428, 1975.
- [2] Du, H.W., Language Features of Chinese Science and technology style, *Language Teaching and Linguistic studies, Beijing*, vol.7, pp92, 1987.
- [3] Huang, D.H., General Geology, *Higher Education Press, Beijing*, 2004.
- [4] Huang .Z.Y., HNC (Hierarchical Network of Concepts)Theory,*Qinghua University Express,Beijing*, 1998.
- [5] Li, Q., Teaching Chinese for Specific Purposes and Its Textbook Compilation, *Applied Linguistics,Beijing*,vol.3,pp111, 2011.
- [6] Su, D.J., Vocabulary and Vocabulary Teaching in TCFL, *The commercial Press, Beijing* , 2006.
- [7] Shan,Y.M. and An.R., Analysis of Chinese courses for Specific Purpose——take ‘Technological Chinese’ course as an example, *Journal of Southwest University for Nationalities,Sichuan*,vol.8, pp261,2009.
- [8] Tao, S.L. and Wan,T.F.,Introduction to Earth Science (Second Edition),*Geological Publishing House,Beijing*, 2010.
- [9] Xu,J.H.,Xie,Y.L.,Li,J.P and Li,K.Q.,Geology(Third Edition), *Metallurgical Industry Press,Beijing*, 2000
- [10] Xie,W.W.,Huang,T.L.,Zhou,R.Y.and Wang,S.L., General Geology,*Geological Publishing House,Beijing*
- [11] 2007.
- [12] Xiao,G.Z., Description of the Lexical Meaning Structure of the Verb ‘da’ and the Construction of its Synset——A study on ‘Synset-Lexeme Anamorphosis’ Shared by Human and Computer:Chinese Computing Technologies and Related Linguistic Issues.In:Proceedings of the 7th International Conference on Chinese Computing, *Electronic Industry Press,Beijing*, pp.4-7, 2007.
- [13] Zheng, D.O., Desiderata for Promoting Chinese Language World-Wide,*Chinese Language Learning,Jilin*,vol.3, pp91, 2008.