A COMPARATIVE STUDY OF SYNTATIC AND PHONETIC FEATURES OF INTERNET QUASI-CHENGYU AND EXISTING CHENGYU

Yautina Zhang

Master program: Language and Communication
Linguistics Institute
Leiden University

Universiteit Leiden

Contents

1. Introduction

Following the fast development of communication technology, the internet is now playing a big role in daily life. People's language behaviors are changing – different grammar use, code-switching and the creation of new words have emerged in large numbers. With the quick spread of internet access in Chinese speaking areas and the large Chinese speaking population base, Chinese has surpassed Japanese as the second most used internet language in the world (Liu 2002, p.35). The huge amount of Chinese text on the internet, along with the fact that Chinese internet language is changing quickly, has also yielded various phenomena of Chinese internet language. As a Chinese native speaker, I was intrigued by these phenomena of internet Chinese. Therefore, this thesis will be focused on one phenomenon: special-made four-morpheme phrases which originally appear on the internet. My first impression was that Internet words remain on the Internet, until one day in a conversation, one of my friends used *rén jiān bù chāi* 人艰不拆, which is a typical internet word which will be explained in more detail later in this thesis It was at that point that I realized that internet words have influenced the language beyond the internet and moved into offline conversation. Some self-made four-morpheme combinations have been used on the internet as a sort of idiom like the one just mentioned. This thesis will be mainly concerned with answering the following question: Since there are already various four morpheme combinations in Chinese, will there be differences and similarities between the existing four-morpheme combinations and the new ones on the internet? To this end, first a few important concepts will be introduced.

1.1. Internet language

Internet language, or netspeak, is mostly used on forums, Twitter, blogs, online-chat rooms, and other forms of social media. Communication on the internet is different from other means of communication, mainly due to its special relationship to time and place. With the help of the internet's memory, the communication can be asynchronous as well as synchronous, so the listener and speaker can, but don't need to join the communication at the same time. As for place, the communication all takes place on the internet platform, therefore the geographic location of each speaker is no longer relevant. The speaker and the listener do not need to be face to face in order to carry out a conversation. Furthermore, internet communication can be bi-directional, in the case of a conversation in a chat room or emails, or one-directional, as in blogs or informational websites [Andrist]. Bi-directional communication is comparable to face to face communication or traditional letters, in which at least a speaker and a listener are involved. It is closer to the spoken language. One-directional communication is comparable to books, which authoritatively represents written language. There is no speaker and listener, but a narrator and the audience.

Although one-directional and bi-directional communication makes internet language comparable with both written and spoken language, it differs from both. Internet language is different from simple written text that is based on spoken language, because people also use a lot of language variations such as abbreviations, acronyms, stripes, punctuation, and even emoticons, which people would not use in a spoken conversation. It is also not written language, because many words or grammar uses are too colloquial for written language. I would suggest that internet language occupies a space between written language and a written form of spoken language, but is closer to the latter.

There are also other factors that influence the internet language: Censorship plays a role in people's word choice, and also influences of what source one is able to have access to; the limitation of text length (such as on Twitter and Weibo) drives internet users to form their utterance differently; the input method also affects

people's word choice. For instance, one may select the first option that appears on the option list after typing the pinyin (Chinese spelling system) of a word, regardless of whether it is the right word for the situation. Some of these factors are universal, yet others are specific to Chinese.

The features of internet language are also different in different contexts. The context of internet language can be news, advertisement, internet BBS, chat room, email, etc. (Liu 2002, p.38-52).Within different contexts, different styles of language are required. For instance, language in a chat room are more informal, compared to a news cast. The use of language (such as using Zzzz for the sound of snoring), creative homophony (such as using a different character which has the same pronunciation as the intended character), lexicon (such as newly created words), and grammar (such as the sentence formation). Changing or expanding of lexicon happens most with internet language (Liu 2002, p.54). An enormous amount of new words is made, and many of them disappear after a while because they are buried under the fast expanding internet vocabulary. In some cases, some words have become so popular that people carry them over to actual conversation.

1.2. Forms of Internet words

With various internet platforms, "netizens" (an internet word, means citizens on the internet) have made diverse kinds of internet words The most common kinds of internet words are the following:

1. Abbreviations

An abbreviation is a shortened version of a long expression, in order to make the communication or expression more convenient. A typical form of abbreviation is to only use the first letter of the words to shorten it. For example, *the US* is the abbreviation of *the United States*. This concept is not new. In Chinese there are also multiple forms of abbreviations (Liu 2013).

a. Abbreviations can be made with pinyin: LS refers to *lóu shàng* 楼上 'upstairs'; BS refers to *bǐshì* 鄙视 'despise'.

b. They can come from English words: GF refers to 'girlfriend'; DIY refers to 'Do It Yourself'.

c. They can also be a merger of two syllables. Two syllables can merge into one, reflecting fast speech. These kinds of abbreviations mostly come from Taiwan Mandarin, such as *jiàng zǐ* 酱紫 (literally 'dark purple') from *zhè yang zi* 这样子 'something like this', or *biǎo* 表 (literally 'watch') from *bú yào* 不要 'no', and *zào* 造 (using a character that is normally used for 'to make') from *zhīdào* 知道 'to know'.

These three types of abbreviations appear with a clear goal: to make words shorter so that internet users do not need to spend as much time typing. Internet users also use abbreviations to avoid using curse words when they need to express strong feelings or sometimes to make fun of words (Liu 2013, p.26). They are still widely used.

Besides these three kinds of abbreviations, a fourth has become popular recently: an abbreviation for a long sentence or phrase, reducing it to a short word. This has nothing to do with pinyin, English or syllables, but rather an issue of semantics or syntax. The long sentences or phrases appeared first online and have become frequently used. Later the internet users picked up several characters of the original ones to make them shorter.

For example:*nǎo bǔ* 脑补 possibly comes from:

在	脑海	里	补充
zài	nǎohǎi	lǐ	bǔchōng
at	mind	in	add

Literal: to add up information in mind
Figurative: to imagine

Here is another example:

Gāo dà shàng 高大上 comes from:

高端	大气	上档次
gāoduān	dàqi	shàngdàngcì
high-end	lofty	high-class

Literal: high-class, advanced
Figurative: same as above

These three words have similar meanings, the abbreviation attracts one character of each words and it means the same.

2. Near homophony

Another way new words are formed on the internet is near homophone. Near homophone (in Chinese: *xieyin*) means one word sounds like another. Sometimes some characters are used to replace the original ones to represent the same meaning. There are four kinds of near homophony:

A. Replacing the original characters with numbers. For instance, 7878 (pronounced as *qī bā qī bā*) sounds like *qùba qùba* 去吧去吧 'Sure, just go'.

B. Replacing the original characters with another character. *Chǎo jī* 炒鸡 (literally means 'fried chicken') is a near homophoneof *chāojí* 超级 'super'.

C. Replacing the original words with near-homophonous English words or letters. For example, *chén guàn C* 陈冠 C is used to replace *chén guàn xī* 陈冠希 (an actor from HongKong).

D. Replacing the original foreign words with Chinese characters or numbers. One of the most commonly used daily expressions 88 (pronounced as *bābā*) is actually referring to the English greeting 'Byebye'. IPhone is also transformed into *ài fēng* 爱疯, which is simply a near homophony.

3. New meaning for an existing word

On some occasions people give new meanings to existing words. Sometimes it is due to the fact that the existing word can be used to describe something new on the internet, such as *shāfā* 沙发 'couch', which has been given a new meaning as 'someone who is the first to reply to a post'. It could be interpreted that the

post is a 'TV show', and the first person who wants to enjoy the show chose to sit on the place with the best view—the couch. This idea is in comparison to the English way to refer to an argument that is about to break out which is to "grab popcorn" to eat while you watch other people fight.

Another possible reason to use an abbreviation that the existing word is taboo, or uncommon. Internet users tend to choose the word which is not taboo, for instance the same word with different characters, or another word which bears some resemblance to the meaning of the taboo. For example, *bōzhǒng* 播种 'to sow' is also used as 'to have sex', because words directly relating to sex are seen as taboo, or can be filtered online.

Another type of netspeak similar to "near homophone" is the replacement of characters with characters suggested by auto-complete. However, a near homophone doesn't need to be an existing word. For example, *ní méng* 泥萌 is a near homophone for *nǐmen* 你们 'you, second person plural', but it is not a word itself. *Lánzhōu* 兰州, which means now also 'the original person who send a post online', is originally a city name in China. This word is chosen to replace *lóuzhǔ* 楼主 'landlord', which is also a new internet word, which describes the person who sends a post. Before *lóuzhǔ* became widely used, when the user types in the pinyin onsets of the word, which is *LZ*, *lánzhōu* appeared first. That's why now *lánzhōu* is used to refer to *lóuzhǔ*. As for *ní méng*, it is more as a mocking of dialect, in which *eng* and *en* are mixed. The second difference between near homophone and auto-completed words is that a near homophone is closer to the original word when it comes to pronunciation. Returning to the examples above, *chǎo jī* and *chāojí* only differ in tone (because Chinese typing software doesn't identify tone), but *lánzhōu* and *lóuzhǔ* only share the same the onsets. Sometimes there can be an overlap: *zhūjiǎo* 猪脚 'Pig feet' is a near homophone of *zhǔjiǎo* 主角 (it is actually an incorrect pronunciation of *zhǔjué*, because the character 角 has two pronunciations, of which jiao is the most common, but in this instance it is incorrect), which means the main role in a film, book, play, etc. As a result, the meaning of *zhǔjiǎo* is also added to the word *zhūjiǎo*.

The third situation in which a new meaning is given to an existing word is that the new meaning of the original word is related to the literal meaning of the word. For example, *lǜchá* 绿茶 'green tea' has now been given a new meaning, referring to a kind of young woman that pretends to be innocent and harmless, but actually is good at taking advantage of others, materialistic, and sometimes even willing to use sex as a bargaining counter to reach her goal. This type of woman is also labled *lǜchábiǎo* 绿茶婊 'green tea bitch', because green tea symbolizes 'pure' and 'fresh', which is what this sort of woman pretend to be. Another simpler example is as follows: the word *wāng* 汪 'onomatopoeia of a dog's bark' is used to refer to a dog, because a dog and its bark are related.

4. Split and merge

Yet another way of forming new words is by splitting and merging. Splitting one character into more characters (which also exist by themselves) is not a new phenomenon. It does not only happen in word games, but also in everyday life. When I introduce myself to someone for the first time, I might say: 'My family name is *zhāng* 张, *gōng* 弓 *cháng* 长张.' That is because there are two family names pronounced as *zhāng* and I would like to distinguish mine from the other *zhāng* 章 (*lì* 立 *zǎo* 早章). Both *zhāng*'s are split by its components; 张 is split vertically into 弓 *gōng* and 长 *cháng*, and 章 is split horizontally into 立 *lì* and 早 *zǎo*. The

splitting of characters is also used on the internet. There are multiple reasons: firstly, some internet users want to be creative and distinguish themselves from the average internet user by using non-traditional ways of expression. Split characters make a different visual impression. Secondly, some words are split because they are taboos or profanities, so the internet users try to avoid being rude in this way. Thirdly, on account of Chinese internet censorship, some sensitive words or topics can be filtered. Internet users tend to split characters to avoid the censor. For example, *qiángjiān* 强奸 'to rape' is horizontally split into 弓虽女干, and *qiāng* 枪 'gun' is split into 木仓.

In contrast to splitting, merging means to merge several characters into one character. The new character already exists in the language, but may have another original meaning. By merging, a new meaning of this character is added. This phenomenon is relatively new in comparison to splitting. The reasons for merging are more or less the same as with splitting, but there can be other reasons: Firstly, the merged character can vividly describe something. Secondly, merging more characters into one makes the conversation go faster. For instance, the character *méi* 槑 (originally is a variant of *méi* 梅, a type flower) has no connection to *dāi* 呆 'dull', but because this character is two 呆 combined together, it is now widely used to refer to *dāi dāi* 'a duplication of "dull", therefore 'very dull'. Other examples like *háo* 壕 'moat' is also used as a merge of *tǔháo* 土豪 'local tyrant, now also means a rich person (in a jocular way)'. *Mào* 氁 'between 80 and 90 years old' is the vertical merge of *lǎo máo* 老毛 'literally old Mao, which refers to Mao Zedong, a famous Chinese political figure'.

5. *Phonetic variant of existing words*

Although Mandarin is the official language in China, different areas have their own dialects. When they speak Mandarin, there are variations in their language because of the influence of their dialects or accents. In the North-Eastern area, 'r' is often pronounced as 'y'; 'z', 'c' and 's' are often mixed up with 'zh', 'ch' and 'sh'. Some North-Western and Southern areas have difficulties separating the front and back nasals. Some Southern areas mix up retroflection, or between 'r' and 'l', 'n' and 'l', or 'h' and 'f', etc. Accents make vocabulary on the internet more varied. Internet users have made words based on different accents: *huīcháng* 灰常 (literal: gray, common) is a variant of *feīcháng* 非常 'very', which comes from the h-f confusion. The same goes to *huīj* 灰机 (literal: gray, machine) and *feījī* 飞机 'plane'. *Nányín* 男银 (literal: male, silver) or 男淫 (literal: male, obscene) is a variant of *nánrén* 男人 'man', due to the r-y confusion. *Liúbì* 流弊 (literal: flow, harm) is a variant of *niúbī* 牛逼 'freaking awesome (profanity)', basing on the confusion of n-l. The main intention is to make informal words by mocking the accents or dialects.

6. *Combination of different languages*

As the internet connects the whole world, other languages are brought into use on the Chinese internet quickly. New Chinese internet words are also influenced by foreign languages, and some combinations are now widely used on the internet.

Some combinations are on a purely lexical level. A classic example is *I fúle* 服了 *U* 'I serve you', which is a combined phrase of Chinese and English. Here the only difference from the proper Chinese sentence is that some words have been replaced by English words.

Other combinations are on the grammatical level. A recent example is *no zuo no die*, in which *zuo* comes from Chinese *zuō* 作 'north dialect: seek for trouble'. This is originally a Chinese saying *bù zuō sǐ jiù bú huì sǐ* 不作死就不会死 'You wouldn't die if you hadn't sought for death'. Internet users make it shorter in order to communicate more effectively by applying English grammar. The *No A no B* structure is common in English, such as *no pain no gain,* this has been adopted by Chinese. However, because *zuō* is hard to translate with one English word, let alone a word with one syllable, it remained the same. So this phrase is a combination instead of a Chinglish word. Moreover, in the original English "no A no B" structure, A and B are always nouns. When Chinese borrow this structure, it is further adapted. Other word classes than nouns are allowed at the position of A and B, such as *die*, which is a verb.

Combinations can even be on the morphological level. For example, *chīfàn* 吃饭 *ing* 'eating'. Compared to similar elements in Chinese, the English *ing* can more easily and clearly indicate the present progressive tense, therefore it became popular. *Niubility* 'awesomeness' is another kind of combination that comes from putting a Chinese element in a self-made English word. *Niubi* comes from *niúbī* 'damn awesome', an adjective, and –lity is a typical English suffix for nouns. In this way *niúbī* is nominalized.

Combinations are not only Chinese –English, other languages such as Japanese or Korean are also widely used. For example, *X sang* 桑 means Mr. X, and *sang* is the Chinese near homophone of Japanese さん *san* 'sir'.

7. Special suffix

The most popular suffixes are *X gē* 哥 'big brother', X *jiě* 姐 'big sister', X *dì* 帝 'king' and X *xīng rén* 星人 'people from planet X'. Most of these suffixes are used to refer to someone, but it is not the traditional way to use it. Traditionally we tend to call people who are in the same generation but older Y 哥 *gē*, Y 姐 *jiě* 'big brother/big sister Y (Y is the person's family name or first name)' to show respect. However, the usage of these suffixes on the internet is not the same. The X indicates their characteristics, or the reason that they are known by the internet users, instead of their real names. In 2010, someone posted a group of photos on a photographers' forum, showing the performance of his new camera. A homeless man in the photo became popular with his unique style of clothing and his sharp expressions. They started to call him *xīlì gē* 犀利哥 'Brother Sharp'. Sharp has nothing to do with the man's real name, but rather a factor that he became famous for. Since then, these kind of "brothers" and "sisters" appeared often on social media (Liao 2011). Special suffixes can be used on things other than people. An example of this is *wāng xīng rén* 汪星人, which literally means someone from the planet of a dog's bark, which refers to a dog.

8. Emoji

It is still disputable whether emoji's can be categorized as internet words, but they are symbols exclusively used on the internet. They are unique ways to express one's feelings in a written way, although not in language. Emoji's can be horizontal (such as ': P' as a smile with the tongue out) or vertical (such as '^_^' as a smile). There might even be an Eastern-Western difference.

A special emoji, which I would like to consider as a word, is the Chinese character *jiǒng* 囧. *Jiǒng* is a pictograph of 'window', but because it sounds the same as *jiǒng* 窘 'embarrassed' and looks like a sad face, it is now used as an emoji. Unlike the other emoji's, 囧 expresses the meaning 'embarrassed' both by how it looks and what its sound relates to. This makes it more like a word.

9. *Literal quotation*

When a news clip becomes a hot issue, or an advertisement becomes popular, what people say in it is sometimes literally quoted. Many are sentences or phrases, such as *nǐ dǒng de* 你懂的 'you know it' or *zhì yú nǐ xìn bú xìn, fǎn zhèng wǒ shì xìn le* 至于你信不信, 反正我是信了 'No matter you believe or not, I believe it anyway.' This quotation comes from a serious train rear-end accident which caused 40 deaths and 172 wounded. When asked why the wreck was buried so soon, the spokesperson Wang said that it was for the convenience of rescue. On this far-fetched explanation, he further commented, "Believe it or not, I believe it anyway". This irresponsible and unprofessional comment enraged the public and therefore this quotation is frequently used as sarcasm. In some other situations, new words are made by extracting the essences of these hot issues instead of literally quoting. For example, *zhènglóng pāi hǔ* 正龙拍虎 'Zhenglong takes photo of a tiger', which is a well-known scandal. In 2007, a Shaanxi citizen Zhenglong Zhou claimed he had found awild South China tiger (an endangered species) and taken aphoto of it. Local government in Shaanxi later confirmed that the photo was real. However, after the photo was officially published, its authenticity was questioned by the public whensome people found a painting which was highly similar to the photo. As the local government had verified the authenticity of the photo, the public started to distrust the credibility of the government. After two rounds of professional identification, the photo was proven to be fake. Although Zhenglong still insisted that the photo was real, he was prosecuted for counterfeit and fraud and sentenced to imprisonment. Now this quotation is used as "to fake things in order to profit from it" and to signify the distrust of the government credibility.

Not all the internet words mentioned above are still used. Some of the internet words slowly disappeared as the news became less known, and some other words have withstood the test of time and become a part of the basic internet vocabulary, such as LZ and *shāfā* (Liu 2013, p.25).

1.3. Introduction of related terms

Aside from the internet language discussed in the previous section, there is also a trend that people start to make four-morpheme phrases, which are like a certain type of idiomatic expressions or proverbs, called *chéngyǔ 成语* in Chinese. The original language sources are mainly from the news, quotes by celebrities, and song lyrics. They have become so popular that people sometimes use them in real conversations. Which kind of internet words mentioned above do they belong to? What do they have in common with existing idioms or how are they different? As mentioned at the beginning of the thesis, these type of phrases becomes the central focus of this thesis.

Since Internet language came into being, various internet words have appeared and become a popular topic for language researchers. Many have already researched different aspects of internet words, such as Yang (2014), Shao (2014), Che (2014), Ma (2014), Liu (2009) etc. Self-made four-morpheme combinations have caught the attention of some researchers. In their research, these types of phrases are given different terms. Yang (2014) and Ma (2014) named these kinds of novel phrases *wǎngluò* chengyu 网络成语'Internet chengyu', Liu (2009) and Shao (2014) called them *wǎngluò xīn* chengyu 网络新成语'Internet new chengyu', and Che (2014) have referred to them variously as *wǎngluò sú*chengyu 网络俗成语'Internet vulgar chengyu', *hòu xiàndài* chengyu 后现代成语'post-modern chengyu' and *wǎngluò sìzìgé súyǔ* 网络四字格俗语'internet four-character adages'. There is not yet agreed on an official name for these phrases, but most researchers tend to add the term chengyu in the title. Given that these self-made four-morpheme combinations are used as chengyu; it is very likely that they have some common linguistic characteristics with real chengyu. What exactly do they have in common and what are their differences? In order to understand that, it is necessary to first explain what a chengyu is.

A chengyu is a kind of idiomatic expression or proverb. There are many different types of such expressions and chengyu is just one of them. Some of them will be briefly introduced below. The general designation for well-known fixed phrases and fixed sentences is *shúyǔ*. Like English proverbs, the usage is strictly standardized, just like simplex words. One of their essential characteristics is 'fixed structure'. The components cannot be arbitrarily added, deleted or changed. For instance:

胸	有	成	竹
xiōng	yǒu	chéng	zhú
chest	have	finished	bamboo

Literal: Having a grown bamboo in the chest.
Figurative: Already having a complete plan of what to do.

Xiōng yǒu chéng zhú 胸有成竹 is a fixed structure. Elements from the structure cannot be deleted+ thus *xiōng yǒu zhú* 胸有竹 is ungrammatical. Elements from the structure also cannot be changed, for example *xiōng yǒu chéng mù* 胸有成木 'Having a grown wood in the chest' does not exist. Finally, the order of the elements cannot be changed: *chéng zhú xiōng yǒu* 成竹胸有.

A second characteristic of *shúyǔ* is 'opacity'. The meaning of the whole *shúyǔ* is independent from the meaning of its components. The meanings of the components are (partly) neglected. The meaning *xiōng yǒu chéng zhú* as a *shúyǔ* is takes priority over the literal meanings of the components, which is 'having a grown bamboo in the chest'.

As mentioned above, there are different types of *shúyǔ.* These different types of *shúyǔ*s are different in function and make up. The ones mentioned here will play no role in this thesis. For this thesis, the most important type of *shúyǔ* is the chengyu. Chengyu is the biggest category of*shúyǔ*. It has the following features Wen (2005):

1. It is a special fixed structure with four morphemes and, like other types of shuyu, always used as an entirety.

2. Most of the chengyu have a figurative meaning.

3. They have been in public use for a long time.

4. Chengyu have various origins, such as fairy tales, fables, historical stories, poetry or prose, colloquial phrases, etc[1].

5. The internal structure of chengyus also differ, but predicate-object (VO) constructions are often not used. Details about the structure will be discussed later.

Wen (2005:70) further suggests a narrower concept of chengyu: an expressive or descriptive fixed structure that is divided in two parallel pairs. In his point of view, the parallelism in chengyu is an important factor, and the parallelism can be structurally or phonetically. This concept will be further discussed in paragraph 2.

Chengyus are also divided into 雅成语 *yǎ* chengyu 'elegant chengyu'and 俗成语 *sú* chengyu 'vulgar chengyu' (Wen 2006). According to Wen, there are three characteristics of elegant chengyu: 1. It comes from literary works; 2. It consists mostly of elements of classical Chinese, including classical lexical words and classical function words; 3. It is only used by a select group of people, which are mostly intellectuals. Vulgar chengyu, on the other hand, can come from ancient, pre-modern or modern spoken language. It has more vernacular elements and is broadly used among the general public. Vulgar chengyu are seen as part of *súyǔ* (another type of shuyu) (p.67).

 Example of an elegant chengyu:

四	面	楚	歌
sì	miàn	chǔ	gē
four	side	Chu	song

Literal: Chu songs from four sides.
Figurative: Besieged on all sides.

This chengyuis from historical documents and consists of classical Chinese words. Therefore, it is seen as anelegant chengyu.

1

　　Based on the terms in *xiàndài hànyǔ zhīshi cídiǎn* p.173-174

Example for vulgar chengyu:

三	长	两	短
sān	cháng	liǎng	duǎn
three	long	two	short

Literal: three long and two short
Figurative: Accident or disaster that causes death.

This example comes from a classical opera, which can be seen as a style of spoken language, and has more vernacular elements and is broadly used among the general public. Because, compared to a book, opera has a larger target audience and the language is more down to earth, it can be seen as a vulgar chengyu.

Although chengyu always consist of four morphemes, not all lexical items with four morphemes are chengyu. All sorts of other lexical items can consist of four morphemes such as other types of *shúyǔ*s, as well as different types of compounds, fixed or otherwise, and other types of phrases. Therefore, when a new four-morpheme structure needs to be verified to seewhether it is a chengyu, all features of chengyu should be present, otherwise it is not a chengyu.

Judging from the definition above, I agree with Che's (2014) way of naming the internet four-morpheme structures. Instead of directly applying the term chengyu, he introduced a new term: *wǎngluò lèi*chengyu 网络类成语 'Internet quasi-chengyu'. Because these four-morpheme structures are not completely accepted by the public it is still unknown whether they will be incorporated into the general language. However, since they are fixed structures and often have some figurative meanings, they do share some features of a real chengyu and therefore the word "quasi" is more precise.

In this thesis, the term "Internet quasi-chengyu" will be used to refer to these internet four-morpheme structures. Some quasi-chengyus are closer to vulgar chengyu, and some others are closer to free phrases. In general, we can say that they are something between chengyu and free phrases, but in an intermediate phase developing towards the structure of chengyus (Che 2014, p.76). According to Che (2014), Internet quasi-chengyus have the following features:

1. The structure is fixed. The elements cannot be randomly replaced, rearranged, added, deleted or split. For example, *zhào shān xiàn guǐ* 兆山羡鬼 'Zhaoshan envies the ghost' cannot be changed into *zhào shān xiàn shén* 兆山羡神 'Zhaoshan envies the god' (changing of elements) or *zhào shān guǐ* 兆山鬼 'Zhaoshan the ghost' (deleting elements, etc.

2. The language style is typically very informal, common, and often derogative.

3. Sometimes the meaning is more opaque, which means that without an understanding of the origin of the word, it is not possible to understand the meaning.

4. Internet quasi-chengyus come from the Internet. Unlike real chengyus, quasi-chengyus are not from legends, fairy tales, historical stories or poetry. Instead, they are from news, current events, and popular quotes and so on.

5. Use of stylistic forms and meaning. For example, rhetorical question and metaphor.

As mentioned above, Internet quasi-chengyu*s* are not real chengyus, but chengyu-like. How are they connected and what are the similarities and differences? In the following sections, I will discuss the linguistic connections between the Internet quasi-chengyus and the real chengyus. In chapter 2 I will make comparisons of the syntactic features between real chengyu and internet quasi-chengyu. Chapter 3 will be the comparison of the phonetic features.

2. Internet quasi-chengyu

2.1. Forms of Internet quasi-chengyu

In the last paragraph the main features of chengyu and internet quasi-chengyu have been discussed. With the changing of hot issues on the internet every day, new Internet quasi-chengyu appear while some old ones are less used than before. Che (2014) has made a selection of these quasi-chengyu, arranged by Internet frequency of use on Weibo. Fifty-eight quasi-chengyu have been selected, which all have a frequency higher than 1000[2]. Whether they are all quasi-chengyu or rather fixed phrases still remain*s* uncertain but seeing as how this is a recent and relatively complete selection, the sources and meanings of these quasi-chengyu are difficult to trace back. Those examples of which the structure is closer to that of fix- or free-phrase rather than that of a chengyu have been deleted. Quasi-chengyu that appear later than this collection will be added in Appendix 1.

As analyzed in paragraph 1.2, there are ten types of internet words. Internet quasi-chengyu belong to abbreviations, and literal quotations. In the latter case, internet quasi-chengyu are made by extracting the general idea out of a literal quotation. The term Internet quasi-chengyu also means that they originally appeared on the internet.

Yang (2014) categorized four forms of Internet quasi-chengyu, which can be applied to this thesis:

A. Abbreviation

Abbreviations mostly come from a sentence, such as a sentence from lyrics. These sentences can be found on the internet, such as forums or blogs. Four core characters or morphemes from the sentence are chosen in order to make a quasi-chengyu. These four core characters represent the meaning of the whole sentence (Yang 2014, p.110).

人			**艰**				**不**	**拆**
人生	已经	如此	艰难，	有些	事情	就	不要	拆穿
rénshēng	yǐjīng	rúcǐ	jiānnán	yǒuxiē	shìqíng	jiù	bùyào	chāichuān
life	already	such	hard	some	thing	just	no	expose

Literal: Life is already so hard; you don't need to expose everything.
Figurative: Same as above.

2

This data selection is on 23rd, June, 2014. Seeing that the main tendency of frequency of use will keep the same direction, although the exact frequency till now may be changed, this data base is recent enough for this thesis.

细		思 (sī)		恐	极	
仔细	想想，	觉得	恐怖	极	了。	
Zǐxì	xiǎngxiǎng	juéde	kǒngbù	jí	le	
careful	think	feel	horrible	extreme	particle	

Literal: Realizing that something is really scary after giving it a second thought.
Figurative: Same as above.

	累	觉	不			爱	
很	累，	感觉	不	会	再	爱	了。
Hěn	lèi	gǎnjué	bù	huì	zài	ài	le
very	tired	feel	no	can	more	love	particle

Literal: Feeling too tired to ever fall in love again.
Figurative: Same as above.

The three examples above illustrate how the core morphemes are extracted from the whole sentence. Sometimes it does not need to be a character precisely from the original sentence as long as the character has the same meaning as the original one. In the second example, *xiǎngxiǎng* is changed to *sī*, because they both mean "to think", but *sī* is more formal and *xiǎngxiǎng* is more colloquial. In the word collection of this thesis, eleven quasi-chengyu belong to this type. Because the four characters contain the core information, although one may not know the original sentence, one can still derive the meaning of the quasi-chengyu. These quasi-chengyu do not have figurative meanings but rather are directly interpretable.

B. Summarize the general idea

This type is consistent with what is said earlier in this section: quasi-chengyu are made by extracting the general idea instead of a literal quotation. For these types of quasi-chengyu there is always a background story. Four morphemes with the core information are chosen to outline the whole story. The core information is typically the main person in the story and what he did, which is called "actor - action" structure in this thesis. Those quasi-chengyu always have a figurative meaning as well, most commonly to expose the darkness of the society or to mock some ridiculous behaviors.

正龙	拍	虎
Zhènglóng	pāi	Hǔ
(name)	clap	Tiger

Literal: Zhenglong takes a photo of a tiger.
Figurative: 1. Someone produces a counterfeit and denies it after
being revealed
2. Lacking of social credibility
3. Out of thin air

At first glance, its literal meaning is not related with its figurative meaning at all. The background story is needed. As mentioned in the first paragraph, Zhenglong took a photo of a painting and claimed that he had seen a real tiger (which is an endangered species), showing the photo as evidence. After the local government had seen the photo and published it, doubts arose among the public. It was soon proven that the photo was fake. The quasi-chengyu *Zhènglóng pāi hǔ* has since then been used to describe situations or articles that were "counterfeit", "making things out of thin air" or "lacking of social credibility".

兆山	羡	鬼
Zhàoshān	xiàn	guǐ
(name)	envy	ghost

Literal: Zhaoshan envies the ghost.
Figurative: Referring to someone cold-blooded.

Without knowing about the background story, it is impossible for someone to derive its meaning. Zhaoshan Wang is a member of theChina Writers Association. He published a poem after the big earthquake in Sichuan in 2008, in which he sang his praises to and curried favor with the government, instead of showing any sympathy to the 60,000 victims. In the poem there is a sentence, *shísān yì rén gòng yì kū, zòng zuò guǐ, yě xìngfú* 十三亿人共一哭，纵做鬼，也幸福 'with 1.3 billion people crying together for me, I could even be happy being a ghost'. His cronyism infuriated the public;the callousness of a man who could celebrate in the face of such a catastrophe was rightly despised. Therefore, this phrasewas made to describe someone cold-blooded.

开	胸	验	肺
kāi	xiōng	yàn	fèi
open	chest	test	lung

Literal: Opening the chest to exam the lungs.
Figurative: The difficult situation for occupational disease patients to protect their rights.

This type ofquasi-chengyu can also be made in other forms than "actor – action": some of them are in a form of "action – details", which represents the main action and what happened exactly with this action. *Kāi xiōng yàn fèi* is such a structure, because compared to the "subject", the action and the object contain more important information. Haichao Zhang was diagnosed with pneumoconiosis (an occupational lung disease), but because the hospitals he went to did not have the qualification for diagnosing occupational diseases, he could not use the diagnosis to apply for treatment and insurance. When he went to the occupational disease precaution clinic, the company he worked for refused to offer him related documents. Later, that clinic diagnosed that he didn't have occupational disease. Left with no choice, Zhang asked for a lung biopsy operation. The clinic finally acknowledged that Zhang did have pneumoconiosis and he successfully received the treatment he was entitled to. This process should have been simple but instead it took much longer with many unnecessary difficulties. Protecting rights in China is a sensitive and difficult issue. In order to emphasize their frustration, Internet users made this phrase.

From the examples above, it is clear that to understand what these quasi-chengyu mean, one has to know their background stories. The meaning cannot be seen from the literal meaning of the composite morphemes. Take *Zhènglóng pāi hǔ* as anexample:*pāi* has multiple meanings, among which are 'photo shoot' and 'clap'. Without knowing the background story, it is not possible to know which meaning

of *pāi* is represented here. Even if one knew the right meaning, one would not won't understand why this word has any relevance. In the research material of this thesis, this type of quasi-chengyu is most common.

C. Adaptation of existing chengyu*s*

This type of quasi-chengyu is more of a word play, but is a reflection of some social issues as well. They are all adapted from existing chengyu either by changing the order, replacing one or more characters with homophones, or merging several chengyu into one. Therefore, the meaning of the quasi-chengyu is still related to the original.

谁	死	鹿	手
shéi	sǐ	lù	shǒu
who	die	deer	hand

Literal: Who dies in the hands of the Deer?
Figurative: A complaint with food safety.

This is an adaption of the chengyu *lù sǐ shéi shǒu* 鹿死谁手 'literal: at whose hand will the deer die. Figurative: who will win the prize?' by switching the place of *lù* and *shéi*. The milk powder from a dairy industry company *sān lù* 三鹿 'Three Deer' was found to contain melamine (which causes kidney stones) and a lot of babies were diagnosed with kidney stones. Being angry and frustrated at the dark minded merchants and corrupt governors, the public made this expression.

十	面	霾	伏
shí	miàn	mái	fú
ten	side	haze	hide

Literal: Haze hiding on ten sides.
Figurative: The public's concern about air pollution.

This is a case of replacing a syllable with a homophone from the original chengyu *shí miàn mái fú* 十面埋伏 'Setting up an ambush on all sides to surround and annihilate the enemy'. In recent years, China is often surrounded by haze, and the PM 2.5 index is unusually high. This has raised strong concerns of the public about the environment and their health since it is impossible to go out without a gauze mask. The Chinese word for 'haze', *mái* 霾, is a homophone of the *mái* from *mái fú* 'ambush'. The meaning of 'being surrounded and in danger' of the original chengyu is also perfectly related to people who are surrounded by severe air pollution. Thus the meaning and form of the original has been translated to this new quasi-chengyu.

D. Borrowing from other language

This type of quasi-chengyu is borrowed from other languages and is adapted with the aforementioned three methods: (abbreviation, summarizing the general ideas and adaptation from existing chengyu). These quasi-chengyu are relatively uncommon. Examples follow:

猪	涂	口红
zhū	tú	kǒuhóng
pig	smear	lipstick

Literal: Putting lipstick on a pig.
Figurative: Someone tries to dress something up, but is still that something

Putting lipstick on a pig is an English saying in reference to someone who may be trying to make something or someone look appealing or attractive when it quite clearly will not work, or will only deceive the dumbest of people. This term can be seen recently during the presidential election in the US when Barack Obama used it in his speech against John McCain and Sarah Palin's new "change" mantra: "You can put lipstick on a pig. It's still a pig. You can wrap an old fish in a piece of paper called change. It's still gonna stink. We've had enough of the same old thing."

压力	山	大
yālì	shān	dà
stress	mountain	big

Literal: The stress is as heavy as a mountain.
Figurative: Under a lot of pressure.

It is an adaption of *yà lì shān dà* 亚历山大, which is a transliteration of the English name Alexander. There is no background story or current topics that caused its popularity, it is only because *yālì* sounds like *yà lì* and the characters of the transliteration vividly describes someone under a lot of pressure.

Although these two examples are both borrowing from other languages, they are two different cases. The first one is a literal translation of an English saying, while the second one is a phonetic translation of an English word and the translation happens to contain some other meaning.

From these four types of quasi-chengyu we can see the overlap between them and the types of internet words discussed in paragraph 1.2. As with real chengyu, most quasi-changyu have an original background and it is necessary to understand the background story of these internet quasi-chengyu in order to understand their full meaning. Most quasi-changyu also have a figurative meaning like the real chengyu does. Besides abbreviations and literal quotations as mentioned earlier, we can also find other strategies such as applying near homophone and adapting existing words. Since there are so many types of Internet quasi-chengyu, do they also overlap with the real chengyu? It is this question that I would like to talk about and explore in the next section.

2.2. Syntactic features of real chengyu and Internet quasi-chengyu

2.2.1. Syntactic features of real chengyu

In paragraph 2.1 the definition and features of chengyu were discussed. Chengyu generally have a fixed structure, are used idiomatically and are still in use today. As was mentioned in the paragraph 2.1, a chengyu always consists of four morphemes, but expressions with four morphemes are not all chengyu. As mentioned in chapter 2.1, a chengyu is considered to have a fixed structure, which should have expressive or descriptive functions. Besides, its four morphemes – also four syllables – should be seen as two pairs. Wen (2005) further suggests that the division can be grammatical, semantic or prosodic. From a grammatical and semantic point of view, lots of chengyu *have* symmetrical structures, such as *hóng nán lǜ nǚ* 红男绿女 'Red men and green women'. They can be seen as two symmetrical pairs: *hóng nán* 'red man' vs. *lǜ nǚ* 'green woman', and *dāo shān* 'a mountain of blades' vs *huǒ hǎi* 'a sea of fire'. Both word groups have the same structure. Phonetically speaking, chengyu consist of two parts prosodically, with two disyllabic feet. People tend to pronounce the chengyu with two prosodic feet, by dividing it in the middle. *Suyu*, on the other hand, do not need to be divided in pairs (Wen 2005, p.70). Liu and Xing (2000) have also mentioned a similar concept: In general, chengyu can be divided into two prosodic feet. The first two syllables make up the first prosodic foot and the last two make up the second prosodic foot. They can be seen as two pairs. These two pairs are semantically related on different levels (Liu and Xing 2000, p.77).

Although the inner structure of a chengyu can be divided into two pairs. These pairs can be semantically parallel, butmay also be divided into two pairs by their phonetic features. That means, there is semantical symmetry (structural symmetry) and phonetical symmetry in a chengyu. In this chapter there will be a comparison of the structual features of chengyu and quasi-chengyu.

Looking at the inner structure, there are two main angles: symmetrical and parallel. Symmetrical means that a chengyu can be divided by two pairs from the middle, while parallel means that a chengyu cannot only be divided in the middle, but also that the structure of the two pairs is the same.

Taking "parallelism" as an entry point, a chengyu is categorized into three main branches: parallel, partly parallel or as non-parallel by Liu and Xing (2000). Wen (2005), on the other hand, taking into account both the semantically symmetrical and the phonetically symmetrical possibilities, categorizes chengyu into two main branches: symmetrical chengyuand nonsymmetrical chengyu. All authors further subcategorize every branch by its semantic structure. In this thesis, the main categorization of Liu and Xing (2000) will be applied, but for the analyses of the subcategories Liu and Xing, as well as Wen, will be consulted.

A. *Parallel* chengyu

In a parallel chengyu, both pairs either have the same structure, or the meanings of the pairs may be the same, opposite or similar (Liu and Xing 2000, p.77). The same structure means that both pairs have the same grammatical combination and corresponding words have the same word class. Parallel chengyu can be nominal expressions:

正人　　　　君子
zhèngrén　　jūnzǐ
gentleman　gentleman
Literal: gentleman

Figurative: Same as above.

The two parts are both nouns, and they both contain the meaning of 'gentleman' like symmetrical pairs, which is also the meaning of the whole chengyu.

Parallel chengyu can also be verbs:

敲诈	勒索
qiāozhà	lèsuǒ
blackmail	blackmail

Literal: blackmail
Figurative: Same as above.

Two verbs *qiāozhà* and *lèsuǒ* both have the meaning of 'blackmail'. In this way they form a symmetrical structure as chengyu. In its entirety it means the same as its components.

Sometimes a chengyu is an adjective:

刁钻	古怪
diāozuān	gǔguài
crafty	weird

Literal: Crafty and weird.
Figurative: Same as above.

Both pairs are adjectives, and both have negative meanings: 'crafty' and 'weird'. Because the two words are always used in the same context, they are often bound together as a chengyu. This chengyu can be used to describe someone's personality, or some unusual and difficult problem. For example, *diāozuān gǔguài de wèntí* 刁钻古怪的问题 'a weird question' refers to an uncommon and difficult question with the intention of embarrassing others.

Parallel chengyu can also be phrases. A phrase is a more complicated structure than a word. There is a syntactic connection between the components. Functionally, a two-word-phrase, can be a subject-predicate structure (SP), predicate-object structure (PO), verb-complement structure (VC), modification structure (which consists of a modifier-noun structure (MN) or an adverbial-verb structure (AV)), a serial verb structure (SV), combinative phrases (CP) etcetera. If a parallel chengyu is a phrase, it means that both pairs belong to the same type of structure. For example:

SP + SP

龙	飞	凤	舞
lóng	fēi	fèng	wǔ
Dragon	fly	phoenix	dance

Literal: dragon flies, phoenix dances.

Figurative: Something magnificent and vivid. (Describing someone's calligraphy) vigorous and forceful.

The first and third morphemes, 'dragon' and 'phoenix', are nouns. They both take the role of subject. The second and fourth morphemes, 'fly' and 'dance', are verbs, and they function as predicates. These two pairs have similar meanings;therefore, they are combined to form a parallel chengyu.

PO + PO

安	家	立	业
ān	jiā	lì	yè
install	home	stand	business

Literal: set up a home and build business.
Figurative: settle down and start one's career.

The odd morphemes, *ān* and *lì*, both mean 'to set up', so they both act as a predicate, while the even morphemes, 'home' and 'career', are both the objects of the predicates. Since 'to set up a home' and 'to set up a career' are connected, these two pairs are often combined and used as chengyu.

VC + VC

斩	尽	杀	绝
zhǎn	jìn	shā	jué
chop	exhaust	kill	absolutely

Literal: To chop and kill till (it's) absolutely exhaust
Figurative: completely annihilate

'To chop' and 'to kill' both contain the meaning of 'to annihilate', and they both work as a verb. The complements here, *jìn* and *jué*, both have the meaning of 'up, finished, exhausted', and they show the result of the verbs.

MN + MN

红	男	绿	女
hóng	Nán	lǜ	Nǔ
red	Man	green	woman

Literal: Red man and green woman.
Figurative: Young people who dress fashionably.

'Red' and 'green', two adjectives are modifiers, which are attributive to the nouns – 'man' and 'woman'. 'Red/green man/woman' actually means a man/woman in red/green clothes. This chengyu describes a scene

of young people dressed in different colors walking down the street, and therefore refers to the modern world.

AV + AV

并	驾	齐	驱
bìng	jià	qí	qū
equally	drive	together	drive

Literal: Drive side by side.
Figurative: Neck and neck. The capabilities of the competitors are at the same level. Hard to tell who is stronger.

The even morphemes here originally mean 'to ride a horse' and 'to drive a carriage', because horses and carriages were the main vehicles back in ancient China. Now they both mean 'to drive a car' as well. The odd morphemes both have the meaning of 'on a symmetrical level, together, side by side', which are attributives of the verbs. This chengyu is used to describe two persons/parties/groups of similar capability, making it hard to tell which is superior.

CP + CP

生	死	存	亡
shēng	sǐ	cún	wáng
live	die	exist	Die

Literal: Live and die.
Figurative: On the razor-edge.

In this example, all four components belong to the same category and the meanings are related. The odd components have the similar meaning and the meaning is opposite to the even components.

B. *Partly parallel* chengyu

Some chengyu are not completely parallel, but part of the morphemes in the chengyu is parallel. It doesn't even need to be symmetrical. As long as part of the chengyu is parallel, it is categorized as partly parallel. This can occur in the following situations:

a. One of the two pairs is internally parallel

不	分	黑	白
bù	fēn	hēi	bái
no	divide	black	white

Literal: Cannot distinguish black from white.
Figurative: Cannot tell right from wrong.

In this chengyu, the first pair is not parallel, but the second pair is. 'Black' and 'white' belong to the same word class and are antonyms.

b. The odd/even morphemes are parallel

因	小	失	大
yīn	xiǎo	Shī	dà
because	small	Lose	big

Literal: Lose the big things because of the small ones.
Figurative: Cause a huge loss because to gain some small benefits.

In this chengyu, the even morphemes 'small' and 'big' are from the same word class – adjective – and are antonyms, while the odd morphemes 'because (conjunctive)' and 'lose (verb)' are not parallel.

大	材	小	用
dà	cái	xiǎo	yòng
big	material	small	use

Literal: Big material for small use.
Figurative: Waste one's talent for a petty job.

This chengyu is an example opposite to the one mentioned above. The odd morphemes – still 'big' and 'small' are from the same word class and are antonyms, while the even morphemes, 'material (noun)' and 'use (verb)' cannot form a parallel structure.

C. *Nonparallel chengyu*

Nonparallel chengyu have a stronger integrity in regard to meaning. Liu and Xing (2000) believe that nonpa-rallel chengyu are also not symmetrical. Grammatically speaking, they cannot be combinative phrases, but are always syntactically connected (Liu and Xing 2000, p.78). However, Wen categorizes the chengyu in another way. Although some chengyu are not parallel, they can still be symmetrical, which means the whole chengyu can be divided into two pairs. As for the others, they are not parallel, nor symmetrical. In line with this reasoning, nonparallel chengyu will be divided into two subcategories.

a. Nonparallel symmetrical chengyu

This type is not mentioned by Liu and Xing, but it is by Wen. The chengyu of this type are not parallel, but can still be divided in the middle, into two pairs. Simply put, this kind of chengyu consists of two two-morpheme-words, the two words can be syntactically connected, or can form a combined phrase.

SP

夜郎　　　　　　　自大
yèláng　　　　　　　Zìdà
Yelang　　　　　　　arrogant
Literal: (The King from) Yelang is arrogant.
Figurative: Someone is ignorant and arrogant.

Yelang is the name of an ancient country in China and it refers to the king of Yelang in this chengyu. *Zìdà* as an adjective, plays the role of predicate. Because Yelang and *zìdà* are two individual words, this chengyu is automatically divided into two parts and forms a 'subject-predicate' structure.

VC

粉饰　　　　　　　　　　太平
fěnshì　　　　　　　　　　Tàipíng
sugarcoat　　　　　　　　Peace
Literal: To gloss over the surface and make it look peaceful.
Figurative: To present a false appearance of peace and prosperity.

The first word is a verb and the second is a noun. At first sight it looks like a predicate – object structure. But 'peace' is not the object of the verb 'sugarcoat', but rather a result of the verb. The object, which is omitted, is the dark and turbulent situation. Since 'peace' is the result, this structure is more of a verb – complement structure.

MN

露水　　　　　　　　　　　　　　　夫妻
lùshuǐ　　　　　　　　　　　　　　fūqī
dew　　　　　　　　　　　　　　　spouse
Literal: Spouse of dew.
Figurative: Illicit lovers who are illegally and unofficially acknowledged as man and wife and are temporarily together.

Both words are nouns and the meanings seem unrelated. However, people use the characteristics of *lùshuǐ* to describe the relationship. *Lùshuǐ* 'dew' only appears at night and vanishes as the sun comes up, which indicates that the relationship of the couple can only exist in the darkness. Dew disappears overnight, which also indicates that the relationship cannot last. In this way, *lùshuǐ* works as a modifier. This example is a modifier – noun structure.

CP

浑浑	噩噩
Húnhún	è' è
Profound	Solemn

Literal: profound and solemn.

Figurative (in the past): deep-thinking and serious

Figurative (now): ignorant and unaware

This example is composed by two duplicate words;therefore, it is divided automatically in the middle. Both words belong to the same word class and together they form a compound phrase.

b. Nonparallel nonsymmetrical chengyu

Nonsymmetrical chengyu are neither parallel, nor symmetrical. They cannot be divided into two symmetrical pairs because they are composed of more than two grammatical units. Some have a main syntactical structure, but within the structure they can be further divided into other structures. The following examples show the structures of nonsymmetrical chengyu.

SP

目	不	转	睛
mù	bù	zhuǎn	jīng
eye	no	turn	eyeball

Literal: [Staring] without rolling the eyeball.

Figurative: Very concentrated.

Generally seen, this is an SP structure, in which the subject is 'eye' and the predicate is 'not turn the eyeball'. However, this chengyu cannot be split into two symmetrical parts, because the subject has one component and the predicate has three components— negation, predicate, and object.

PO

暗	送	秋	波
Àn	sòng	qiū	bō
dark	send	autumn	wave

Literal: Send the wave of autumn in the dark.

Figurative: Make secret overtures to someone.

The basic structure of this chengyu is predicate- object: the verb 'to send' and 'what is sent'. Within these two parts, the predicate can be further divided into an adverbial-verb structure – *àn* works as an adverb, meaning 'secretly', and the object can be divided into a modifier (autumn) – noun (wave) structure.

VC

嫁	祸	于	人
jià	huò	yú	rén
marry	misfortune	to	people

Literal: Impute the misfortune to (the other) people.

Figurative: Shift the misfortune to others; cast the blame on someone else.

According to Wen (2005), the predicate is *jià huò* 'to shift the misfortune'. Complementing the predicate, *yú rén* shows to whom the misfortune is shifted. Assessed on a deeper grammatical level, the predicate is also a predicate-object phrase, in which 'misfortune' is the object of the predicate 'to shift'. The complement is a preposition (*yú*) – object (*rén*) phrase. One could also say that this chengyu is a predicate – direct object – preposition object structure. This means that all three grammatical units are on the same level. Whichever the case may be, this is a nonparallel chengyu.

MN

后	起	之	秀
hòu	qǐ	zhī	xiù
later	get up	of	excellent

Literal: Excellent people that get up later.

Figurative: Excellent people that have newly appeared.

The grammatical structure of this example is modifier – noun on the first level, in which 'get up later' is the modifier of 'excellent people'. Because *zhī* 'of' is a particle which is closer to *hòuqǐ*, the components are split into 3:1, which is neither parallel nor symmetrical.

Pivot Construction

指	鹿	为	马
zhǐ	Lù	wéi	Mǎ
point	Deer	for	Horse

Literal: Pointing at a deer and claiming it is a horse.

Figurative: Confound right and wrong.

Pivot construction, jianyu in Chinese, is a structure combined by a predicate-object structure and a subject – predicate structure. The subject in the SP structure is at the same time the object of the PO structure:

In the example above, there is first a PO structure in which 'point' is the predicate and 'deer' the object, while the subject is unknown. The second part is an SP structure lù wéi mǎ 'deer is horse', in which 'deer' is the subject.

Serial Verb Construction

闻	鸡	起	舞
wén	jī	qǐ	wǔ
hear	rooster	get up	dance

Literal: Hear the rooster and get up to (sword) dance.
Figurative: Be diligent.

Serial verb construction is a combination of two or more predicate phrases. Those predicate phrases can also include objects. The predicate phrases do not have to be related per se, but they share the same subject. The structure is as follows: In this chengyu, there are two predicate groups, 'hear the crow', 'get up and sword

(Subject) predicate1 (object)

Predicate2 (object)

…

dance'. All predicates share the same subject, which is omitted from the chengyu. The predicates occur sequentially: the person in question needs to first hear the crow, then get up and practice the sword dance. Therefore, this is also a nonparallel nonsymmetrical structure.

Coordinate Construction

生	老	病	死

| shēng | lǎo | bìng | sǐ |
| Birth | old age | Illness | death |

Literal: Birth, old age, illness and death.

Figurative: refers to childbearing, care for the elderly, medical care and to funerals.

The components of coordinate constructions belong to the same word class and their meanings are also related. But unlike parallel chengyu, it cannot be divided into two pairs, instead, all components are on the same level. In this example, all four components are nouns which are linked to different phases of one's life. Since all four components are seen separately, there is no reason to categorize this type of construction as a parallel or partly parallel chengyu.

D. *Unidentified* chengyu

Beside all the types of chengyu that can be categorized, there are also chengyu for which the structures cannot be analyzed. Some chengyu are from classic literature, thus have been strongly influenced by ancient Chinese. The structures are not clear anymore and cannot simply be analyzed with modern Chinese grammar. Some chengyu have also been influenced by ancient dialect (Liu and Xing 2000, p.78 note 3). For example:

一	败	涂	地
yí	bài	tú	dì
one	fail	smear	earth

Literal: (not grammatical)

Figurative: Complete failure.

精	益	求	精
jīng	yì	qiú	jīng
refined	more	require	refined

Literal: Being refined and required for more refined

Figurative: Constantly strive for perfection.

Liu and Xing (2000) consider these chengyu to be nonparallel chengyu, and in this thesis, these chengyu without clear structure are also seen as nonparallel chengyu. In the research of Liu and Xing (2000), they analyze *Hànyǔ chéngyǔ cíhǎi* 'Chinese chengyu archive', in which there are 32,335 four-morpheme- chengyu. The result indicates that the parallel and partly parallel chengyu make up 40.82% (see table 1) of the total amount. If other four-morpheme phrases in this archive are deleted, the percentage could be higher. This means that parallelism is a noticeable feature of chengyu.

Chengyu	amount	percentage
parallel	12703	39.29%
partly parallel	495	1.53%
non-parallel	19137	59.18%

total	32335	100%

(Table 1. Four-morpheme-chengyu structures from *Hànyǔ* chengyu *cíhǎi* by Liu and Xing.)

From the structural features of real chengyu and their percentages above, I would like to make a comparison with the quasi-chengyu of my collection. Will they be all four syllables? Will there be the same structures in the collection and will the percentage also be similarly distributed? These will be the basis I use to define whether a quasi-chengyu can be seen as chengyu. Judging from the small size of the collection, I do not expect that for every single type of structure there will be an equivalent, or that the percentage will be evenly distributed. As long as the main structure overlaps and the percentage of distribution is similar, it will be seen as a possibility for the quasi-chengyu to become real chengyu.

2.2.2. Syntactic features of Internet quasi-chengyu

A. *Parallel*

As mentioned in paragraph 1.3, if internet quasi-chengyu are seen as a chengyu-like phrase, then it is possible that the inner structures of quasi-chengyu have some similarities with the real chengyu. Indeed, some examples as the following show the features of parallel chengyu.

SP + SP

范	跑	郭	跳
fàn	pǎo	guō	tiào
(name)	run	(name)	jump

Literal: Fàn runs and Guō jumps.
Figurative: Birds of a feather; jackals of the same lair; pot calling kettle black.

Fàn is a high school teacher in Sichuan. When the earthquake started, he was the first who *ran* out of the building, leaving his students behind. Guō is an anchorman, who interviewed Fàn after that incident. During the interview, Guō became furious and *jumped* up to scold him. This expression is used to mock people who try to exaggerate their morality. The first and third morphemes are the subjects, while the second and the fourth ones are the actions of the subjects. Therefore, this is an SP + SP structure.

男	默	女	泪
nán	mò	nǚ	lèi
man	silent	woman	weep

Literal: (After hearing/seeing this,) men become silent and women start to weep.
Figurative: Same as above.

In general, this is also an SP – SP structure. As a noun, *lèi* means 'tear' and as a verb it means 'to weep', so it is reasonable to consider *nǚ lèi* as an SP structure. *Mò*, as an adjective, means 'silent, tacit', whereas as an adverb it means 'privately' and as a verb it means 'to recite from memory'. Although it is also a verb, the meaning is not what the morpheme indicates in this quasi-chengyu. Judging from the meanings of *mò* from different word classes, the *mò* in this quasi-chengyu should be an adjective. Since an adjective can also function as a predicate, *nán mò* can also be seen as an SP structure. In this case, these two pairs are not perfectly parallel, because the predicates are not from the same word class.

MN + MN

林	貌	杨	音
lín	mào	yáng	yīn
(name)	appearance	(name)	sound

Literal: Lin's appearance, Yang's voice.

Figurative: deceit for some reason or goal. Mixing the false with the genuine. Doing vile and immoral things under the name of 'national benefit'.

On the stage of the 2008 Beijing Olympics opening ceremony where the whole world was paying attention, Lin Miaoke began the ceremony with a song. Later, it was found out that the song was not performed live, but recorded in advance. What was even worse was that the recorded song was sung by another girl, Yang. All Lin did was lip-synch the song. Yang had a beautiful voice, but was regarded as not beautiful enough to represent Chinese children. She was then replaced by Lin, who had an advantage in appearance and a privileged political family background, and was considered more suitable as a symbol of China's warm welcome. When this charade was revealed, this expression was created. Seeing that Lin is the possessor of *mào* and Yang is the possessor of *yīn*, these are two 'modifier – noun' pairs which form a parallel MN-MN chengyu.

B. *Partly parallel*

In the collection of internet quasi-chengyu, only one example is seen as partly parallel.

化	危	为	机
Huà	wēi	wéi	jī
convert	danger	for	opportunity

Literal: To change the crisis into opportunity.

Figurative: Same as above.

In this example, the odd morphemes, *huà* (verb) and *wéi* (prep), are not parallel because they do not belong to the same word class. The even morphemes, *wēi* and *jī*, are both nouns, and they are antonyms. Therefore, this is a second type of partly parallel chengyu.

C. *Nonparallel*

a. *Nonparallel symmetrical*

In the collection of internet quasi-chengyu, an appropriate example is as follows:

钓鱼	执法
Diàoyú	zhífǎ
Fishing	enforce the law

Literal: To enforce the law by means of fishing.

Figurative: Entrapment.

This internet quasi-chengyu is made by two predicates, of which the meanings seem unrelated. But just like *lùshuǐ fūqī*, this is an example of using the characteristics of the first word to modify the second. The trick of fishing is to put bait on the hook to attract the fish, while the trick of entrapment is also to lure the criminals into 'swallowing the bait'. So *diàoyú* is used as an adverbial – with the meaning 'in the way of fishing' – and this whole quasi-chengyu is thus an adverbial – verb structure.

The examples above match the structure of nonparallel symmetrical chengyu. However, some examples in the collections are questionable. The problem is either that the syntactical structure is not appropriate, or that the structure is not fixed.

处女　　　嫖娼
chùnǚ　　　piáochāng
virgin　　　prostitute visiting
Literal: Virgin visit prostitute.
Figurative: (not applicable)

This internet quasi-chengyucomes from a sex trafficking case. A girl named Ma was arrested by the local police service, and accused of 'having sex with a prostitute. As Ma denied any criminal wrongdoings, she was mistreated by the policemen and forced to sign a statement of confession. Later she was pre-sentenced to 15 days' detention. The story took a turn after Ma showed a document from the local hospital proving that she was still a virgin, meaning that she could not possibly have committed the crime she was accused of. She was released, but for her emotional trauma was not compensated.

Judging from the syntactic relation of 'virgin' and 'prostitute visiting', this internet quasi-chengyu is a subject – predicate structure. But if we go through the whole story, it is not hard to find that the police agency wanted Ma to confess to 'solicitation', instead of 'visiting a prostitute'. They actually arrested Ma on the suspicion that she was prostituting herself, but instead, they made up an incorrect accusation, which was also proven false. Ma proved that she was still a virgin, thus proving that she could not possibly have prostituted herself. Therefore, although the structure of this internet quasi-chengyu is in line with the real chengyu, the meaning is not in line with the story. For this reason, I doubt whether this should be categorized as an internet quasi-chengyu.

b. *Nonparallel nonsymmetrical*

Within the collection of chengyu presented in this thesis, there are also examples of nonsymmetrical chengyu. In the earlier section an internet quasi-chengyu is mentioned about someone who did a long examination in order to prove that he has an occupational disease. This quasi-chengyu is an example of a serial verb construction. First, this quasi-chengyu can be divided into two predicate-object structures: 'open' – 'chest' and 'test' – 'lung'. Second, because these two PO structures are not parallel ('open' and 'test' are semantically unrelated in this context), this structure cannot be seen as a parallel chengyu. Furthermore, these two predicates have a sequential relation: the chest first needs to be opened, then the lung can be tested. Thus, the PO structures are not symmetrical, but rather serial. Last, these two predicates have the same subject which is omitted here. In conclusion, this is a nonparallel nonsymmetrical chengyu.

开	胸	验	肺
Kāi	xiōng	yàn	fèi
Open	chest	test	lung

Literal: Opening the chest to exam the lungs.

Figurative: The difficult situation where occupational disease patients have to protect their rights.

The most common structure in this collection of internet quasi-chengyu is the nonsymmetrical SP structure:

欧阳	挖	坑
Ōuyáng	wā	kēng
(name)	dig	hole

Literal: Ouyang digs a hole.

Figurative:Celebrities giving irresponsible and careless speeches.

In 2007, a photo of the moon taken by a Chinese satellite was released by the Chinese space agency. The authenticity of this photo was soon questioned, as it appeared to be a plagiarized photo taken by NASA. In order to prove its authenticity, Ouyang, top academician of Chinese Academy of Sciences, pointed out that there were two holes on the surface of the moon on the Chinese photo, but only one hole on the photo taken by NASA. The reason might be that the NASA photo has a lower resolution, or that the second hole came into being after the first photo was taken. However, a complete match line on the Chinese photo was found, and the second hole seems to be a mistake by the one who processed the photo. 'The second hole is dug by Ouyang', internet users joked. The main structure is subject – predicate, which is the most regular pattern for the quasi-chengyu, the predicate part can be further divided into a predicate – object structure.

D. *Unidentified Chengyu*

In the collection of internet quasi-chengyu, there are also examples of which the structure is unclear or ungrammatical. They are mostly extracted from a whole sentence, such as a sentence from song lyrics. For example:

人	艰	不	拆
Rén	jiān	bù	chāi
People	hard	no	break up

Literal: People are hard, don't break them up.

Figurative: Life is already so hard, you do not need to expose everything.

This example already appeared in the earlier section. The full sentence is a sentence from a song, which reads as follows:

人			艰				不	拆

人生	已经	如此	艰难，	有些	事情	就	不要	拆穿
rénshēng	yǐjīng	rúcǐ	jiānnán	yǒuxiē	shìqíng	jiù	bùyào	chāichuān
Life	already	such	hard	some	thing	just	no	expose

Literal: Life is already so hard; you do not need to expose everything.

Figurative: Same as above.

The structure of this example is considered to be unclear, mainly because the meanings of the chosen component are grammatically not connected. *Rén* means 'person', 'people', 'others', 'man', 'adult' and so on, but cannot represent the meaning of 'life'. *Chāi* means 'open' or 'break up', and it cannot represent the meaning of 'expose'. Therefore, if someone sees this quasi-chengyu, but has no knowledge of the original sentence, this quasi-chengyu is meaningless to him. The internet users have picked four words from the sentence which carry the core meaning – 'life', 'hard', 'no' and 'expose', and took out one morpheme from each word to combine them together, like a piece of telegraph text. Since the quasi-chengyu is limited to four morphemes and the four chosen morphemes cannot fully represent the original meaning, the structure becomes unclear.

The internet quasi-chengyu in this thesis are analyzed based on the research conducted by Liu and Xing (2000). Phrases that are part of this collection, but for which it is questionable whether they are quasi-chengyu, are also considered nonparallel (see table 2).

Chengyu	Amount	percentage
parallel	13	24.07%
partly parallel	2	3.70%
non-parallel	39	72.22%
total	54	100%

(Table 2. Structures of internet quasi-chengyu in the collection of this thesis.)

Table 2 shows that there are 54 internets quasi-chengyu in total, among which 13 are parallel and 2 are partly parallel. Thirty-nine are nonparallel (nonparallel symmetrical, nonparallel nonsymmetrical or unidentified), which make up 77.05% of the whole amount. This result is not consistent with the research of Liu and Xing.

	Liu and Xing		This thesis	
chengyu	amount	percentage	amount	percentage
parallel	12703	39.29%	13	24.07%
partly parallel	495	1.53%	2	3.70%
non-parallel	19137	59.18%	39	72.22%
total	32335	100%	54	100%

(Table 3. Comparison between table 1 and table 2.)

In this comparison, the percentages of the two researches are not exactly in line, but the main trend is the same. Nonparallel chengyu make up the biggest portion, partly parallel the smallest, and parallel make up roughly one-third or one-fourth of the total. Parallelism is a feature for internet quasi-chengyu, but not as noticeable as for real chengyu.

2.3. Summary

In this chapter, idiom-related terms have been introduced and compared. The four-morpheme chengyu-like phrases are defined as internet quasi-chengyu. In order to find out the syntactic difference and similarities between this type of quasi-chengyu and real chengyu, the inner structure of both types need to be analyzed. In general, the inner structures of real chengyu and quasi-chengyu are similar. For some structures of real chengyu, equivalents can be found in the quasi-chengyu collection. Compared to the earlier data research of real chengyu conducted by Liu and Xing, there are some similarities between internet quasi-chengyu examined in this thesis and real chengyu, such as parallelism. The percentage distribution is not exactly equal, but the trend that non-parallel makes up the biggest portion of the collection, then parallel, and partly parallel the smallest, is alike.

3. Phonological features of real chengyu and internet quasi-chengyu

In this chapter, I would like to introduce the phonological features of Mandarin first. As mentioned in the previous chapter, chengyu can be divided into two pairs, according to structural features and qua phonetic features. In the previous chapter I discussed thestructural features of chengyu. From the examples in the last chapter we can also see that in Mandarin, bi-morphemic words are very common. Although many words are not bi-morphemic, people tend to combine two morphemes. This is according to Duanmu (2014) a constraint that requires words to be composed of a prosodic "foot" consisting of stressed and weak syllables. The most common foot is with two syllables (p.291). Wu (2003) also mentions that a phrase with four syllables is typically divided into a 2+2 structure (p.101). The aim of this chapter is to point out the phonological features of real chengyu, compared tointernet quasi-chengyu. This chapter will first take a look at the tone structures of the real chengyu, then at the tone structures of the quasi-chengyu collection. The criteria whether a quasi-chengyu can be seen as chengyu is whether the tone structures are in line, and whether the percentage distribution is close.

3.1. Phonological features of Mandarin

The smallest unit of phonology is a phoneme, which is combined with other phonemes to form semantically meaningful units, i.e. the morphemes. The phoneme is described as the "smallest contrastive linguistic unit which may bring about a change of meaning" (Gimson 2008, p.41). The difference between the English words *bat* and *bag* is a result of the exchange of the phoneme /t/ for the phoneme /g/. Two words that differ in meaning through the change of only a single phoneme make a minimal pair. The system of phonemes contains vowels and consonants, and in different languages the number of vowels and consonants are different. For example, in Mandarin there are no consonants such as /ð/ or /Ө/, but they exist in English. In Mandarin, there are 24 consonants and 10 vowels (see table 3).

| Consonants	b /p/, p /p'/, m /m/, f /f/, (w /w/), d /t/, t /t'/, n /n/, l /l/, g /k/, k /k'/, h /x/, j /tɕ/, q /tɕ'/, x /ɕ/, (y /j/), zh /tʂ/, ch /tʂ'/, sh /ʂ/, r /ʐ/, z /ts/, c /ts'/, s /s/, ng /ŋ/
vowels	a /ɑ/, o /ɔ/, e /ɤ/, i /i/, e /ê/, u /u/, ü /y/, i /ʅ/, i /ɿ/ (retroflex), r /ɚ/

(Table 3. Consonants and vowels in Mandarin. Summarized based on Wu (1992).)

Furthermore, Mandarin is a tone language. Tones are also part of the Mandarinphonological system and are phonemic. The important role that the tones play will be discussed later.

A combination of phonemes can form a syllable. A syllable is a prosodic unit of a word which may divide the word into several parts. For example, the English word *water* is composed of two syllables: *wa* and *ter*. The essential component of a syllable is a nucleus (which is often a vowel), with optional onsets or codas in most languages (which are consonants). If a word consists of only a single syllable it is called a monosyllabic word, such as *hat*; a word which consists of two syllables is called a disyllabic word, such as *chil – dren*; a trisyllabic word is a word with three syllables, such as *at-ten-tion*; if a wordhas more than three syllables, it is then called a polysyllabic word. In Mandarin most words are composed of two syllables.

The structure of the syllable can be described as containing two parts: onset and rime. An onset can consist of a consonant, a consonant cluster, or null. It is obligatory in some languages, optional or forbidden in some other languages. The rime can be further split into the nucleus and the coda. The Nucleus is always a vowel

or syllabic consonant, and it is obligatory in most languages. The Coda is also a consonant or consonant cluster. In some languages a coda is optional, but in some other languages it is highly forbidden. The structure is like in picture 2:
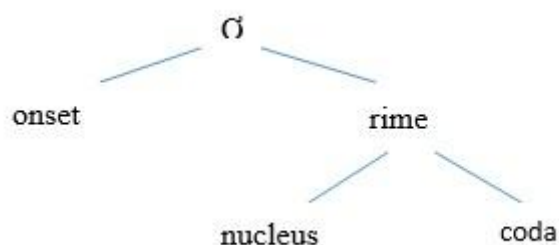
(Picture 2. Structure of a syllable.)

Researchers in ancient China already used syllable structures to analyze traditional Chinese. The idea of Mandarin syllable structure is similar to the typical syllable structure, but the rime can be further split into medial, nucleus and coda. In Mandarin, the onset is optional, and if it exists, it is a consonant. The medial and coda are also optional. The medial is often a glide or semi-vowel (i, u, ü) and the coda can be a consonant (n, ng) as well as a vowel (i, u). The nucleus is obligatory and is always a vowel. There are no consonant clusters in Mandarin. For example, the syllable chuāng 窗 'the window' has four segments: *ch* is the onset,
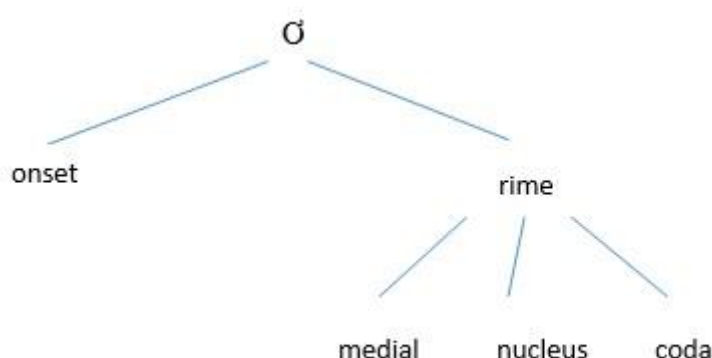
uang is the rime, in which the medial is *u*, the nucleus is *a*, and the coda is *ng*.

(Picture 3. Syllabic structure of Mandarin.)

Although onsets are always consonants, they cannot be considered as equivalent to consonants. Not all consonants can be onsets, for example ng /ŋ/. There are 23 onsets in Mandarin. Although codas can also be consonants only ng /ŋ/ and n /n/ can be used as coda. A Mandarin syllable can have up to all four segments, or be limited to the nucleus.

The Syllable constitutes the phonological prosodic unit. Besides this, the length, strength and pitch of the syllable can also affect the language practically. In Mandarin, length does not cause a difference in meaning. A [i:] and a [i] are allophones. Sometimes strength can affect the meaning; if a syllable is pronounced weakly, the tone becomes unnoticeable, and the meaning can change as well. Pitch plays an essential role in distin-

guishing meanings in Mandarin. The height of a pitch and the way it changes constitute the tones, and they can make two words with otherwise identical structure have enitrely different meanings.

There are four tones in Mandarin: high, rising, low and falling. High is a flat high tone, the starting point and the endpoint are on the same pitch level. Rising starts from a middle pitch and rises to the high point. Low starts from a middle pitch, first goes down to reach a low point, and then rises again to a higher level. Falling starts from a high level, and drops all the way down to the lowest point (see picture 4). There is also a "weak tone" or "zero tone", which is when a syllable is pronounced without a noticeable tone. It doesn't have a certain pitch level, and its pitch is affected by the tone of the previous syllable. A toneless syllable cannot exist independently nor can it be the first syllable of a word. In the Pinyin system, the tones are marked above the nucleus. Taking *ma* as example, the four tones will be marked as: *mā* (high), *má* (rising), *mǎ* (low) and *mà* (falling)
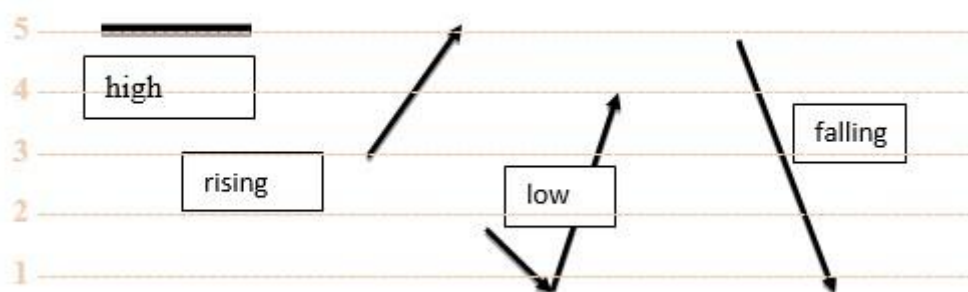
(Picture 4. Tones in Mandarin.)

According to Liu (2004), the four tones in modern Mandarin have evolved from "level", "low", "falling", and "entering" in the old Chinese. High and rising are also called "even" tones, while "low" and "falling" belong to "oblique" tones. Even tones and oblique tones form a binary opposition (Liu 2004, p.45). The alternation of even and oblique tones, together with rhyme, are important factors in Chinese poetry.

Although all four tones have their own pitch level, a tone can be affected by the tones of its adjacent syllable; the pitch of the first tone then changes. This phenomenon is called tone sandhi. Sometimes it is only a matter of the affected tone becoming relatively higher or lower, but in some other situations, the affected tone changes so much that it becomes another tone, and may change from even to oblique, or the other way around. For example, if two low tones are together, the first one changes to a rising tone, so it changed from oblique to even. Some morphemes change their tone before certain tones, such as yī 一 'one' which changes to rising if it is before a falling and changes to falling if it is before a rising tone.

Another common factor is the prosodic "foot", as mentioned at the beginning of this section. A standard foot most often consists of two syllables (Yu 2012, p.64). One syllable is lightly stressed and the other is stressed. This exchange of stress adds flow and melody to poetry or phrase, and also to chengyu. A standard foot is also a standard prosodic word. Two prosodic words make a composite prosodic word (Yu 2012, p.64). Rele-

vant tot his thesis is that a chengyu is composed of four syllables forming two feet. Therefore "foot" and "prosodic word" might also be important factors of a chengyu.

3.2. Tone patterns in chengyu

All chengyu have four syllables, so there are diverse ways of distributing even and oblique tones. According to Liu (2004), there are 16 types of even-oblique distribution patterns comprising 5 categories. In this thesis, even tones are indicated with "O" and oblique tones with "X". If there is a sandhi in the chengyu, such as two low tones next to each other, it will be analyzed by the surface form rather than the supposed underlying form. As in this case, it will be analyzed as rising – low, thus an OX pattern.

3.2.1   Opposite pattern

In this category, the amount of O and X tones are the same, but the two kinds of tones don't alternate. There are two possible types: OOXX and XXOO. In the previous chapter, it was mentioned that chengyu can be seen as parallel from the perspective of phonology. In this category, the syllables are split into two pairs: the "even" and "oblique" syllables are related.

风	花	雪	月
fēng	huā	xuě	yuè
wind	flower	snow	moon

Literal: wind, flower, snow and the moon
Figurative: Literature with empty content. Dissolute life full of love affairs and entertainment.

This is an OOXX pattern, in which the tones are: high – high – low – falling.

剑	走	偏	锋
Jiàn	zǒu	piān	fēng
Sword	walk	side	blade

Literal: The sword walks to the side of the blade.
Figurative: Take an unusual way.

This is an XXOO pattern, in which the tones are: falling – low – high – high.

3.2.2   Alternating pattern

In this category there are thesame number of "even" and "oblique" syllables and they appear alternately. There are two types: OXOX and XOXO. Either O-X combination makes up a pair, so the chengyu is also parallel.

An example of OXOX is as follows:

七	上	八	下
qī	shàng	bā	Xià

seven   up   eight   down
Literal: Seven up and eight down.
Figurative: feeling perturbed.

The tones are: high – falling – high – falling.

作　奸　犯　　科
zuò　jiān　fàn　　Kē
do　evil　commit　　article of the law
Literal: To do evil things break the law.
Figurative: Same as above.

The tones of this chengyu are: falling – high – falling – high. Therefore, it is an XOXO pattern.

3.2.3　Looping pattern

Looping means that the whole pattern is a circle such that the first syllable and the last syllable belong to one category while the middle two syllables belong to the other. Again, there are two types: OXXO and XOOX. This mirrored structure makes the chengyu also symmetrical.

万　　　　　人　　　空　　　巷
wàn　　　　rén　　　kōng　　xiàng
ten thousand　　people　　empty　　lane
Literal: Then thousand people (leave) empty lanes.
Figurative: A huge public celebration.

Tones: falling – rising – high – falling. Therefore, it is an XOOX pattern.

纲　　　　　举　　目　　　张
gāng　　　　jǔ　　mù　　　zhāng
scheme　　　lift　category　open
Literal: Lift the schema and open the category.
Figurative: Well organized, logical (eg. An article)

The tones are: high – low – falling – high, thus an OXXO pattern.

3.2.4　Irregular pattern

The above three categories are the basic categories, or "totally harmonious" (Liu 2004, p.45). In the irregular pattern, there are both Os and Xs, but they are not evenly distributed. There are 8 types:

OOOX

	人	中	龙	凤
	rén	zhōng	lóng	fèng
	people	middle	dragon	phoenix

Literal: A dragon or phoenix of people.
Figurative: An outstanding person.

Tone	**rising**	**high**	**rising**	**falling**

XXXO

	九	死	一	生
	jiǔ	sǐ	yì	shēng
	nine	die	one	live

Literal: Nine die and one lives.
Figurative: (survive from) an urgent situation.

Tone	**low**	**low**	**falling**	**high**

In this chengyu, because *yi* appears before a high tone, it is pronounced as a falling, so an X pattern. Therefore, this chengyu has an XXXO pattern.

OXOO:

	相	辅	相	成
	xiāng	fǔ	xiāng	chéng
	mutual	supply	mutual	complete

Literal: Mutually supply and mutually complete.
Figurative: Things that are mutually supporting each other and cannot be missed.

Tone	**falling**	**low**	**falling**	**rising**

XOXX:

	五	湖	四	海
	wǔ	hú	sì	hǎi
	five	lake	four	sea

Literal: Five lakes and four seas.
Figurative: All around the country/world.

Tone	**Low**	**rising**	**falling**	**low**

OOXO:

形	单	影	只
xíng	dān	yǐng	zhī
shape	single	shadow	single

Literal: Single shape and single shadow.

Figurative: Being alone.

Tone	**rising**	**high**	**low**	**high**

XXOX:

指	鹿	为	马
zhǐ	lù	wéi	mǎ
point	deer	as	horse

Literal: Point the deer as the horse.

Figurative: Mix the right with the wrong.

Tone	**Low**	**falling**	**rising**	**low**

OXXX:

欺	上	罔	下
qī	shàng	wǎng	xià
deceive	up	hide	down

Literal: Deceive from above and hide from the under.

Figurative: Lie to the higher authorities for their trust, and hide the truth to the public.

Tone	**high**	**falling**	**low**	**falling**

XOOO:

水	滴	石	穿
shuǐ	dī	shí	chuān
water	drop	stone	through

Literal: Water drops through the stone.

Figurative: Success will eventually be achieved with cumulative efforts.

Tone	**Low**	**high**	**rising**	**high**

3.2.5 *Special patterns*

This category consists of two types: OOOO and XXXX. There is no exchange of even-oblique patterns, but the tones often vary within the patterns, such as a high tone combined with a rising tone, or a low tone combined with a falling.

作	茧	自	缚

zuò	jiǎn	zì	fù
make	cocoon	self	tie

Literal: Make a cocoon to tie oneself.
Figurative: Make troubles for oneself.

Tone	**falling**	**low**	**falling**	**falling**

人	微	言	轻
rén	wēi	yán	qīng
people	micro	speech	light

Literal: Micro people's speech (weighs) light.
Figurative: If people have no authority, their words will not be sincerely considered.

Tone	**rising**	**high**	**rising**	**high**

3.3. Data and comparison between chengyu and internet quasi-chengyu

Based on the categorization of paragraph 3.2, Liu (2004) made an analysis from *Zhonghua Chengyu Cihai*, 中华成语辞海 'Chinese chengyu archive' and the percentages are as follows (see table 4):

1	OOXX	4408	13.63%
2	OOOX	3584	11.08%
3	XXOO	3394	10.50%
4	OXOO	3109	9.61%
5	XOOX	2454	7.59%
6	XOXX	1923	5.95%
7	OOXO	1855	5.74%
8	OXOX	1747	5.40%
9	OXXO	1492	4.61%
10	XOOO	1488	4.60%
11	XOXO	1395	4.31%
12	OOOO	1327	4.10%
13	XXOX	1262	3.90%
14	XXXO	1225	3.79%
15	OXXX	983	3.04%
16	XXXX	689	2.13%
Total		32335	100.00%

(Table 4. Liu's (2004) research of *Zhonghua Chengyu Cihai*, percentage ranking from downwards, specified to two decimal places.)

Three of the top five patterns – OOXX, XXOO and XOOX – belong to the basic categories. The other two patterns are irregular, but both with a –X–O or –O–X pattern. As an explanation, Liu applied the theory of Wu

(1992): A four-morpheme- chengyu has two feet. Therefore, the last syllable of either foot is important while the tone of the first and third syllable of a chengyu is not (p.129-135). This opinion is consistent with a guideline of the rhyme of seven-character- poetry: *yī sān wǔ bú lùn, èr sì liù fēn míng* 一三五不论，二四六分明 "the first, third and fifth syllable can be ignored, while the second, fourth and the sixth should be clear" (Wu 2003, p.101). In this way, all the top five patterns can be seen as a variation of the basic categories: –O–X, –O–X, –X–O, –X–O and –O–X. The two special patterns – OOOO and XXXX – take up relatively low percentages. But because there are both highs and risings in an OOOO pattern, the rhythm has more fluctuation than a low – falling variation in an XXXX pattern, and therefore OOOO has a higher percentage than XXXX (Liu 2004, p.48). However, this is not totally true. A closer look at the results shows that some –O–O or –X–X patterns are higher ranked than some –X–O or –O–X patterns. Furthermore, patterns with more than two X rank low, from the table we can see that they rank among the last four places. Meter is important for a chengyu, but that is not the only explanation for the ranking. Attention will be paid to the –X-O and –O-X patterns, but the analysis will not be purely subdivided by the even syllables of the material. Combing the categories by the "foot", it can be seen that the –X–O and –O–X patterns do have a higher occurrence (Table 4.1).

1	–O–X	12369	38.25%
2	–X–O	9220	28.51%
3	–O–O	6065	18.76%
4	–X–X	4681	14.48%
Total		32335	100%

(Table 4.1 Liu's (2004) study collapsed by foot patterns.)

Is this pattern unique to chengyuor does it also appear in other kinds of four-morpheme phrases? In order to verify the phonetic patterns of chengyu, Liu did another sample study. The research material is based on literature, news, commentaries and social science articles (around 2.4 million words in total). Four-morpheme phrases have been chosen, including existing chengyu, four-morpheme phrases that are constantly used in their entirety (such as terms and fixed phrases), parody idioms (such as *bì ér yuǎn zhī* 'stay away from something'), and four-morpheme free phrases with a symmetrical structure. The results are shown in table 5.

1	OOOX	524	11.49%
2	OXOO	461	10.11%
3	OOXX	439	9.63%
4	XXOO	388	8.51%
5	XOXX	317	6.95%
6	XOOX	313	6.86%
7	OOXO	262	5.75%
8	OOOO	253	5.55%
9	OXXO	247	5.42%
10	XXXO	241	5.29%
11	XOOO	224	4.91%
12	OXOX	215	4.71%
13	XXOX	196	4.30%
14	OXXX	180	3.95%
15	XOXO	177	3.88%
16	XXXX	123	2.70%

Total		4560	100.00%

(Table 5. Liu's (2004) random sample research, percentage ranking downwards, specified to two decimal places.)

1	–O–X	1593	34.93%
2	–X–O	1337	29.32%
3	–O–O	916	20.09%
4	–X–X	714	15.66%
Total		4560	100%

(Table 5.1 Collapse version of the data of Table5.)

The ranking of Table 5 is different from Table 4, but the top six patterns have the –X–O and –O–X patterns. For Table 5.1, it is also clear that –O–X and –X–O pattern appear more often than the other two patterns. However, the patterns with three X don't stay at the bottom anymore, XOXX even reaches the top five. This means, four-morpheme phrases and chengyu are not totally agreed on the phonetic patterns. Will the phonetic pattern of internet quasi-chengyu be closer to either of them? Further comparison will be shown in the later paragraphs.

Another study has been doneby Cheng (2011), in which she analyzed a different chengyu dictionary *Hanyu Chengyu Cidian*. Among 8004 chengyu of this dictionary, her results are as shown in table 6.

1	OOXX	1093	13.66%
2	OOOX	888	11.09%
3	XXOO	842	10.52%
4	OXOO	769	9.61%
5	XOOX	607	7.58%
6	XOXX	476	5.95%
7	OOXO	459	5.73%
8	OXOX	432	5.40%
9	OXXO	369	4.61%
10	XOOO	368	4.60%
11	XOXO	345	4.31%
12	OOOO	328	4.10%
13	XXOX	312	3.90%
14	XXXO	303	3.79%
15	OXXX	243	3.04%
16	XXXX	170	2.12%
Total		8004	100.00%

(Table 6. The research of Cheng (2011) of *Hanyu chengyu cidian*, according to percentage ranking in descending order, specified to two decimal places.)

1	–O–X	3065	38.29%

2	–X–O	2283	28.52%
3	–O–O	1500	18.74%
4	–X–X	1166	14.57%
Total		8004	100.00%

(Table 6.1 Collapsed version of the data of Table 6.)

Table 6 indicates a strong agreement with table 4. The top 5 ranking patterns are the same, which seems to confirm Liu's explanation. Moreover, the last four patterns are the ones with more than two X, which is also in agreement with the results of Liu. The percentage of different foot patterns in Table 6.1 also strongly agrees with table 4.1.

Since the internet quasi-chengyu are also four-morpheme structures, will they also have the phonological features of the real chengyu, or will they be phonetically closer to general four-morpheme phrases? To answer this question, I have looked at the internet quasi-chengyu collection. The results are indicated as in table 7.

1	XOOX	9	16.67%
2	OXOX	6	11.11%
3	OOOX	6	11.11%
4	OXOO	5	9.26%
5	OOXO	4	7.41%
6	XOXX	4	7.41%
7	XXXO	3	5.56%
8	XOOO	3	5.56%
9	OXXX	3	5.56%
10	XXOX	2	3.70%
11	OOOO	2	3.70%
12	OOXX	2	3.70%
13	XXOO	2	3.70%
14	XOXO	1	1.85%
15	OXXO	1	1.85%
16	XXXX	1	1.85%
Total		54	100.00%

(Table 7. Data research of this thesis. Percentage ranking downwards, specified to two decimal places.)

1	–O–X	21	38.89%
2	–X–X	12	22.22%
3	–X–O	11	20.37%
4	–O–O	10	18.52%
Total		54	100.00%

(Table 7.1 Collapsed version of this collection.)

As indicated by table 7, the –O–X and –X–O patterns don't take the dominant higher places anymore. An –X–X pattern is in the second place. Patterns with more than two X are scattered, which is also not in line with the two researches on chengyu. The collapsed version shows that the ranking of foot patterns is different

from what we found in the other studies. Although different, the percentage of the last three patterns are close to each other. Only the –O–X pattern has a significant high percentage, which is agreed to the other studies.

	Liu 1		Liu 2		Cheng			This thesis	
1	OOXX	13.63%	OOOX	11.49%	OOXX	13.66%	1	XOOX	16.67%
2	OOOX	11.08%	OXOO	10.11%	OOOX	11.09%	2	OXOX	11.11%
3	XXOO	10.50%	OOXX	9.63%	XXOO	10.52%		OOOX	11.11%
4	OXOO	9.61%	XXOO	8.51%	OXOO	9.61%	4	OXOO	9.26%
5	XOOX	7.59%	XOXX	6.95%	XOOX	7.58%	5	OOXO	7.41%
6	XOXX	5.95%	XOOX	6.86%	XOXX	5.95%		XOXX	7.41%
7	OOXO	5.74%	OOXO	5.75%	OOXO	5.73%	7	XXXO	5.56%
8	OXOX	5.40%	OOOO	5.55%	OXOX	5.40%		XOOO	5.56%
9	OXXO	4.61%	OXXO	5.42%	OXXO	4.61%		OXXX	5.56%
10	XOOO	4.60%	XXXO	5.29%	XOOO	4.60%	10	XXOX	3.70%
11	XOXO	4.31%	XOOO	4.91%	XOXO	4.31%		OOOO	3.70%
12	OOOO	4.10%	OXOX	4.71%	OOOO	4.10%		OOXX	3.70%
13	XXOX	3.90%	XXOX	4.30%	XXOX	3.90%		XXOO	3.70%
14	XXXO	3.79%	OXXX	3.95%	XXXO	3.79%	14	XOXO	1.85%
15	OXXX	3.04%	XOXO	3.88%	OXXX	3.04%		OXXO	1.85%
16	XXXX	2.13%	XXXX	2.70%	XXXX	2.12%		XXXX	1.85%
Total		100.00%		100.00%		100.00%			100.00%

(Table 8. Comparison of tables 4 through 7. Tables 4 to 7 are named respectively Liu 1, Liu 2, Cheng, and This thesis.)

	Liu 1		Liu 2		Cheng		this thesis	
1	–O–X	38.25%	–O–X	34.93%	–O–X	38.29%	–O–X	38.89%
2	–X–O	28.51%	–X–O	29.32%	–X–O	28.52%	–X–X	22.22%
3	–O–O	18.76%	–O–O	20.09%	–O–O	18.74%	–X–O	20.37%
4	–X–X	14.48%	–X–X	15.66%	–X–X	14.57%	–O–O	18.52%

to-tal		100.00%		100.00%		100.00%	100.00%

(Table 8.1 Collapsed version of table 8.)

As seen in table 8, the difference between this thesis and the other studies is clear to see. In the first three studies, although the material and the percentages are not the same, the rankings are similar. The top four rankings are the same patterns, even though the ranking of each study is not exactly the same. For chengyu and four-morpheme phrases, the pattern with the lowest ranking is XXXX. That means, the results of these different material are very consistent. However, the fourth study, mine on the internet quasi-chengyu, is not in agreement with the rest. First of all, the ranking of this thesis hugely differs from the other three studies. The top pattern, XOOX, which makes up 16.67% of the whole collection, ranks relatively low in the other researches. The percentages are only 7.59%, 6.86% and 7.58%. Second, only two of the top four patterns from the other three studies are in the top four of this thesis while the other two rank low: both only 3.70% of the whole collection. Third, even though XXXX also ranks the lowest in this thesis, the other two patterns with the lowest ranking as well – XOXO and OXXO – rank relatively higher in the other studies.

When comparing the foot patterns, it can also be seen that the fourth study is not consistent with the other three. In the first studies, the percentage of –X–O and –O–X patterns surpasses that of–X–X and –O–O patterns. However, in the fourth study, three of the four patterns have more or less the same percentage. In all four studies, –O–X pattern takes the obviously higher percentage, which may show some consistency.

Furthermore, when comparing the difference between chengyu and internet quasi-chengyu, there is more similarity between the results of this thesis and Liu's sample research. As mentioned before, the ranking of –X–O and –O–X patterns are more dispersed, and the ranking of patterns with more than two X are also not regular. The reason can be that in the sample research of Liu, there are not only chengyu, but also other types of four-morpheme phrases. All these factors suggest that the tone patterns of internet quasi-chengyu do not completely follow the path of realchengyu. In the previous chapter it is said that internet quasi-chengyu belong to an intermediate phase between four-morpheme phrases and chengyu, while developing towards to chengyu. Therefore, the internet quasi-chengyu may be closer to four-morpheme phrases, which have a broader range, and then reach towards real chengyu, which have stricter criteria. During this process, some quasi-chengyu may be eliminated because they do not meet the criteria of a real chengyu.

3.4. Other phonological features of chengyu and internet quasi-chengyu

Aside from tonesthere are other features of a syllable that are of importance to chengyu. For example, a regular repetition of the same phoneme is said to make a chengyu more melodyious. There are three styles of repeating of phonemes found in chengyu: alliteration, rhyme and repetition (Liu and Xing 2003, p.50).

3.4.1 Alliteration

Alliteration means repetition of the onset of a series of multiple words, such as the English example *Peter Piper picked a* peck *of pickled peppers* where there is a repetition of /p/. Alliteration in Chinese does not limit itself to "words": the repetition of onsets of a series of multiple syllables/morphemes is also a productive form of alliteration. Since a chengyu has four syllables, there are multiple types of alliteration.

琳	琅	满	目
lín	**láng**	**mǎn**	**mù**
jade		full	eye

Literal: Full of jades in the eyes.

Figurative: An eyeful of goodies.

Both pairs are alliterating, as seen from the example above, the onsets l- l and m- m are both alliteration.

得	天	独	厚
dé	**tiān**	**dú**	**hòu**
get	sky	unique	thick

Literal: Get the unique thickness from the sky.

Figurative: Be richly endowed by nature.

Alliteration can also happen when the odd/even syllables share the same onset, such as d- t- d- h in the example above.

千	秋	万	代
qiān	**qiū**	**wàn**	**dài**
thousand	autumn	ten thousand	generation

Literal: Thousands of autumns and generations.

Figurative: Throughout the ages.

Alliteration also exists when only one pair of onsets are the same, as in q- q- x- y. If the onsets of the middle syllables are the same, or the onsets of the head-tail syllables are the same, it also forms alliteration, such as x- q- q- y in the example below.

自	取	其	辱
zì	**qǔ**	**qí**	**rǔ**
self	get	one's	embarrassment

Literal: Get one's own embarrassment.

Figurative: Bring disgrace on one's own head.

3.4.2   Rhyme

While alliteration in Chinese is a repetition of onset, rhyme is the repetition of rimes. Similar rimes (such as *ian* and *uan*) can also rhyme.

An at the head and tail syllable make a rhyme.

烟	消	云	散
yān	**xiāo**	**yún**	**sàn**
smoke	demolish	cloud	dismiss

Literal: Smoke demolish and clouds dismiss.
Figurative: Completely vanish.

玉	洁	冰	清
yù	**jié**	**bīng**	**qīng**
jade	clean	ice	clear

Literal: Clean like jade and clear as ice.
Figurative: Pure-hearted.

Rhyme can exist within one pair of a chengyu such as u- ie- ing- ing; if both pairs rhyme, it is called a double-rhyme, such as an- an- uo- o in the example below.

寒	酸	落	魄
hán	**suān**	**luò**	**Pò**
scrubby		sorehead	

Literal: Scrubby and sorehead.
Figurative: Same as above.

It is even possible for all four syllables to rhyme:

断	简	残	篇
duàn	**jiǎn**	**cán**	**piān**
broken	bamboo slip	incomplete	page

Literal: Broken bamboo slips and incomplete pages.
Figurative: Stray fragments of text.

3.4.3   Repetition

Repetition in Chinese means that the same syllable is repeated. Repetition can happen within one pair:

惺	惺	相	惜
xīng	**xīng**	**xiāng**	**xī**
clearheaded	clearheaded	mutual	cherish

Literal: Clearheaded people cherish each other.
Figurative: People who have the same personalities, hobbies or circumstances sympathize with each other.

It can also happen in both pairs:

朝	朝	暮	暮
zhāo	**zhāo**	**mù**	**mù**
morning		night	

Literal: Morning and night.

Figurative: For days and days.

Sometimes a repetition of syllables that do not belong to the same pair can also make repetition.

人	山	人	海
rén	**shān**	**rén**	**hǎi**
people	mountain	people	sea

Literal: Mountains and seas of people.

Figurative: A lot of people.

Sometimes, alliteration, rhyme and repetition can exist in combination in a chengyu, such as *xīng xīng xiāng xī*, is actually a combination of alliteration and repetition.

Because of the limited number of Chinese onsets and rimes, and the fact that every morpheme is only made up of one syllable, it is easy to make alliteration or rhyme. Liu and Xing (2003) did a research about four-morpheme chengyu and find out that alliteration, rhyme and repetition in chengyu is common (see table 9).

Type	quantity	percentage
alliteration	4516	13.97%
rhyme	9423	29.13%
alliteration – rhyme combination	2451	7.58%
repetition	1841	5.69%
total	18231	56.37%

(Table 9. Percentage of phonetic repetition in four-morpheme chengyu, based on the research of Liu and Xing (2003). Specified to two decimal places.)

Table 9 shows that more than half of four-morpheme chengyu have some kind of phonetic repetition. Phonetic repetition is thus an important property of a chengyu. The phonetic repetition of internet quasi-chengyu in this thesis have also been examined, and 42.59% of the whole collection has at least one type of phonetic repetition (see table 10). However, it doesn't reach more than half (like regular chengyu). For the types of phonetic repetition that are found in the internet quasi-chengyu collection, the percentage distribution is following the path of that of that of real chengyu: for example, alliteration in real chengyu is 13.97%, while 11.11% in quasi chengyu; rhyme is 29.13% in real chengyu and 27.78% in quasi-chengyu; repetition is 5.69% in real chengyu and 3.7% in quasi-chengyu. The distributions have the same tendency.

Type	quantity	percentage

alliteration	6	11.11%
rhyme	15	27.78%
repetition	2	3.70%
subtotal	23	42.59%
non-identified	31	57.41%
total	54	100.00%

(Table 10. Percentage of phonetic repetition in internet quasi-chengyu. Specified to two decimal places.)

In this section, phonetic repetition of chengyu and internet quasi-chengyu have been examined. On one hand, internet quasi-chengyu do not have all the phonetic repetition types like chengyu does; but on the other hand, the percentage distribution of all types of phonetic repetition that are found in internet quasi-chengyu is similar to the distribution of the same types across real chengyu. In this way, internet quasi-chengyu and real chengyu are connected.

3.5. Summary

This chapter discusses the phonological features of Chinese chengyu. Comparison between researches of four-morpheme chengyu and internet quasi-chengyu are made and this comparison indicates some similarity between real chengyu and internet quasi-chengyu, but to a limited extent. Phonetic patterns of internet quasi-chengyu do not strictly follow the path of real chengyu. The foot pattern that appears significantly the most is –O–X, which is similar to real chengyu. These results might be due to the relatively small quantity of research material of this thesis. Another factor that may affect this difference between quasi-chengyu, real chengyu and four-morpheme phrases, is the frequency of use. Although chengyu with highly ranked patterns are large in number, these chengyu are not necessarily frequently used in daily life. The phonetic patterns of the internet quasi-chengyu are more likely to be influenced by frequently used chengyu. The materials of the sample research of Liu are more up to date, which means the chengyuand other types of four-morpheme phrases are likely to be more common, and the results are closer to what I found in my investigations of internet quasi chengyu. This may suggest that internet quasi-chengyu are not only influenced by common chengyu, but also by other types of common four-morpheme phrases. Quantitive analysis of frequency is unfortunately outside the scope of this paper. Focus on frequency and its influence on internet quasi-chengyu is expected for future studies.

4. Conclusion

Having noticed that self-made four-morpheme chengyu-like phrases on the internet – which are termed as internet quasi-chengyu in this thesis – are popular, this thesis makes a comparison between these quasi-chengyu and the real chengyu. Both syntactic and phonologic features have been researched. Comparing with other previously conducted researches and with the data research of this thesis, some similarities and differences are discovered. Syntactically, the inner structure of internet quasi-chengyu and real chengyu follow a similar trend: The most common form being non-parallel, the smallest being partly parallel, and about

one-fourth to one-third being parallel. Parallelism is also a non-ignorable feature for both real chengyu and internet quasi-chengyu. Phonologically, the tone patterns of real chengyu display similar trends. Generally speaking, the –O–X and –X–O patterns are more common, and patterns with more than two X rank less. This phenomenon is not seen in the results of internet quasi-chengyu. The emergence of tone patterns across the collection for this thesis appears more divided. The ranking of this thesis is closer to Liu's sample research, which doesn't only consist of real chengyu, but also other kinds of four-morpheme phrases. This shows that internet quasi-chengyu may also have been influenced by other four-morpheme phrases. For real chengyu, diverse types of repetition such as alliteration, rhyme and repetition are also important factors. More than half of the chengyu collection contains one or more types of repetitions. This proportion is not as high in internet quasi-chengyu. Although the percentage is non-ignorable, the total amount is only 42.59%, which is less than half. However, the distribution pattern which is displayed by the different types of repetition is more or less the same.

Judging from the difference and similarities between internet quasi-chengyu and real chengyu, although there is accordance between them to some extent, it is hard to say whether internet quasi-chengyu will remain used in the long term. Even though some may be used in the long term, it is still possible that they will be used as fixed phrases or special terms instead of chengyu. So whether internet quasi-chengyu will meet the criterion of chengyu and whether they will be in public use for a long time remains uncertain.

Due to practical reasons, there are still other factors that affect the results and conclusion. First, the amount of real chengyu is much larger than the amount of internet quasi-chengyu. This is reasonable, since the internet has only existed for a few decades, but real chengyu have been added up throughout the whole history of Chinese language. The size of such a small collection brings difficulty for data analysis. For some patterns only one example has been found, or even none, but it doesn't mean that those patterns won't exist. If the amount ofinternet quasi-chengyu keep growing in the future, the results might become clearer. Second, although the other researchers have done research based ondifferent chengyu dictionaries, it is possible that a lot of the chengyu are not used anymore. There is no test of frequency. Since the internet quasi-chengyu are modern, the patterns are likely to be influenced by the most frequently used chengyu. If a data analysiscan be made based on the most frequent chengyu, the results might be more in accordance with what I found when looking at the quasi-chengyu in this thesis. Quantitive analysis of frequency is unfortunately outside the scope of this paper. Focus on frequency and its influence on internet quasi-chengyu is expected for future researches. Will internet quasi-chengyu eventually reach the status of chengyu? This can be another question in future researches.

Bibliography

Che, F. (2014). "Wǎngluò xīn chéngyǔ" gàiniàn jí xiāngguān wènt´t de xīn tànsuǒ. 'A New Exploration of Quasi- idioms of Network and Their Related Issues'. *Shàoxīng dàxué xuébào (zhéxué yǔ shèhuì kēxué bǎn), 34, 5,* 75-78.

Cheng, J. (2011). Hànyǔ píngzèlǜ qiǎn lùn. 'A general discussion of the tonal patterns of Chinese' *Cāngzhōu shīfàn dàxué xuébào, 27, 1,* 17-18.

Dong, S., & Yan, J. (1996). *Hànyǔ Zhīshi Cídiǎn*. Beijing: Jǐngguān jiàoyù chūbǎn shè.

Duanmu, S (2014). Stress theory and stress phenomena. *Contemporary Linguistics, 3,* 288-302.

Gimson, A.C. (2008), Cruttenden, A., ed*., The Pronunciation of English* (7 ed.), Hodder.

Liao, G. (2011), *Chuánméi* zhōng xìxuè nǐqīnshǔ chēngwèi cíyǔ de yǔyòng fēnxī.'Analysis of the usage of kinship terms in the media', *Biānjí zhīyǒu, 7,* 78'79.

Liu, H. (2002), *Wǎngluò Yǔyán 'internet language'*. Beijing:Zhōngguó guǎngbō diànshì chūbǎnshè.

Liu, Z. (2004). Hànyǔ sìzìgé chéngyǔ píngzè dāpèi de duìchènxìng yǔ rènzhī. 'Symmetrical Features of Ping-ze Arrangement in Four character Chinese Idioms and Cognition'. *Shāndōng dàxué xuébào (zhéxué yǔ shèhuì kēxué bǎn), 4,* 44-51.

Liu, Y. (2013). Cóng cíhuìxué jiǎodù jiěxī 2013 nián wǎngluò rè cí de gòuchéng. 'The formation of popular words on the internet in 2013 from a morphological point of view'. *Jiāozuò shīfàn dàxué xuébào, 29, 4,* 25-27.

Liu, Z., & Xing, M. (2000). Hànyǔ sìzìgé chéngyǔ yǔyì jiégòu de duìchènxìng yǔ rènzhī. 'The symmetry and cognition of the semantic structure of Chinese four-morpheme-chengyu'. *Shìjiè hànyǔ jiàoxué, 1,* 77-81.

Liu, Z., & Xing, M. (2003). Sìzìgé chéngyǔ de yīnyùn duìchèn yǔ rènzhī. Symmetry in 'Phonological Patterns in Four-character Idioms and Cognition'. *Yǔyán jiàoxué yǔ yánjiū, 3*, 48-56.

Ma, R. (2014). Wǎngluò chéngyǔ –cóng "rén jiān bù chāi" shuō qǐ. 'Internet chengyu, starting from "rén jiān bù chāi"'. *Xiàndài hànyǔ, 4,* 131-132.

Shao, Q. (2014). Wǎngluò xīn chéngyǔ de yǔyán piānlí jí qí guīfàn. 'Language Deviation and Standardization of New Network Idioms'.*Tàiyuán dàxué xuébào, 15, 1,* 75-78.

Wen, D. (2005). *Hanyu Yuhuixue*. Beijing: Shāngwù yìnshūguǎn.

Wen, D. (2006). *Hanyu Yuhuixue Jiaocheng*. Beijing: Shāngwù yìnshūguǎn.

Wu, J. (1992). Hànyǔ jiézòu de zhōuqī jí chéngcì –jiézòu guīlǜ yánjiū zhīyī. 'The circle and level of Chinese rythm – a research of rythm patterns'. *Zhōngguó yǔwén, 1992, 2,* 129-135.

Yang, K. (2014). Wǎngluò sìzì duǎnyǔ jí "wǎngluò chéngyǔ" diàochá yánjiū. 'A research on internet four-morpheme phrases, also known as "internet chenyu"'. *Chángchūn lǐgōng dàxué xuébào (shèhuì kēxué bǎn), 27, 7,* 110-113.

Yu, G. (2012). Chéngyǔ yùnlǜ mànyì. 'A general discussion of the rhyme of chengyu'.*Chángzhōu gōngxuéyuàn xuébào (shèhuì kēxué bǎn)*, *30, 6,* 64-66.

Zhang, Q. (1990). *Xiandai Hanyu Zhishi Cidian*. Chengdu: Sìchuān rénmín chūbǎnshè.

Appendix 1. Internet quasi-chengyu with a frequency of use higher than 1000 based on Che (2014)

zhènglóng pāi hǔ 正龙拍虎

Someone defrauds and deniesit after being found out; Lacking of social credibility

qiūyǔ hán lèi 秋雨含泪

A metaphor for crocodile tears

zhàoshān xiàn guǐ 兆山羡鬼

Referring to someone cold-blooded

jù dǎ jiàngyóu 聚打酱油

Having no opinion about something

ōuyáng wā kēng 欧阳挖坑

Celebrities giving irresponsible and careless speeches.

fàn pǎo guō tiào 范跑郭跳

Birds of a feather; Cut from the same cloth.

guān zǒu shēng zuò 官走生坐

The officers should leave first (when there is danger) and the students must sit and wait.

shì jiān tiào hǎi 世坚跳海

Making strong political statements.

sānmáo chāo sì 三毛抄四

Drawing conclusions without making any effort to find out the truth; purposely confusing the public

hé chuán yán jìn 核传盐尽

Expresses the strength of rumorand how blinded the public is

tángjùn dú bó 唐骏读博

Defrauds; lacking of social credibility. See *zhènglong pāi hǔ*

dōng pǎo xī diān 东跑西颠

Going everywhere.

wǔ pán xī shòu 捂盘惜售

Hoarding property and rigging its price. (in the real estate industry)

guó jìn mín tuì 国进民退

Reversion of marketization and enhance of macro-control.

fēifǎ xiàn huā 非法献花

Refers to government's interference in human rights.

diào yú zhí fǎ 钓鱼执法

entrapment

jiǔ gēng tuō shì 久耕托市

Attempt something which exceeds one's limitations

xué lì tuán gòu 学历团购

Group purchasing diplomas. Pure corruption.

kuà shěng zhuī bǔ 跨省追捕

Supression of free speech.

chùnǚ piáochāng 处女嫖娼

The police works inefficiently.

dīnglěi yǎng zhū 丁磊养猪

Speculation; action under pressure; the food safety problem is extremely severe.

lù biān gāo qiáng 路边高墙

To build a wall in order to block the sight of a sewage.To cover up mistakes instead of solving them.

gùgōng qiáng hàn 故宫强撼	Making simple writing mistakes; pretending to be a all-knowing; reluctant to admit one's mistakes
ōu zhū wǔ guó 欧猪五国	A negative name for five European countries with the lowest bond credit rating: Portugal, Italy, Ireland, Greece and Spain.
bàidēng chī miàn 拜登吃面	The unrealistic hope of getting the biggest benefit at the lowest cost.
qián lǘ sān chēng 黔驴三撑	[Pejorative]Someone is just capable of playing by themselves
kāi xiōng yàn fèi 开胸验肺	The difficult situation for occupational disease patients to protect their rights.
shuā qī lǜ huà 刷漆绿化	Painting the rock in green instead of planting more trees. Coveringup mistakes instead of solving them.
	Imposture for reason or goal.
	Mixing the false with the genuine.
lín mào yáng yīn 林貌杨音	Doing vile and immoral things in the name of "national benefit"
zhèng qí dú quán 证齐毒全	Using certificates as a shield to make illegal profits.
hán shí měi shǐ 韩食美史	Things that arenot worth mentioning
yì zhèng yì xié 亦正亦邪	Being good and bad at the same time.
qiǎo yán lì kǒu 巧言利口	Good in debates or arguments.
nì zéi bài dǎng 逆贼败党	A defeated party of traitors.
nán mò nǚ lèi 男默女泪	After seeing something, men become silent and women start to cry. Mostly used for some essays that are always about emotion and love.
rén jiān bù chāi 人艰不拆	Life is already so hard that things should better be left unsaid. Mostly used when someone said something which is true but harsh, the other person knows it as well, butis reluctant to face the truth. Sometimes used as sarcasm to describe some social phenomenon.
xì sī kǒng jí 细思恐极	Feeling terrified by something after giving it a second thought.
lèi jué bú ài 累觉不爱	Feeling too tired to ever fall in love again.
bù míng jué lì 不明觉厉	To not understand what someone is talking about, but still being impressed. Sometimes it is also used in a sarcastic way to mock someone who tried to

be a know-it-all.

xǐ dà pǔ bēn 喜大普奔

Referring to good news that should be spread. Can also be used in a sarcastic context.

shuō nào jué yú 说闹觉余

Feeling left out.

shí dòng rán jù 十动然拒

Being really moved by some kind act but still rejecting it afterwards. Mostly used in the context of relationship.

xiào ér bù yǔ 笑而不语

"Smilence". To laugh without speaking.

shè bìng wǒ yào 社病我药

It is the society that is sick, why should I take medicine? Mostly used to comment on news on the internet.

shéi sǐ lù shǒu 谁死鹿手

Showing the concern and frustration with food security, and the feeling of vulnerability to the unscrupulous merchants.

nán kē yí mèng 南科一梦

Unrealistic fantasy.

shí miàn mái fú 十面霾伏

There is haze everywhere. A complain about the air condition.

fán róng chāng shèng 繁荣娼盛

The spring up of solicitation.

qián fǔ hòu jì 前腐后继

Corruption.

gǎn dòng zhōngguó 敢动中国

Other people or countries that dare challenge China.

yālì shān dà 压力山大

The pressure is as heavy as a mountain. It is an adaption of 亚历山大 (the transliteration of the English name Alexander).

huà wēi wéi jī 化危为机

Turn a crisis into opportunity.

zhū tú kǒuhóng 猪涂口红

Coming from the English saying "putting lipstick on a pig" during Barrack Obama's presidential campaign. Making things seems more attractive in order to tempt or deceive others, but not changing the essence.

páng qí qū jìng 旁岐曲径

Borrowed from Korean. Refers to government being arbitrary and ignoring the public opinion.

dì mìng hǎi xīn 地命海心

Being of low social statusbut staying concerned about state affairs.