The Senses of LUO in Internet Text

Quan Zhang, Yi Yuan, Xiangfeng Wei, Jianming Miao
Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190
Zhq@mail.ioa.ac.cn

Abstract
Hanzi (Chinese character) is unique in Chinese, and “a Hanzi’s meaning is primitive, and a word’s meaning usually is combined” is the nature of modern Chinese. Based on this nature, we investigate the paraphrase of the words including the Hanzi LUO (裸) from the view that Hanzi dynamically combines to be a word. We collect out-of-vocabulary words from the Internet text, and compare them with the dictionary. We find that the senses of the LUO (裸) in out-of-vocabulary words are developed, i.e. the senses in dictionary are “nude, especially for human body, and bare, without wrapping” and the senses in out-of-vocabulary words add the meaning of “without any additional content, and abandon all the thing which a person had hold (such as property, social position, and so on)”. Furthermore, we analyze the senses of LUO (裸), and find that the nuclear meaning of LUO (裸) does not changed fundamentally, and the development is the modified object, from state of individual to the state of event, and the content of object also becomes abstract from concrete. Meanwhile, we discuss the meaning development reason too. The presented result will help for the Chinese out-of-vocabulary words processing which include LUO (裸).

Key Words
Out-of-vocabulary Words; Dynamically combined words; Lexical semantics; Sense analysis; Hanzi sense development, Senses of Hanzi LUO (裸)

1. Introduction

As the world develops continually, the new concept comes into being, and the neology occurs ceaselessly. What is neology in Chinese? Shao(2004) figured out that “from the view of middle, the neology is the words which the last revision of Modern Chinese Dictionary does not contain ”, and he presented the category for Chinese neology as shown
in Fig. 1. We also call neology as out-of-vocabulary word (in short OOV).

![Diagram of neology categories]

**Fig.1** The category for neology

The OOV has been an interested researching field for linguistics and other related studies. The investigation focuses on four issues: the source of Chinese OOV, the reason for Chinese OOV appearance, the regulation for Chinese OOV, and the recognizing and processing for Chinese OOV. The forth issue relates with language information processing, and is the focus of this paper.

On the other hand, the progress of information technology has made the Internet very popular in the world, and many people use it to communicate and acquire information more and more. Internet presents an opening communicating platform, the words in Internet text is more unbending and colloquial than the traditional plane media, and the Intern text is an important source of OOV.
Tab.1 The 6 kinds of OOV words in Internet chat

<table>
<thead>
<tr>
<th>The format of chat OOV words</th>
<th>Rate</th>
<th>Example sentence</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chinese words</td>
<td>61.8%</td>
<td>稀饭直来直去。</td>
<td>“稀饭(Xifan, conjee)” = “喜欢(Xihuan, prefer)”</td>
</tr>
<tr>
<td>Chinese phrase</td>
<td>5.2%</td>
<td>细八细要开个会议?</td>
<td>“细八细(Xibaxi, thin eight thin)”=”是不是(Shibushi, isn’t?)”</td>
</tr>
<tr>
<td>English capital letters</td>
<td>16.3%</td>
<td>PF 他们的做事态度。</td>
<td>“PF”=”佩服(Peifu, admire)”</td>
</tr>
<tr>
<td>Arabic numerals</td>
<td>7.9%</td>
<td>9494 , 该打。</td>
<td>“94(jiushi, nine four)”=”就是(jiushi, yes)”</td>
</tr>
<tr>
<td>Combination of above</td>
<td>8.0%</td>
<td>8 错, 怎么弄得?</td>
<td>“8 错(bacuo, eight error)” =“不错(bucuo, right or OK)”</td>
</tr>
<tr>
<td>Face icon</td>
<td>0.8%</td>
<td>天气真好, : -)</td>
<td>“: -)” = “愉快(yukuai, happy)”</td>
</tr>
</tbody>
</table>

As a new thing in the information times, the diction in Internet text becomes a prevalent researching field, and many researchers pay attention to the study. Jiang (2004) analyzed the language in Internet text by the view of culture and pragmatics, and pointed out that the feature of the language is farraginous, fashionable, and variable. Farrago includes the farrago of various homophony substitutes, the farrago of spoken language and written language, the farrago of ancient Chinese and contemporary Chinese in written language, and the farrago of Chinese and Western language, character and icon, character and numeral, and so on. Variability means one concept can be expressed in different ways, which all the ways can be transformed and replaced each other and they also can overlay each other, then there is chance for new meaning coming across. And fashion refers to the direction of farrago and expression ways always selects vogue and popularity. His research helps to understand the language in Internet text generally and massively. Guo and Hao
Quan Zhang, Yi Yuan, Xiangfeng Wei, Jianming Miao (2006) put forward 12 types of OOV words in Internet text. Furthermore, they analyzed the reason of the OOV words occurrence: saving time and Internet fee, protecting a person’s right to privacy, showing one’s personality, disdaining tradition, and advocating innovation. Their concrete analysis offers more detail to understanding the language in Internet text. Besides the above that focus on linguistic investigation, Xia (2007) did his research based on language information processing. He adopted the real chat language in Internet, and presented the variation types of the chat OOV words. He constructed a mapping model from the OOV words to the normal words according the phonetic feature of words, and processed the chat OOV words. In order to processing the chat OOV words, he divided chat OOV words into 6 kinds as shown in Tab. 1.

We can draw a conclusion from above introduction that the language in Internet text holds some characters, especially in Internet chat. However, this language usually prevails in a given crowd, and it is endowed with meaning in a definite context. And one person can’t understand the words’ meaning if he doesn’t know the language in Internet text. This expression becomes hard to be understood if it leaves its context, and if someone want to understand its meaning, he needs the knowledge of the language in Internet text.

Besides the above mentioned OOV words which appears in Internet and depends on its context, there are also some OOV words in Internet, which occurs in the public text with different character from the before. The Hanzi LUO(裸) recently appears in Internet frequently, and it always composes a OOV word, such as: 裸奔(luoben, to streak), 裸演(luoyan, to perform with nude), 裸诵(luosong, to intonate with nude), 裸替(luoti, replacer with nude), 裸闹(luonao, to make trouble with nude), 裸聊(luoliao, to chat with nude), 裸教(luojiao, to teach with nude), 裸画(luohua, to paint with nude), 裸退(luotui, to saunter with nude), 裸写(luxie, to write with nude), 裸吃(luochi, to eat with nude), 裸秀(luoxiu, to show with nude), and so on. Besides the above words, there are some OOV words including LUO(裸), such as: 裸捐(luojuan, the meaning of the word will show in the following part of this paper, the same as the next words.), which correlates to famous entertainer; 裸考(luokao), which correlates to the national college common entrance examination; and 裸退(luotui), which correlates to leader retrial. According to the category by Shao (2004), the OOV words about LUO(裸) belongs to the combination of sememe class. Furthermore, the meaning of LUO(裸) in “裸捐(luojuan) 裸考(luokao) 裸退(luotui)” is different from the front group obviously by their context. And its expression is not same as the variation in Internet which presented by Xia(2007), there is not phonetic mapping. Then in order to understand these words, we need to analyze the words’ meaning according to their contexts, and to find what they refer to. On the other hand, we also need to investigate the development of the meaning of LUO(裸), then we can realize the words’
meaning by the meaning of LUO(裸).

This paper will aim to the OOV words which are composed of normal Hanzi as combination of sememe. We wish to understand the OOV words which are composed of LUO(裸) and to distinguish the various senses of LUO(裸) through the way which analyzes the composition of the OOV words.

2. The Method of Investigation

Hanzi is unique in Chinese, and it conveys both meaning and pronunciation. The pronunciation, form and meaning are united artfully in one Hanzi. And Chinese holds the feature that “a Hanzi’s meaning is primitive, and a word’s meaning usually is combined” (Huang, 1998). Therefore, although a word is OOV word, the Hanzi which composes the OOV word is not a new Hanzi in Chinese, the Hanzi is already existent. The total of Hanzi does not increase though the concepts in language are more plentiful than before. In other words, although the development of the world makes many OOV words add into Chinese, the total of Hanzi doesn’t increase and it even decreases. This is determined by the Chinese feature. Then, we shall relate the Chinese OOV words understanding with the Hanzi sense recognizing, and realize OOV words meaning according to this feature.

There are many words in real Chinese text, which general dictionary does not contain, even a lexicon for language information processing is very large, however it embodies just a part of them (Tang, 2007). Tang (2007) put forward the concept of dynamically combined words, and deliberately blurred the difference between word and phrase in dynamically combined words, and did not differentiate the two linguistic concepts strictly. The dynamically combined word is a meaning unit for Chinese understanding processing, which is composed of a Hanzi and its neighboring Hanzi or word according constructing mode temporarily, and which needs to be treated as a whole unit by recognizing. Dynamically combined word doesn’t appear in the lexicon for language information processing. It isn’t necessary that any single Hanzi to construct dynamically combined word. The construction of dynamically combined words needs obey some rules or regulations in order that the meaning of the dynamically combined words can be realized. Then the viewpoint of dynamically combined words is to investigate the combination rules between Hanzi and Hanzi or Hanzi and word.

To bring forward dynamically combined word illumines a way for Chinese OOV words which are not entity names: the OOV word meaning can be deduced by the composing Hanzi and their structure form. This way also embodies the feature of Chinese.

In our investigation, LUO(裸) is seldom used as a single Hanzi word, it always appears
in the multi-Hanzi words, and the lexicon don’t include these words. Meanwhile, the feature of combined meaning in the OOV words with LUO(裸) is obvious, and we can guess the word’s meaning from the sense of LUO(裸), and accomplish the processing of OOV words with LUO(裸).

We acquire the primal words and senses of LUO(裸) from Modern Chinese Dictionary (2002 Edition). Further, we collect words including LUO(裸) with Chinese searching engine (Baidu, http://www.baidu.com; and Google, http://www.google.cn), and select significant words. Then we guess the words meaning by their context, in order to obtain reasonable explanation, we also use Chinese Internet online cyclopedia (Baike, http://baike.baidu.com).

3. The words and the sense of LUO(裸) in Dictionary

In Modern Chinese Dictionary, the meaning of Luo(裸) is: bare, and naked.

Meanwhile, it lists the words with Luo(裸), we partition these words into two groups according to the combining direction of Luo(裸), fore combining and back combining.

Fore combining:
赤裸(Chiluo): bare; naked, without any cover.
赤裸裸(Chiluoluo): naked, undressed; naked, without any cover, direct.

Luo(裸) just combines with Chi(赤) in the fore combined words. Chi(赤) holds many senses, such as: (i) a kind of color which slightly fleeter than vermilion; (ii) red; (iii) standing for revolution; (iv) loyalty; (v) bare, naked; (vi) empty; (vii) pure gold. Among them, there is a sense close to the sense of Luo(裸), the two Hanzi combine together to emphasize this sense. This is a common way for constructing a word in Chinese. Both of the words hold two senses, and the second sense is metaphorical meaning. We analyze the instance sentences in the dictionary in detail, and find that, for 赤裸(Chiluo), its metaphorical meaning is a kind of development which the described object becomes nature thing from the human body, and as to 赤裸裸(Chiluoluo), the described object becomes event from the human body.

Back combining:

Luo(裸) + noun

裸体(luoti): nude.
裸眼(luoshi): naked eye.

Luo(裸) + verb

裸视(luoshi): to watch with naked eye; the eyesight of naked eye.

Constructing terminology
裸麦(luomai): Qingke
裸线(luoxian): mental lead without insulated wrap.
裸子植物(luozi zhiwu): gymnosperm.
裸机(luoji): the mobile phone which has not connected to the communication net; the computer without operating system and other software.

In the back combined group, the majority of the words are Luo(裸) and noun combined words. The 视(shi) in 裸视(luoshi) actually holds two syntax functions: as noun, means eyesight; and as verb, means to watch. We sum up that Luo(裸) modifies the state of the noun in the back combining group, means naked and undressed. The senses of Luo(裸) in terminology are more complex, especially the sense in 裸机(luoji) means without basic software or setting, and not to be used or operated directly for. Meanwhile, the senses of Luo(裸) in terminology also lack the common sense of Hanzi for combining, its usage is more special, then we shall fix these senses in the term of Luo(裸).

4. The Dynamically combined words composed of Luo(裸)

Besides the above words listed in the dictionary, there are many OOV words with Luo(裸) in the Internet text. We find that Luo(裸) is an active Hanzi for constructing a word by combining other Hanzi, and list the words with Luo(裸) in Tab. 2.

To learn from Tab.2, the fore combining words is few, only two words, and they are composed of 全(quan) and 半(ban). 全(quan) and 半(ban) convey measure and range in the words. The meaning of Luo(裸) in the two words is same with two sense: (i) nude, the combined word is adjective, for instance, 半裸模特(banluo mote, a partly nude model); (ii) naked, the combined word is verb, for instance, 半裸臂膀(banluo bipang, naked arm partly).

We divide the words in back combining group into two kinds according to the combined Hanzi as noun or verb. Combining with noun, Luo(裸) holds two senses: (i) nude; (ii) naked, undressed. Luo(裸) shall adopt the second sense when the noun is used to convey a part of human body, in others it shall adopt the first sense. Because the person in the nouns with Luo(裸) is regarded as a whole conventionally, Luo(裸) depicts the apparent state of the person, it is appropriate to consider the sense of Luo(裸) as nude. If the noun conveys a part of human body, it is not a whole body that is naked, and only is partly naked. Further, it needs to point out the naked part when it is partly naked, then the relevant part need to occur in the word.
<table>
<thead>
<tr>
<th>No.</th>
<th>Words</th>
<th>Structure</th>
<th>Combining direction</th>
<th>Meaning of Luo(裸)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>半 裸 (banluo, partly naked)</td>
<td>Modifier + Luo(裸)</td>
<td>Fore</td>
<td>Bare, naked</td>
</tr>
<tr>
<td>2</td>
<td>全裸 (quanluo, completely naked)</td>
<td>Modifier + Luo(裸)</td>
<td>Fore</td>
<td>Bare, naked</td>
</tr>
<tr>
<td>3</td>
<td>裸 男 (luonan, nude man)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>4</td>
<td>裸 女 (luonv, nude women)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>5</td>
<td>裸身 (luoshen, nude body)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Bare, naked</td>
</tr>
<tr>
<td>6</td>
<td>裸 尸 (luoshi, nude corpse)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>7</td>
<td>裸 替 (luoti, replacer with nude)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>8</td>
<td>裸 图 (luotu, drawing of nude)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>9</td>
<td>裸 戏 (luoxi, play of nude)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>10</td>
<td>裸 胸 (luoxiong, bare breast)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Bare, naked</td>
</tr>
<tr>
<td>11</td>
<td>裸 照 (luozhao, picture of unde)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>12</td>
<td>裸 足 (luozu, naked foot)</td>
<td>Luo(裸) + noun</td>
<td>Back</td>
<td>Bare, naked</td>
</tr>
<tr>
<td>13</td>
<td>裸 奔 (luoben, streaking)</td>
<td>Luo(裸) + verb</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>14</td>
<td>裸 吃 (luochi, to eat with nude)</td>
<td>Luo(裸) + verb</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>15</td>
<td>裸 逛 (luoguang, to saunter with nude)</td>
<td>Luo(裸) + verb</td>
<td>Back</td>
<td>Nude</td>
</tr>
<tr>
<td>16</td>
<td>裸 画 (luohua, to paint with nude)</td>
<td>Luo(裸) + verb</td>
<td>Back</td>
<td>Nude</td>
</tr>
</tbody>
</table>
In the verb combining group, the situation of 裸露 (luolu) is same as 赤裸 (chiluo), which we introduced above. The sense of Luo(裸) in other OOV words means nude, and depicts the actor’s state when he acts. And all the combined verbs are about human activity. Therefore, we shall realize that Luo(裸) refer to the state of the actor who conducts the action if it combines with a single Hanzi verb about human activity.
So, we can find the senses of Luo(裸) in dynamically combined words:
(i) nude, (ii) bare, naked, undressed

5. New meaning of “裸”

We seem to end our investigation. However, we come across some OOV words with Luo(裸) when we analyze the OOV words with Luo(裸) in Inter text furthermore:

裸捐 (luojuan)
尽管之前《笑里逃生》剧组曾几次辟谣，澄清该片和牛群裸捐并无关联，而该片总监制彭定邦告诉记者，这个地点是政协方面负责安排的，并非他们有意而为之。
(From http://gb.cri.cn/6851/2004/11/27/108@373348.htm)

裸考 (luokao)
为保障“裸考”考生的公平权益，有人建议取消高考加分政策。
(From http://news.xinhuanet.com/comments/2006-06/04/content_4637614.htm)

裸退 (luotui)
“一府两院”新任领导成员分析：裸退是进步的表现。
(from http://news.xinhuanet.com/politics/2008-03/19/content_7818540.htm)

If we understand these three words according to the sense of Luo(裸) which we presented above, although they are reasonable in someway, they are ridiculous in their context. Then, we think that the sense of Luo(裸) in the three words must be different from the above. We can find the explanation of the words from the Chinese Internet online cyclopedia Baidu
Baike (http://baike.baidu.com).

裸捐 (luojuan): To donate one’s all property, including the corpus after death.

裸考 (luokao): To take part in the national college common entrance examination without any additional score beyond the examination.

The explanation in Baike aims at the national college common entrance examination, however the meaning of 裸考 (luokao) does not limit to the college common entrance examination, for example to the civil service examination. Then, we understand 裸考 (luokao) as taking part in an important examination without any additional score beyond the examination.

Perhaps the occurrence of 裸退 (luotui) is relatively short, we can not find its explanation in Baike. However, we can guess its explanation from its appearing text, and find the meaning: for a leader, to retire without holding any social position.

From the three words, we find that the senses of Luo (裸) are different from the above. Although there are the many divergences among the words, the function of Luo (裸) is same, as a modifier to modify the next single Hanzi word. Deep analysis shows that the sense of Luo (裸) in 裸考 (luokao) is same as Luo (裸) in 裸视 (luoshi),裸线 (luoxian) and 裸眼 (luoyan), means without any attachment or accessory. This sense belongs to usage development of the original sense, to make the sense used in dynamic combined word. The senses of Luo (裸) in 裸捐 (luojuan) and 裸退 (luotui) are similar, both of them mean to abandon all the thing (property, social position) which a person had held. This sense was not included in the dictionary, and the word with this sense was not contained. This sense belongs to sense development of the original sense. Then, we need to add two senses for Luo (裸), (i) there is not any any attachment or accessory. (ii) to abandon all the thing (property, social position) which a person had held

6. Discussion

Here we review the senses of Luo (裸) further. There are four senses of Luo (裸) as above shown. Is there any common component of all sense of Luo (裸)? If yes, what is it? We find that there are some common in the senses of Luo (裸) when we compare the senses each other in detail. The common of the senses is that to take out the accessorrial content makes the main body stands out. In 裸休 (luoti), the clothes is taken out, and the human body stands out; in 裸捐 (luojuan) and 裸退 (luotui), the property and social position, event the corpus, are taken out totally in order to accomplish donation and retiring. Therefore, the senses of Luo (裸) is developed in OOV words in Internet text, but the kernel semantic content is not changed fundamentally, the development occurs in the modified object, the
object from state of individual to the state of event, and the accessory content also from concrete to abstract.

In addition, there are many OOV words which are composed of Luo(裸) combining with a verb, and a verb always conveys the syntax feature. The Hierarchical Network of Concept (in short HNC) theory classified the sentences in two kinds at the top level: action and effect (Huang 1998). In simplification, the action sentence can correspond with active voice expression, and the effect sentence with the passive voice expression. This induces our interest to investigate the OOV words of Luo(裸) in the aspects of action and effect. As to Luo(裸) itself, it only expresses a kind of state, and belongs to effect expression. Moreover, it can be distinguished action expression and effect expression how to come to this state, that is to say that whether the main body actively comes to this state or not. As Luo(裸) itself belongs to effect expression, and a noun is not able to express the action or effect, the OOV words which Luo(裸) combines with noun is belongs to effect expression. In the OOV words which Luo(裸) combines with verb, except the verb in 裸死 (luosi) belongs to effect expression, the verbs in the other words belongs to action expression. We find that these words with action expression verb belong to action expression, and the word 裸死 (luosi) belongs to effect expression. In the fore combining group, the combined Hanzi either emphasizes the sense of Luo(裸) or modifies the degree of Luo(裸), they all belong to effect expression.

Finally, there are two aspects to discuss. Why Luo(裸) holds high frequency in Internet text? Xia(2007) pointed out that the nature of the language in Internet is anomalous. In order to be anomalous, the language needs to be larruping, and shows its nature. The OOV words with Luo(裸) hold this nature, and Luo(裸) becomes the activist in Internet text.

Furthermore, we need to discuss the development senses of Luo(裸) deeply. Besides the anomalous nature, the reason that Luo(裸) involves in the OOV words also relates to the linguistic expression vacancy (Hu, 2007). Chinese lacks a Hanzi to express to take out abstract content, then develops the sense of Luo(裸), the modified object from state of individual to the state of event, and the accessory content also from concrete to abstract.

7. Conclusion

With the development of the world, the OOV words come forth continually. The research on OOV words has already been an important issue for many fields including language information processing and lexical semantics. Meanwhile, the Internet grows and prevails rapidly, and it offers a convenient and expedite communication platform for human beings. The language in Internet also becomes a source for linguistic investigation. This paper
The Senses of Hanzi LUO(裸) in Internet Text

differs from other papers which concerns the anomalous OOV words in Internet, it focuses on the OOV words which is composed of normal Hanzi, and selects the words with Luo(裸) to research. Chinese holds the feature that “a Hanzi’s meaning is primitive, and a word’s meaning usually is combined”, we analyze the explanation of the words with Luo(裸) in order to acquire the senses of Luo(裸) for OOV words processing. We collect the OOV words with Luo(裸) in Internet text by Chinese searching engine, and compare the senses of Luo(裸) in the OOV words with its senses in the dictionary. We discuss the senses and its development of Luo(裸), and also try to explain the reason for the sense development.

The next step of our work is to formalize the semantic knowledge of Luo(裸), and to form rules for dynamically combined words with Luo(裸). These will help to OOV words processing.

Acknowledgement

The paper is supported by the National Fundamental Research Project (973 Project) (Grant No. 2004CB318104), the Knowledge Innovation Engineering Project of Institute of Acoustics, CAS (Grant No. 0654091431), the Fund for Excellence by the Director of the Institute of Acoustics, Chinese Academy of Sciences (the contract No. GS13SJJ04), and the Fund for Young Excellence in Advance Field of CAS(Grant No. O754021432).

References

Shao, Y., 2004, Neology investigation in recent years, Academic study, vol. 9, pp. 131-134.