A Study on the Language Bias of Neologism on the Internet

Guoxin Miao
College of the Humanities, Jilin University, Changchun, Jilin, China
Foreign Language Teaching and Research Institute, Jilin University of Finance and Economics, Changchun, Jilin, China

Ping Yu
College of the Humanities, Jilin University, Changchun, Jilin, China

Lin Lv*
School of Media at Jilin Engineering Normal University, Changchun, Jilin, China
*Corresponding author

Abstract
The current era of highly developed computer technology, the Internet has become the main way for people to communicate. While the Internet is widely used, various new Internet words have also appeared and developed. With the diversification of Internet language, this has also attracted many scholars’ interest in network language, so as to carry on the research to its related content. In the meantime, it not only gets the research conclusion, but also promotes the development of new Internet words. The neologism is the most vivid and vital part of the network language system. This paper is mainly explained from the perspective of constructive grammar, qualitative and quantitative methods are used to analyze the new Internet words in recent years. Internet neologism is not only widely spread on the Internet, and it’s very common in everyday communication. This paper makes a relatively comprehensive analysis of the structure of the collected neologisms, reveals the working mechanism and linguistic phenomenon of neologism on different levels. During the generation of new words on the Internet, the construction grammar mechanism has been used in a variety of new words, vivid display of ordinary people’s concern and concern in social life, in a simple and clear way to present the ordinary people.

Key words: Network Neologism, Construction Grammar, Phenomenon Study

1. Introduction
As one of the greatest inventions of the 20th century, the Internet has had an incredible impact on people’s lives. It is not only an unprecedented technological revolution, but also a language revolution that brings about major changes in the use of language. According to the 42nd “China Internet Development Statistics Report” released in July 2018, as of June 2018, the number of Chinese Internet users was 802 million. In the first half of the year, the number of new Internet users increased by 386,682,800, an increase of 3.8% from the end of 2017. The Internet penetration rate reached 57.7%. As of June 2018, China Mobile Internet users reached 788 million, and the number of new mobile Internet users reached 35.09 million in the first half of the year. Compared with the end of 2017, the growth rate was 4.7%, and the proportion of netizens using mobile phones reached 98.3%.

In today’s rapid development of the Internet, we must have a deeper understanding of network culture and network language, not only to define what is a network language, but also to clarify the connotation and extension of the network language in order to better analyze and explore. There is always debate about what is a new word in the network. Some scholars have summarized it into a network-related language form. Broadly speaking, network languages can be divided into three categories: first, network-related terminology. Such as: a variety of computer-related network language vocabulary, such as JAVA, PHP, etc.; the second is to reflect new culture, new economy and new words related to the network, such as grassroots industry, leading enterprises, post-8090, iron rice bowl, Wenshan Huihai, etc.; third is the network words commonly used by netizens in online chat rooms and Weibo. And some network symbols, such as rotten female, handsome pot, scorpion, official Xuan, prawn, bar fine and so on. The narrowly defined network language refers only to the third type. This article mainly discusses the third type of network words and symbols, which we call network new words.

Many scholars have done in-depth research on the understanding of construction grammar. They have contacted the development of social language from many other aspects and made great contributions to the research and development of human language. Hoffmann argues that construction grammar is cognitive structuralism, and he explores the interaction between tectonic networks and the diachronic evolution of processing in English Hoffmann, (2017). He introduced a constructivist approach based on usage that enabled researchers to reinterpret the blank classical structuralist concepts in the system as gaps in the psychological structure network. This type of cognitive structuralist analysis explains the existence of less iconic C2C1 structures in OE (and no more iconic C1C2 structures), the origin of C1C2 structures at the end of the OE phase,
and the thus predicted processing effects. Karl discusses the semantic analysis of several words that encode emotions in Croatian. The constructive grammar approach was used and some of its basic theoretical principles were fully demonstrated in terms of NooJ’s computational power Karl et al., (2018). Using the example of the noun strah ‘heep’, the purpose of the study is to point out the possibility of annotating specific structural meanings in NooJ, such as the different connotations of the selected vocabulary, the general use of the structure, their relationship to other structures, and their specific meanings. Significant features (pragmatic features, as well as semantic and morphological syntax) that reflect different language and cognitive phenomena in the language. Mirjam explores the connection between the Slavic language and the theoretical principles of constructive grammar, a cognitive and functionally oriented approach to linguistic analysis Fried, (2017). The advantages of traditional Slavic linguistics are in particular their concern for diachronic problems, lexical semantics and morphological issues. Architectural analysis provides a solid theoretical foundation for these traditional areas, and it also raises concerns about phenomena and problems that Slavic linguists pay less attention to. This involves various syntactic patterns, but also relates to the discourse organization and grammatical device fields that serve specific utterance functions, whether it is the nature of pragmatic particles, the specific clause structure, and the expression of subjective cognitive positions. Interest is also the origin and evolution of this device. In Slavic linguistics, this field is generally unaffected, but it represents a lot of interesting data and is directly related to the current theoretical issues at the forefront. Dunn proposed an algorithm for learning grammar of language structure from a large corpus Dunn, (2017). This grammatical inductive algorithm has two goals: first, to show that the construction grammar is learnable, without a highly specified innate structure; and second, to develop a model in which units constitute or do not constitute a construct in a given data set. The basic task of constructing grammatical induction is to identify the smallest set of constructs that represent the language in question and to have the greatest descriptive adequacy. These structures must (1) summarize unspecified number of cells, and (2) contain mixed representation levels internally (project-specific and schema-represented representations), and (3) allow unfilled and partially filled Time slot. Additionally, these structures may (4) contain recursive structures within a given time slot that need to be reduced to produce sufficient schematic representations. Yuanyuan systematically compared the English and Chinese sentences with a constructional grammar from a new perspective Yuanyuan, (2017). Torre studied the structure of Italian idioms from a perspective that combines the insights of constructivism and dynamic systems approaches with language Torre, (2015). First, he analyzed the appearance of structural samples downloaded from the TenTen corpus based on experience, and conducted queries through the online corpus query system Sketch Engine. Based on the trends observed in the analysis, he observed that the patterns of stability and idiom changes in use can be satisfactorily considered in dynamic system terminology. Then, the use of the conventional structure is considered to be governed by the principle of causality, where the attractor state limits the possible use of the structure, but at the same time a large number of idioms shape the attractor into a continuous, non-linear. In addition to idioms, he proposes a similar mechanism to regulate the function of the entire language system, which is consistent with the constructivist’s view of language as a network of interconnected units. Kelli studied the application of construction grammar in the core English form, considering the examples of English cognition and root mode verbs Slijmp, (2010). The modal expression in English provides many ambiguous explanations and unique realizations, because there are conversational situations in which they can be spoken. The form of English has proven to be a serious background and in some cases depends on the subject. Spuy considers constructive grammar to be a language framework that assumes that all grammars are based on a relatively simple single entity, construct, defined as a traditional mapping between form and meaning van der and Andrew, (2017). Tectonic morphology is a sub-theory of constructive grammar that deals with word-level domains. He introduced how the structural form explains all the types of inflection. It argues that constructive morphology is a type of morphological theory, which Stump (2001) calls “inference-implementation,” which he considers to be the only theory that explains all known morphological facts. Fischer believes that constructive grammar is a language concept deeply rooted in cognitive linguistics Fischer, (2015). Its main idea is that linguistic knowledge can be described in detail as a structured list of “form-meaning” pairs. Although this view has proven to be very useful for describing a large number of linguistic phenomena, it has a great impact on the development of the construction. Brems studied the construction grammar and its application in English Brems, (2015). Dunn proposed an algorithm for learning the grammar of language structure from a large corpus Dunn, (2017). This grammatical inductive algorithm has two goals: first, to show that the construction grammar is learnable, without a highly specified innate structure; and second, to develop a model in which units constitute or do not constitute a construct in a given data set. The basic task of constructing grammatical induction is to identify the smallest set of constructs that represent the language in question and to have the greatest descriptive adequacy. Shefer examined the hypothesis that the Hebrew pattern conforms to the structure of the construction grammar hypothesis Shefer, (2017). The pragmatic function of negation is discussed, but the significance of this model is that it exhibits a structural feature: it consists of a large number of expressions with extensive internal variation, a system’s internal structure, productivity, traits, and rules. In addition, he analyzed the location of this structure in a structured network that
represents the language of the speaker. This is achieved through a classification hierarchy that embodies the inheritance relationship between construction and other more general language rules. The significance of adopting this view is that it allows us to think of a language as a network that contains different levels of schematic construction.

The new network words discussed in this article are new words and phrases that are produced and popularized on the Internet. According to different standards, network new words can be divided into many categories. Someone has classified them according to the stability of semantics and divided them into new words and new words. The emergence of new words in speech must rely on a certain context. If you leave this context, its semantics will point to some other meaning or people may not understand the meaning it refers to, such as “pig foot (protagonist)”, “Table sauce purple (not like this)”, “7456 (gas mad me)”, “diving (in the forum only see posts do not reply)”, “white bones (white collar + backbone + elite)”, “porridge (like)” And the network address “pro”, etc., this type of network new words can only be used in the specific communication environment of the network. If you leave this context, these new words will make people listen to the clouds.

2. Proposed Method

2.1. Definition of Construction

Construction Grammar (CG) is a grammatical theory that gradually emerged in the late 1980s. It applies to research methods and genres of almost the entire language category. In the early days of the rise of construction grammar, people’s understanding of the word construction is not very thorough. In fact, the definition of construction is only “construction”, but this representative is to construct two or more words. Later, people I feel that the word “construction” is more vivid. Traditional grammar works borrow the term to describe the language structure. This usage is found in many English grammar works, such as Quirk and Chomsky, which are used in their writings. The construction in these works mainly refers to the specific structure. The meaning of the word retains an important aspect of the original meaning; the language structure it refers to must be a combination of two or more components. However, in most of these works, it is not a special term. From the example, it refers to the syntactic structure (body), and generally does not refer to structural units such as words and morphemes. In the current most popular Goldberg view, the two core ideas of construction grammar are clarified: construction is the pairing of form and meaning and the meaning of the whole construction cannot be predicted from the constituents.

Before the construction grammar was officially born as a theoretical method, there were already some ideas in the linguistic works. Constructivism treats construction as a symbol, and not only has its form but its own meaning. In ancient Greek grammar theory, language is the “symbolic reaction of psychological process”, and language units are considered to be form-meaning correspondence. Currently, the “construction” referred to by the construction grammar is most popular with the definition of Goldberg: C is a CONSTRUCTION if def C is a form-meaning pair <Fi, Si> such that some aspects of Fi or some aspects of Si is not strictly predictable from C’s components parts or from other previously established constructions.

2.2. Fractional Grammar

The real rise of construction grammar was in the 1980s, and the construction of construction grammar was based on the reflection of the theory of transformational generative language proposed by Chomsky. The theoretical system of construction grammar is not only based on a specific theory, but a combination of various grammatical theory models. At present, the main theories of the construction grammar are: the construction grammar represented by Goldberg, the cognitive grammar represented by Langacker, the construction grammar represented by Fillmore & Kay, and the experiential structure represented by Bergen & Chang. Syntax and the radical construct syntax represented by Croft. (Figure 1 from the network, www.baidu.com).

2.3. Construction View of Construction Grammar

The core idea of constructing grammar is that the generation of a grammatical view consists of two parts: vocabulary and rules. Constructive grammar maintains a rebellious attitude towards generative grammar. In the theoretical system of constructing grammar, the meaning of a phrase is not explained by rules. The meaning of many sentences cannot be obtained by the semantic synthesis of sentence components, so Goldberg believes that the acquisition of sentence meaning is the result of the combination of structural meaning and meaning of each component.
2.4. Database

This paper mainly studies the data in the annual language living status report issued by the National Language Resource Monitoring and Research Center as a database. The National Language Resource Monitoring and Research Center was created by the Department of Language and Information Management of the Ministry of Education. As a series of projects, the National Language Resource Monitoring and Research was included in the Ministry of Education’s New Education Revitalization Action Plan. The sub-centers that have been built so far include: print media language, online media language, educational textbook language, and audio media language. Five sub-centers of overseas Chinese communities. The establishment of the National Language Resource Monitoring and Research Center aims to provide more quantitative understanding of Chinese language and national conditions, use real-time monitoring of modern Chinese language applications, conduct dynamic analysis, statistics and research, and establish real-time monitoring and normative guidance of social language life. Long-term mechanism, strengthens the dynamic management and scientific utilization of national language resources, promote the healthy development of social language life, and provide basic support for the development of national language information processing. According to the Ministry of Education, the “Top Ten Network Terms of Each Year” is based on the National Language Resource Monitoring Corpus (Network Media Part), which is extracted by means of “Intelligent Information Processing Technology, supplemented by artificial fine-tuning”. The monitoring corpus extracts a large number of annual language resources in different media formats, such as representative online forums, blogs, microblogs, and online news. The corpus is nearly 6 billion words, representing the language lifestyle of the Chinese Internet. Internet language is the crystallization of the wisdom of netizens. It is humorous and distinctive. Its identification and usage are extremely high. It is one of the top ten online terms of the current network life, or it comes from movies, TV variety shows, or From the live broadcast platform, all vividly portrayed the concern and joys and sorrows of netizens in the past year, reflecting the language and culture of netizens.

On the basis of the construction of language resources, through the development of language resources, the investigation of language living conditions, and the dynamic monitoring of media language, it provides a data foundation for the formulation of national language policies and provides services for language life, language teaching and language information processing. Among the 612 new words extracted from the National Language Resource Monitoring Corpus in 2011, the three words accounted for the most, accounting for 51.68%; the second was the four-character word, accounting for 21.04%; the second word accounted for 15.66%. The proportion of three words continues to dominate and has been associated with the use of popular formats in recent years.

3. Experiments

3.1. Network New Word Classification

New words in the language refer to words and sentences that are clearly defined, used frequently in the
network, recognized by netizens, and sometimes used by other netizens, such as “XX on the tip of the tongue”. “The most dazzling democratic style”, “Gao Fu Shuai”, “Bai Fumei” and so on. This paper believes that this classification standard has certain rationality, but it also has its shortcomings, and there are problems of crossover and unclear standards. Therefore, combined with the above classification, according to the characteristics of new words in the network, this paper divides the new network words studied into two categories.

The first category: computer network terminology. Computer network terminology is a number of professional terms related to some network technologies and network usage. It is the most active and fundamental component of the network. If you use a computer to access the Internet, you must know some computer network terminology, otherwise it is difficult to operate. It is mainly used to refer to network technology results, computer operating systems or words of computer operations objects during the Internet, such as “communication protocols, communication devices, servers, clients, peers, network devices, local area networks, synchronous transmission, routing, hubs., switches, firewalls, NTFS permissions, high-speed Ethernet, Fast Ethernet, etc. There is also some networking terms that are used across regions. The meaning of computer network terminology is relatively difficult to understand and belongs to the field of technology industry. Some computer network terminology has been broadened and used in other industries. The use of such words is relatively easy to accept, such as “software, hardware, accessories, backup”, and is also applied to economic, social life and other fields. Such as information highways, viruses, account numbers, Da Vinci’s tears and so on. It has an implication in a particular context. There is also some network computer terminology developed from everyday idioms or other terms, such as “paths, conflicts, convergence points, congestion, and backbone networks”.

The second category: new terms for the network. The new term network refers to the new name and new title of new and widely used netizens in virtual communication. This type of word reflects new politics, new economy, new culture, new society, and they have strong vitality. The new terminology of the network mainly includes two categories: one is the new term about modern network culture, and it refers to new things and new images that are constantly emerging in the information age, such as “post-PC era, email, bulletin board, e-mail, email, disconnection”. The other is the word about popular culture, referring to some new things and new phenomena that are close to the lives of netizens, such as: “Scorpion, Yuan Fang, how do you see this, jiong, to force, pear, soy sauce, crab, god Horses are clouds, people don’t want too, CNN, cottage, bamboo, cups, brewing purple, friends, crush, Taobao, house slaves, card slaves, car slaves, Benbens, etc. The new term network is a true reflection of the life of modern netizens. It has the characteristics of fresh activity and individuality. These languages are more realistic, alive, and in line with the present life and the present mood.

3.2. New Word Development

The National Language Resource Monitoring Corpus Statistics show that the media vocabulary usage survey was conducted on the 2012 National Language Resource Monitoring Corpus with a total of 1.2 million texts and 1 billion words. A total of 10,807 words and 2,382,565 words were found in all corpora, and the distribution of Chinese characters and words was strong.

In 2012, 585 new words were extracted from the National Language Resource Monitoring Corpus, still based on the second, third and fourth words. Among them, 574 new words all composed of Chinese characters, accounting for 98.12%. There are Chinese characters plus letters, such as “H family, aircraft carrier type”, as well as Chinese characters plus Arabic numerals, such as “1314 marriage tide, 90-year-old body, the most beautiful after 90”. With the changes in society and people’s perceptions and concerns, the number of words that were particularly popular in 2012 has decreased. The pattern of words that continue to be popular is “~ body, micro ~”. The word model that best reflects the characteristics of social life in 2012 is “the most beautiful ~, Chinese style ~, room ~, doomsday ~”.

According to years of investigation, the number of words used in the 1 billion word corpus is stable at around 10,000, and the number of words is stable at around 2.3 million. Among them, the number of words with coverage rate of 80%, 90%, and 99% is 590, 960, and 2400, respectively, and the coverage rate of 80%, 90%, and 95% is 4800, 14000, and 30,000. They form a relatively stable part of the annual media language. These data provide the basis for the grading and quantification of words, which can be used for dictionary editing, writing of native language and Chinese as a foreign language. Words with a large change in frequency of use reflect the hot spots of social life and media attention.

In 2013, the National Language Resource Monitoring and Research Center extracted 364 new words based on 1.25 million texts and 1.2 billion corpora. It reflects the characteristics of social life and language use in China and the world in 2013. Internet language: “Chinese aunt, high-end atmosphere, where my father went, my friends are shocked, waiting for my long hair and waist, Xi Dapu Ben, female man, local tyrant (gold), stalling big things, rising Network terms such as “posture” represent the distinctive features of the use of network language in 2013.

The following two tables are the top ten online terms for each year of 2011-2018:
Table 1. The top ten hot words of the network in 2015-2018

<table>
<thead>
<tr>
<th>Years</th>
<th>2018</th>
<th>2017</th>
<th>2016</th>
<th>Important things</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Koi</td>
<td>Da call</td>
<td>The power of the wild</td>
<td>say three times</td>
</tr>
<tr>
<td>2</td>
<td>Bar fine</td>
<td>Chat</td>
<td>Friendship boat</td>
<td>The world is so big, I want to see</td>
</tr>
<tr>
<td>3</td>
<td>Skr</td>
<td>Does your conscience not hurt?</td>
<td>Set a small goal</td>
<td>People in your city will really play</td>
</tr>
<tr>
<td>4</td>
<td>Buddhist system</td>
<td>Surprise, not unexpected</td>
<td>Eating melons</td>
<td>National support</td>
</tr>
<tr>
<td>5</td>
<td>Confirmed eyes</td>
<td>Pippi shrimp, let’s go</td>
<td>Ge You</td>
<td>Obviously, you can rely on your face to eat, but you have to rely on talent.</td>
</tr>
</tbody>
</table>

Table 2. The top ten hot words of the network in 2011-2014

<table>
<thead>
<tr>
<th>Years</th>
<th>2014</th>
<th>2013</th>
<th>2012</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Does your family know?</td>
<td>Local tyrant, let’s be friends</td>
<td>Positive energy</td>
<td>Hold zhu</td>
</tr>
<tr>
<td>2</td>
<td>Cherish what you have at the moment</td>
<td>Me and my little friends are shocked</td>
<td>Are you happy?</td>
<td>Do you have</td>
</tr>
<tr>
<td>3</td>
<td>Sprouting</td>
<td>Waiting for me to grow my hair and waist, can you please me?</td>
<td>Yuan Fang, how do you see?</td>
<td>Dear</td>
</tr>
<tr>
<td>4</td>
<td>Now the whole person is not good.</td>
<td>Why did you give up treatment?</td>
<td>Diao si</td>
<td>I don’t believe in love anymore.</td>
</tr>
<tr>
<td>5</td>
<td>I have less reading, don’t lie to me.</td>
<td>I think I would not love anymore</td>
<td>XX Style</td>
<td>What is swollen?</td>
</tr>
</tbody>
</table>

4. Discussion

4.1. Internet Development

As the number of Internet users continues to increase, new online vocabulary, creative vocabulary and expressions are rapidly emerging and growing on the Internet. Internet expression can be seen almost everywhere in our lives: journalists are keen to use new online vocabulary as headline news; young people are keen on new online words; students even apply these newly created expressions to their work; The use of new words in advertising is also not uncommon. In the process of surfing the Internet, young netizens often have some small innovations in their expression. Initially it may only be used in a small range, but due to the spatial extent and speed of the network, these inspiring expressions are gradually spreading on the Internet and becoming the common language of the Internet community. The emergence of new words on the Internet comes from the creation of netizens, and the other part is the process of communication between netizens. Some words are used more frequently and are slowly fixed to become new words in the network.

![Figure 2. Scale of Chinese netizens and Internet pervasion rate](image)
4.2. Data and New Word Analysis

In 2016, the total number of Chinese characters used by the media was 975 million, and the number of words was 12041, covering 585 high-frequency Chinese characters, accounting for 80% of the entire corpus. Covering 90% of the entire corpus, 941 high-frequency Chinese characters, covering 2378 high-frequency Chinese characters, accounting for 99% of the entire corpus, 4335 Chinese characters, frequency less than 10, only 1517 Chinese characters used once. In 2016, the media used a total of 575 million words and more than 2.45 million words. There are 32,829 words with a coverage rate of 95% and 240,916 words with a coverage rate of 99%. In the annual Chinese inventory activities: “Regulations and small goals” were elected as domestic and domestic words respectively; “Change, One Belt, One Road” was elected as the international and international words respectively; “two studies and one done, frozen production, expression packs, and floods” Power, Alpha Go, Internet Movies, Mobike Bicycles, Shanzhai Community, Eat Melon People, and Friendship Gates were elected as the Top Ten New Words of Chinese Media in the Year; “Long March Spirit, Two Schools and One Doing. Hangzhou G20 Summit, Nanhai, Rio The Olympic Games, Brexit, US elections, confession, Tiangong No. 2, Alpha Go are elected as the top ten buzzwords of the year; “the power of the wild, the boat of friendship, the setting of a small goal, the eating of the masses, the Ge You lying, the spicy eyes, the whole It is a routine, blue-skinned mushroom, old driver, and my brother.”

In 2017, the total number of Chinese characters used by the media was 994 million, the number of characters was 11583, and the monitoring corpus was 99% of the high-frequency Chinese characters 2377. There are 4004 Chinese characters, the frequency is lower than 10, and only 1452 Chinese characters are used at a time. In 2017, the media used a total of 575 million words, more than 2.56 million words, 28, 1024 words, with a coverage rate of 99%. The following figure shows the histogram of the total number of characters and the number of words used by the media in 2016-2017. It can be seen from the figure that compared with previous years, the total number of characters and the number of characters in the Chinese characters are generally stable.

![Bar chart showing comparison of the total number of characters and number of characters used by the media in 2016-2017](image)

Figure 3. Comparison of the total number of characters and the number of characters used by the media in 2016-2017

The following table shows the number of words in the top ten hot words of the network in the 2011-2018 years:

<table>
<thead>
<tr>
<th>Word count</th>
<th>One</th>
<th>Two</th>
<th>Three</th>
<th>Tour</th>
<th>Five or more</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number</td>
<td>4</td>
<td>13</td>
<td>13</td>
<td>10</td>
<td>40</td>
</tr>
</tbody>
</table>
As can be seen from the histogram in the above figure, the number of words in the new words of the network varies, the number of words is 5 and 5 or more, accounting for 50%, the proportion is the most, and the word proportion of one word is the least, only 5%. Among the 612 new words extracted from the National Language Resource Monitoring Corpus in 2011, the three words were the most, accounting for 51.68%; the second was the word of 21.04%; the second word was 15.66%. The proportion of three-word words continues to dominate, and it has been related to the use of popular formats in recent years.

In this section, the structure of the lexical level is analyzed by referring to “xue X” (“Study X”) as an example. “xue X” consists of two parts: X is an autonomous component, and “xue” is a dependent component of the structure. On the one hand, in terms of non-context, the formation of the semantic, semantic and syntactic features of “xue X” is not self-generated, but under the influence of multiple coercions and the inheritance of its dominant structure. At each level, the “xue X”, as a Chinese phonetic structure (PC), can be divided into sub-structures, such as monosyllabic construction, two-syllable construction, three-tone construction and other multi-syllable constructions. These substructures compete with each other during multiple enforcement and inheritance. The two-syllable structure wins and is chosen as a force to form new expressions because its information is compatible with new information. Therefore, the two-syllable structure is the dominant structure of the speech level “xue X”, and all Xs in the collected data are monosyllabic. The derivation of “xue X” is influenced by the structure of Chinese two-syllables. Under the force of the Chinese two-syllable structure, any three-syllable words will be compressed into two-syllable words in “xue X”.

Then, we quote an example of “wo keneng x de shi jia y” (“I may X is false Y”) to illustrate the cognitive mechanism at the syntactic level. On the one hand, in terms of non-context, the formation of the phonetic, semantic and syntactic features of “wo keneng x le jia y” is not self-generated, but under the influence of multiple coercion. First, from the perspective of syllables, X and “de” together constitute the constructed predicate, and according to the data we collect, most verbs can be used as their predicates “review”, “eat”, “drink”, for example. “enter”, “listen”, “use”, “go” and so on. They are monosyllabic and double syllables. However, the predictive “X de” is imposed by the Chinese two-syllable structure. In most cases, the “X” is monosyllabic. Second, from a syntactic point of view, the entire sentence is a subject-predicate structure and a positive sentence structure, usually used as a small independent clause. Constructed predicates and objects are not fixed. “jia Y” is a modifier and the adjective “jia” is used as a modifier. Therefore, this structure can be rewritten as “wo keneng V le jia N”.

In this chapter, a relatively comprehensive analysis of the collected network new word structure is carried out. The purpose is to reveal the working mechanism of the network new word structure at different levels, and explore the multiple coercives in these structures. The structure of the lexical level is usually enforced by syllables, grammar, metaphor and context. If there is no syllable forcing, syntactic coercion, semantic coercion and pragmatic coercive efforts, the syntactic construction mechanism is difficult to achieve. However, the construction mechanism at the discourse level is much more complicated than the construction mechanism at the other two levels. Forced to occur between registers and genres, in registers, and in genres. In addition, the pragmatic effects of its use are described in the process of discovering its working mechanism.

4. Conclusions

1. The total number of words used in Chinese characters and the number of words are generally stable, and the development of language and culture is stable and progressing. The rapid development and popularity of
new online words reflects the characteristics of social life, and clearly reflects the richness of language and culture.

(2) A relatively comprehensive analysis of the collected network new word structure, revealing the working mechanism of the network new word structure at different levels, and exploring the multiple coercives in these structures.

(3) The structure of the lexical level is usually enforced by syllables, grammar, metaphor and context. If there is no syllable forcing, syntactic coercion, semantic coercion and pragmatic coercive efforts, the syntactic construction mechanism is difficult to achieve.

(4) Comparative studies have found that there is a large imbalance in the productivity of CXG construction at all levels. The number of CXGs at each level varies. There are 100 CXGs at the lexical level, 158 CXGs at the discourse level, and only 20 CXGs at the syntactic level. Development imbalances exist not only at different levels of language, but also at the same level. The production capacity of different CXGs at the same level is also different.

Acknowledgments
This work was supported by 2019 Social Science Foundation Project of Jilin province “Research on English Learning Ability Cultivation of College Students in Jilin Province” (No.2019B173) and “Investigation and Countermeasure Research on Language Service Problems of Enterprises in Jilin Province Under the Background of Going Abroad” (No.2019B186).

References