

Is the Spreading of Internet Neologisms Netizen-Driven or Meme-driven? Diachronic and Synchronic Study of Chinese Internet Neologism *Tuyang Tusen Po*

Zongwei Song

School of Foreign Languages, Sichuan University of Arts and Science, Dazhou, China

Abstract—With the development of internet corpus and sophisticated searching engines, it is possible and convenient to apply the linguistic big data to quantitatively conduct research to Chinese internet neologism *tuyang tusen po*. Based on linguistic big data from the Web Corp Live and Baidu Index, the author has conducted the diachronic and synchronic research to the Chinese internet neologism *tuyang tusen po*. It is found that: (1) *tuyang tusen po* is used as noun, adjective and exclamation by Chinese netizens; (2) during using *tuyang tusen po*, there exist several spreading centres in the macro, middle and micro scope; (3) the spreading of Chinese internet neologism *tuyang tusen po* is netizen-driven, not the so called meme-driven. The primary conclusion is that there are several spreading centers during the spreading of the Chinese internet neologisms, which is caused by netizen-driven. The findings can not be quantitatively revealed without linguistic big data from the internet.

Index Terms—internet neologism; corpus, quantitative research, spreading features

I. INTRODUCTION

The term neologism is defined as: ‘a new word, usage or expression’ in the *Webster’s Third New International Dictionary* (Gove et al. 1993: 1516). With the fast development of internet and the popularization of mobile communication tools, there are a great number of internet neologisms produced by netizens every year. The emerging internet neologisms are used to name new things, express new ideas, construct new identities, and describe new phenomena in the internet world by creative ways. The neologisms somewhat reflect the cultural development, social changes, personal emotions, and netizens’ attitudes. There are a host of Chinese internet neologisms emerging in the past decade that embody the conspicuous novelty characteristic of internet language. Consequently, Chinese internet neologisms have attracted the attention of Chinese government, linguists and lexicographers: the book-length ‘word lists’ of internet neologisms was annually published over the past decades, documenting and reporting the latest coinages; Zhou (2007) viewed the Chinese internet language that is full of neologisms as ‘a totally different language’; Chinese internet neologisms are considered to occupy a unique place in Chinese lexicography (Zhuo Jing-Schmidt and Shu-Kai Hsieh, 2018).

With the sharp increase of netizens, Chinese internet neologisms are mushrooming. The internet users in China are surging in recent years, and according the statistics of *Internet World Stats 2017*, the penetration of population is up to 53.2%. It is surprised that there are 0.75 billion netizens until August, 2017 in China (China Internet Information Centre, 2007). The increasing trend of netizens in China will continue. The top five Chinese internet neologisms appearing in 2017 are: 点赞之交 (pinyin: *dianzan zhijiao*, ‘being friends by ‘give a like’ in cyberspace’), ‘劫’后余生 (pinyin: *jiehou yusheng*, ‘lack of money after the double-11 shopping frenzy alike’), 无‘期’徒刑 (pinyin: *wuqi tuxing*, ‘without wife, being single for long time’), 友荒 (pinyin: *youhuang*, ‘lack of friends’), 撩妹 (pinyin: *liaomei*, ‘make a pass at’). The numerous Chinese netizens who are energetic, learned and active in mind are creating and using a large number of Chinese neologisms in the internet. The Chinese internet neologisms are a crystallization of the wisdom of Chinese netizens, which are humorous, vivid, easily recognizable and high frequent use in the China internet community. With the popularization of mobile phones in China, another feature of Chinese internet neologisms is the instant and wide spreading in cyberspace, which are initially used by netizens of high schools or universities, and then are quickly spreading among the other netizens.

However, not all the Chinese internet neologisms are always welcome and popular in the China cyberspace. For most of the Chinese internet neologisms, the time of popularization is limited, from several months to one year, while the life span of some internet neologisms with cultural characterization or *Zeitgeist* is long, which can even be conventionalized or lexicalized. The conventionalized Chinese internet neologisms are part of linguistic system of Mandarin. The diachronic researches on the conventionalized neologisms can reveal the evolution processes, usage patterns, and distribution features. Hence, the long-life-span neologisms are worth researching. Some Chinese internet neologisms

are quickly spreading in the Chinese internet world, especially in youth netizens, and after several years some of them might be accepted by much more formal media, with the increase of usage frequency. With the development of web corpora and searching engines, the detailed spreading processes of internet neologisms can be observed and trailed, which is helpful for understanding the laws of spreading and the usage features of internet neologisms. Consequently, it is possible and necessary to quantitatively study the spreading features, and the grammatical and semantic features of Chinese internet neologisms based on internet linguistic big data.

The internet neologism ‘图样图森破’ (pinyin: *tuyang tusen po*, ‘too young’, ‘too simple’, ‘too naive’) is a typical Chinese internet neologism and takes on the typical diachronic and synchronic spreading and usage features. It has lasted seven years, initially appearing in the internet in 2011, which still shows the tendency of continuing to be popular in the Chinese cyberspace. Based on the data from Web Corp Live and Baidu Index, the detailed spreading processes and usage features of Chinese internet neologisms can be observed and revealed through *tuyang tusen po*.

II. LITERATURE REVIEW

With the emergence of vast majority of Chinese internet neologisms, Scholars have extensively conducted researches on them. The researches can be classified into three categories. The first category is the linguistic studies on the Chinese internet neologisms: some scholars concentrated on the unprecedented lexical production revolution driven by internet (Wang & Zhong, 2008); some viewed them, from the variationist sociolinguistic perspective, as elements of a unique sociolect of the internet community (Kuang & Jin, 2008). The second category is CMC (computer mediated communication) studies on Chinese internet neologisms, focusing on the mechanism of production. While theoretical validity of memetics has been questioned in the Chinese scholarly community, memetics-based research on Chinese internet neologisms was fashionable (Wu, 2009; Cao, 2012; He, 2014, 2016). The third category is computational studies on the Chinese internet neologisms: researching neologisms based on internet corpus (Zhang, 2017).

There are a lot of progress and advantages in the corpus-based research on internet neologisms. Corpus-based computational lexicology and lexicography have made progress in extracting new words, senses, and automatically parsing morphological patterns in recent years (Zhuo Jing-Schmidt & Shu-Kai Hsieh, 2018), which supplies the techniques to the internet neologism research. The neologisms studied on empirical basis shows the life cycle: birth, settling down, obsolescence, death, and re-birth (Schmid, 2008). Quantitative studies on the neologisms have made significant progress in the life-stage statistics (create, evolve, die out), assuming that word frequency plays the leading role on explaining life stages, the prediction force of whether a neologism might survive after being created (Altmann et al., 2011, 2013). Based on the context of Mandarin Chinese in Taiwan, Liu et al. (2013) think that the frequency is a determinant in lexical diffusion and changes of neologisms. Actually, there are a number of factors affecting the diffusion, the changes, the distribution, and the life-span of internet neologisms: the age and gender of the netizen, the media, the event, the values (carried by neologisms), economic condition of netizens, and so on. The volume of corpus or data will also affect the results of the research on the internet neologisms. Generally speaking, the larger the corpus or data are, the more accurate the research results of neologisms are. There are substantial advantages of internet corpus or big data over the research on internet neologisms: (1) the larger volume: billions, trillions or even zillions words; (2) the dynamics and openness of words or language resources; (3) much more macro and powerful retrieval functions; (4) the multifarious origin of language resources, different ages, genders, education backgrounds, and so on; (5) netizen-driven perspective; (6) diachronic and synchronic description. These advantages of internet corpus or big data made the researches on the internet neologisms much more accurate, macro and much deeper, and make the research results be in line with the actual usage of the internet neologisms. In a word, the research on internet neologisms based on the internet corpus or big data can overcome the shortcomings based on the traditional corpus or small data, and can much more accurately depict the spreading state of internet neologisms in detail.

In order to thoroughly and quantitatively study the life span, and the diachronic and synchronic spreading features of a typical Chinese internet neologism, the paper conducts the research on the Chinese internet neologism *tuyang tusen po*, based on the web corpus Web Corp Live and Baidu Index.

III. METHOD

A. Research Questions

The paper mainly focuses on answering the following four questions: (1) What are the semantic and grammatical features of the Chinese internet neologism *tuyang tusen po*? (2) What are the diachronic and synchronic spreading features of *tuyang tusen po*? (3) What are the causes of the spreading features? (4) Are the neologism *tuyang tusen po* affected by native English?

B. Research Tools

Kilgariff and Grefenstette (2003) said ‘The web, teeming as it is with language data, of all manner of varieties and languages, in vast quantity and freely available, is a fabulous linguists’ playground.’ There are generally two approaches applying the web data to the corpus research: web as corpus and web based corpus. The former directly views web as a corpus and conducts linguistics research on it. The latter just crawls texts from the web, uses them to

construct corpus and then conducts research on the constructed corpus. There are several advantages for the web as corpus approach: data is much larger and more multifarious than web based approach, but there are also some unavoidable shortcomings such as the ‘raw’ or ‘unmarked’ linguistic materials, and much more ‘noise’ materials. The web based corpus approach can overcome the shortcomings of the web as corpus approach. Considering the advantages of web as corpus and web based corpus, this paper adopts the combination research tools of the web based corpus tool called Web Corp Live¹ and web as corpus tool called Baidu Index². Web Corp Live is a large-scale search engine with more search options, part-of-speech tags and quantitative analyses³. Web Corp Live can analyze authentic texts of billions of words (collected from the web) to identify not only instantly what is typical in language and what is rare, unusual or emerging usage, but also the grammatical and semantic features of words or phrases. Hence, the paper applies the Web Corp Live to explore the semantic and grammatical features of internet Chinese neologism *tuyang tusen po*.

Baidu Index is developed by Baidu, Inc., which is one of the largest internet search engine companies in the world. Baidu is the second largest search engine in the world, and holds 76.05% market share in China’s search engine market for 0.751 billions netizens⁴, in 2017. Baidu search engine deals with trillions of data produced by netizens everyday, which in fact is truly big data. Baidu Index allows users to look up the search volume and the trend for certain hot keywords and phrases. It is claimed by the Baidu, Inc. that Baidu Index is based on the Baidu big data to search and analyze the hot keywords, exploring the public sentiments, market needs or features of users⁵. The data of Baidu Index are updated very day, supplying the PC search data from June, from 2006 to the present, and supplying the mobile search data from January, from 2011 to present. Hence, Baidu Index is properly suitable for the research on the diachronic and synchronic spreading features of the internet Chinese neologism *tuyang tusen po*.

C. Research Procedures

1. Using Web Corp Live to analyze the grammatical and semantic features.

After Retrieving all the concordances containing the node word *tuyang tusen po* from the Web Corp Live, the author reads the concordances one by one and then generalizes the POS and meanings of *tuyang tusen po*.

2. Using Baidu Index to analyze the diachronic and synchronic changes, the spreading features.

(1) Searching the diachronic spreading process of *tuyang tusen po* on PC end

In order to observe the diachronic spreading process of *tuyang tusen po* on PC end, the author searches *tuyang tusen po* on Baidu Index PC end, see Fig. 1 for the result.

(2) Searching the diachronic spreading process of *tuyang tusen po* on mobile end

In China, there is an internet trend moving from the PC end to mobile end, which means the Chinese netizens use mobile net more frequently than PC net in passing several years. According to the CNNIC report, in December, 2016, there are 0.695 billion mobile netizens out of 0.731 billion netizens in China⁶. It can be said that most of the Chinese netizens are currently using the internet by mobile end. In order to more comprehensively observe the diachronic spreading process of *tuyang tusen po* on the mobile end, the author retrieves data of the diachronic process in Baidu Index mobile end, see Fig. 2 for retrieving result.

(3) Searching the synchronic spreading state of *tuyang tusen po* on mobile end

In order to observe the synchronic spreading state of *tuyang tusen po* on mobile end, the author retrieves the synchronic mobile end data (time span of a month: from November 26, 2017 to December 26, 2017) on the Baidu Index, see Fig. 3 for the retrieving result.

(4) Macro-spreading of *tuyang tusen po* in China internet world

In order to research the macro-spreading features, the author retrieves the macro-spreading state of *tuyang tusen po* in different provinces and larger districts in China Mainland, see Fig. 4 and Fig. 5 for the results. In order to observe the gender and age difference of the users of *tuyang tusen po*, the author uses the ‘portrait’ function of Baidu Index to retrieve it, see Fig. 6 for the result.

(5) Middle-spreading of *tuyang tusen po* in Chinese internet world

In order to observe the middle-spreading features of *tuyang tusen po*, the author retrieves the netizen distribution according to the cities in China, checking which cities are the spreading centers using *tuyang tusen po*, see Fig. 7 for the result.

(6) Micro-spreading of *tuyang tusen po* in Chinese internet world

In order to get the micro-spreading features of *tuyang tusen po*, the author retrieves the netizen distribution within the province called Guangdong, see Fig. 8 for the result.

(7) Verifying the statement: netizen-driven-internet-neologism

In order to verify the statement that The spreading of internet neologism is netizen-driven, but not so-called meme-driven, the author conducts the retrievals based on Baidu Index ‘portrait’ function: retrieve three different Chinese internet neologisms *打call* (pinyin: da call, ‘cheer for’), *hold住* (pinyin: hold zhu, ‘control’), *tuyang tusen po*. The three Chinese internet neologisms appeared in different periods in the internet: *打call* (in 2017), *hold住* (in 2014), *tuyang tusen po* (in 2011), see Fig. 9-10 for the result.

IV. RESULTS AND DISCUSSION



Picture4. 1 Calligraphy work with *tuyang tusen po*

C. The Synchronic Spreading State of *Tuyang Tusen Po* on Mobile, Within Recent 30 Days, from November 26, 2017 to December 26, 2017

Figure 3 show that, from the synchronic perspective, the fluctuation range is not sharply like the diachronic process with a peak value 16000 (see Fig. 2). The average value in Fig. 3 is just 373 on the Mobile, which is smaller than that (value 655, in Fig. 2) of diachronic process. The value of the synchronic spreading within 30 days also proves that the frequency of *tuyang tusen po* is continuing decreasing in recent years, but there are still relatively stable netizens who daily use this neologism to express their ideas and attitudes in the China internet.

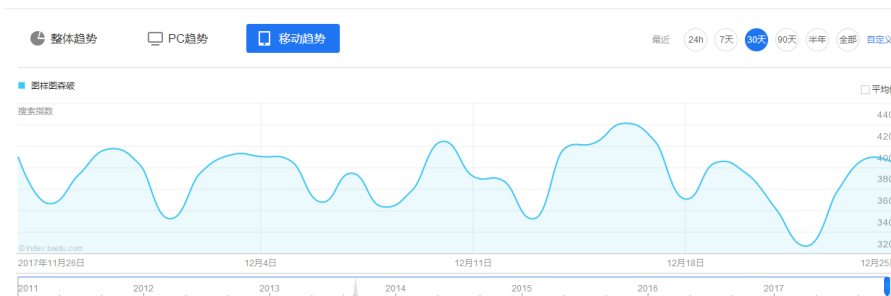


Figure 4.3 The synchronic spreading state of *tuyang tusen po* in recent 30 days.

D. The Macro-scope Spreading Features of *Tuyang Tusen Po* in China

The internet, especially, the mobile net, connects the 0.751 billion Chinese netizens together, who distribute over the land of the 9, 600, 000 square kilometre. Though there is no ‘district’ boundary in the internet world in China, the neologism of *tuyang tusen po* is not equally used in the different provinces of China. That’s to say, the use of this neologism shows ‘several spreading centers’ in China internet world where the netizens in the areas such as Guangdong province, Jiangsu province are most frequently using *tuyang tusen po*, but netizens in some provinces such as Tibet, Sinkiang, hardly use this neologism. More details of macro-scope spreading centre of *tuyang tusen po* can be seen from the following Fig. 4.

To be more macro, the author can also find the larger using spreading centers of this neologism. China is geographically divided into seven larger districts: Huadong, Huannan, Huabei, Huazhong, Xinan, Xibei, and Dongbei. Based on the data of Baidu Index, Huadong is the largest using spreading center of *tuyang tusen po*, and Huanan is the second one, see Fig. 5. To be more macro, netizens in Xibei, Xinan, and Dongbei hardly use the neologism of *tuyang tusen po* so that there are no spreading centers of this neologism in the above three larger districts.



Figure 4.4 The macro spreading state of *tuyang tusen po* in different provinces of China

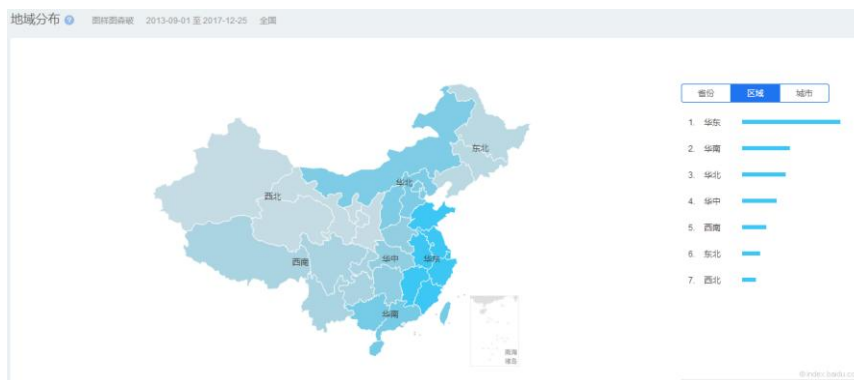


Figure4. 5 The spreading centres of *tuyang tusen po* in the seven districts

E. The Gender and Age Features of Using *Tuyang Tusen Po*

From September 2013, to December 2017, gender and age difference in the using of *tuyang tusen po* is significant. From Fig. 6, it can be seen that male Chinese netizens account for 79%, and female Chinese netizens account for 21%. In terms of age, netizens aged 20 to 29 account for the highest proportion (43%) in using the neologism of *tuyang tusen po*. The second highest proportion is 28% with the age from 30 to 39. The author can make the statement that the Chinese netizens aged 20-39 are especially fond of using Chinese internet neologisms like *tuyang tusen po*, in the internet communication. The Chinese netizens under 19 or over 40 somewhat seldom use Chinese internet neologisms. The retrieving result that the netizens under 19 seldom use the neologism of *tuyang tusen po* a bit overpasses the author’s expectation: before retrieving this neologism on Baidu Index, the author hypothesizes the netizens under 19 is the highest proportion in using it because the netizens under 19 are usually viewed as being the most creative in mind and prone to welcoming and using new words or expressions in the internet communication. However, the result proves the author’s hypothesis is wrong. The big data of the neologism *tuyang tusen po* from Baidu Index reveal the truth: netizens aged 20-29 are the highest proportion of netizens to use *tuyang tusen po* in Chinese internet communication, see more details in Fig. 6.

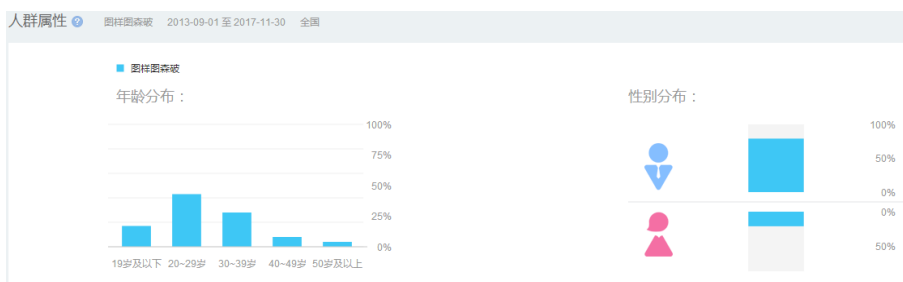


Figure4. 6 Gender and age distributional difference of netizens in using *tuyang tusen po*

F. Middle-and Micro-scope Spreading Features of *Tuyang Tusen Po*

In the middle-scope, the using frequency of *tuyang tusen po* still shows the feature that there are several spreading centers in using this neologism such as Beijing, Shanghai, Guangzhou, and so on, see more details in Fig. 7. Similarly, in the micro-perspective, limiting the area to the province called Guangdong, it can be found that there are still two spreading centers in using this neologism: Guangzhou, and Shenzhen, see more details in Fig. 8. Hence, the author may get a tentative conclusion: whether the area is larger or smaller, there is a spreading center with the highest frequency using the Chinese internet neologisms *tuyang tusen po* in the China internet world.

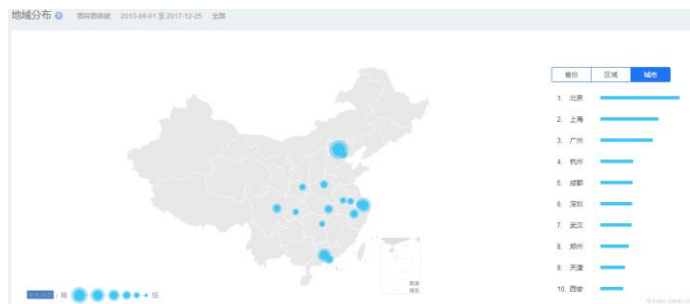


Figure 4.7 The spreading centres of *tuyang tusen po* in key cities of China

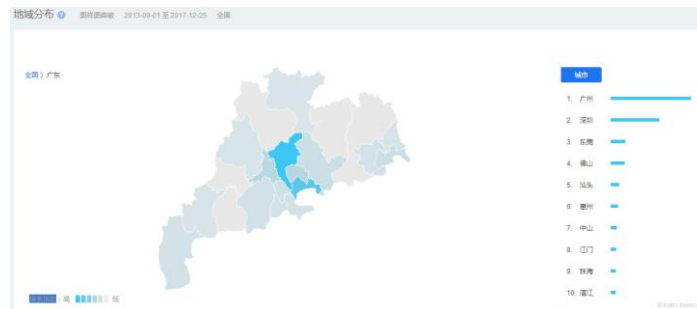


Figure 4.8 The spreading centres of *tuyang tusen po* in Guangdong province

G. Reason for the Spreading Centers: Netizen-driven

The reason for forming the spreading centers of *tuyang tusen po* can be expounded by netizen-driven. In order to further validate the attentive statement that the Chinese netizens aged 20-39 are especially fond of using internet neologisms, the author retrieves the other two Chinese internet neologisms *打 call* (popular in 2017), *hold 住* (popular in 2014). It can be seen from the following Fig. 9: there are almost the same spreading centers for the three different internet neologisms, though the three neologisms are emerging and popular in different time. What is the reason resulting in the spreading centers in using the internet neologisms? Memetics can not explain this phenomenon. According to the memetics, Chinese netizens can copy or imitate the neologisms with ‘meme’, and in the China internet world, there is no ‘district’ boundary so that the netizens can equally (or have the same opportunities to) use Chinese internet neologisms to exchange ideas. Hence, spreading centres should not exist. However, the big data on the Chinese internet neologisms (*打 call*, *hold 住*, and *tuyan tusen po*) from Baidu Index reveal that the spreading centres do exist. Can the density of population in each province explain why there are spreading centres in using the internet neologisms? There is no direct relationship between the density of population and using frequency of neologism *tuyang tusen po*. According to the statistical results⁸, the top seven provinces or cities with the highest population density in China are: Hong Kong, Macao, Shanghai, Beijing, Tianjing, Jiangsu, and Shangdong. However, the spreading centres of *tuyang tusen po* don’t lies in the above seven provinces or cities. If there is direct relationship between population density and using frequency of internet neologisms, the spreading centers of using *tuyang tusen po* should be Hong Kong, Macao, and Shanghai. Nevertheless, Fig. 4 reveals the three true spreading centres are Guangdong, Jiangsu, and Zhejiang. Similarly, Fig. 9 also reveals that the population density ranks and the using frequency ranks of the three neologisms (*打 call*, *hold 住*, *tuyang tusen po*) do not match! Actually, it is revealed by quantitative linguistics and corpus linguistics that language is a complex-dynamic adoptive system and human-driven system (Liu & Lin, 2018; Kretzschmar, 2015; Ellis & Larsen-Freeman, 2009). The change and development of language is caused by language users. The language is human-driven, and words consist of the language. Hence, the author argues the Chinese internet neologisms are also netizen-driven, which is verified and explained in the following.

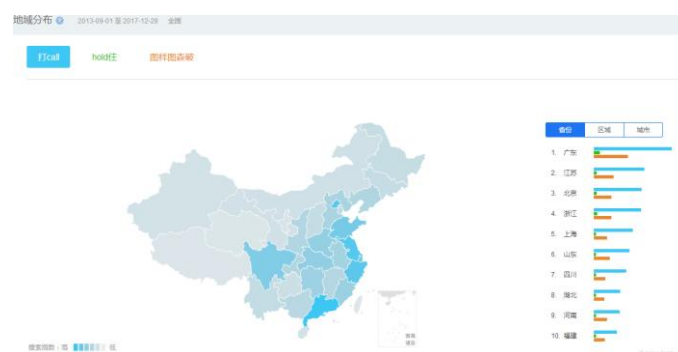


Figure4. 9 The ranks of three neologisms *打 call*, *hold 住*, *tuyang tusen po* in provinces

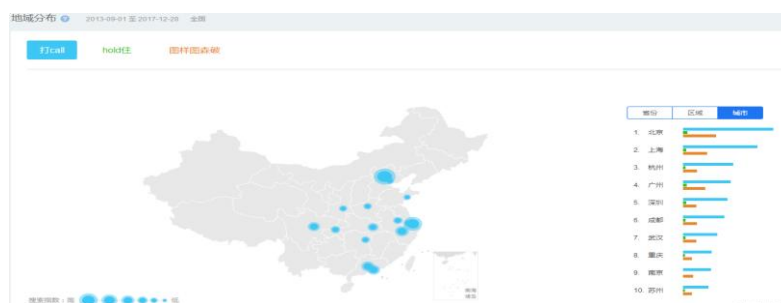


Figure4. 10 The ranks of three neologisms *打 call*, *hold 住*, *tuyang tusen po* in cities



Figure 4.11 The ranks of the Chinese internet neologisms *hold 住* in cities

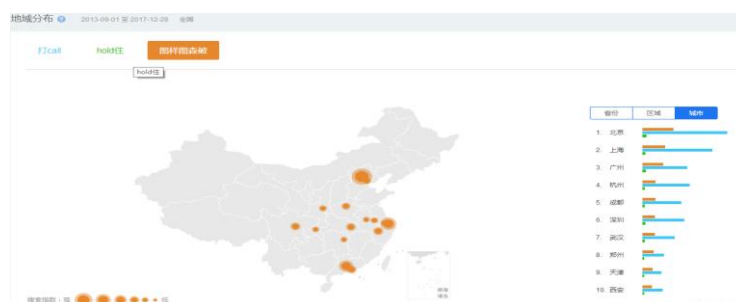


Figure 4.12 The ranks of the Chinese internet neologisms *tuyang tusen po* in cities

The spreading centres of internet neologisms revealed by the big data from Baidu Index can be explained by the netizen-driven. General speaking, Chinese netizens are active, creative, and learned. Particularly, the Chinese netizens aged from 20 to 39 are rich, creative and prone to accepting new matters so that they are the main group to use Chinese internet neologisms (see Fig. 6). According to China Statistical Report on Internet Development (China Internet Information Centre, 2017), the internet penetration rate shows that the top three cities or provinces are Beijing (77.8%), Shanghai (74.1%) and Guangdong (74.0%). Surprisingly, the top three biggest spreading centres of using the three Chinese internet neologisms, *打 call*, *hold 住*, *tuyang tusen po*, are exactly located in Beijing, Shanghai and Guangdong, which shows that the spreading centres are closely related to the internet penetration rate, see Fig. 10-12 for more details. The author may make the tentative statement that the use of Chinese internet neologisms is affected by internet penetration so that the spreading of Chinese internet neologisms is netizen-driven.

H. Is *Tuyang Tusen Po* Affected by Native English?

The Chinese internet neologism *tuyang tusen po* is not affected by the native English. According to the explanation of *Baidu Baike*⁹: *tuyang tusen po* is the result of the homo-phonic translation of the China English 'too young, too simple', which means the neologism *tuyang tusen po* derives from 'too young, too simple'. In order to check whether the English native speakers use the expression 'too young, too simple' or not, the author retrieved the English expression 'too young, too simple' in the Now Corpus¹⁰, whose volume is 5.5 billion words. The retrieval result is 'no matching record'! It can be said that the expression 'too young, too simple' is not used by the native English speakers in the internet world, daily life or even academic texts. So it can be said that the expression 'too young, too simple' is Chinglish or China English, which reflects the Chinese thoughts and conceptions. The Chinese concepts or ideas of *很年青* (pinyin: hennianqin, 'very young'), *很天真* (pinyin: hentianzhen, 'naive') are coded into the China English 'too young, too simple', and then be homophonically interpreted into *tuyang tusen po* by Chinese netizens, which are quickly spreading in the China internet world through netizen-driven. In essence, the neologism *tuyang tusen po* is Chinese-rooted neologism and it widely spreads through netizen-driven.

V. CONCLUSION

Quantitatively based on the linguistic big data from Web Corpus Live and Baidu Index, the most important findings of the research are: (1) *tuyang tusen po* is used as noun, adjective and exclamation by Chinese netizens; (2) there exist several spreading centres in the macro, middle or micro scopes when this neologism is used in the Chinese internet world; (3) the spreading of the Chinese internet neologism *TuYang TuSen* is netizen-driven, not the so called meme-driven. The findings can not be quantitatively revealed without linguistic big data from internet.

Quantitatively revealed by the linguistic big data, the spreading of the Chinese internet neologism *tuyang tusen po* is driven by Chinese netizens. All in all, Chinese netizens cause the diachronic and synchronic spreading and fluctuation of *tuyang tusen po*. The origin of neologism *tuyang tusen po* is from China English 'too young, too simple', which is coded by the Chinese concepts of 'very young', 'naive'. Chinese netizens are the major forces to coin, use, and make

Chinese internet neologisms spread in the Chinese internet world.

ACKNOWLEDGMENTS

Funding: This work was supported by A Project Supported by Center for Rural Childhood Education Research [Grant NO.: NYJ20150607] and by A Project Supported by Sichuan University of Arts and Science [Grant NO.: 2017JY18]

Note

¹ <http://www.webcorp.org.uk/live/index.jsp>

² <http://index.baidu.com/>

³ <http://www.webcorp.org.uk/live/>

⁴ <https://en.wikipedia.org/wiki/Baidu>

⁵ <http://index.baidu.com/Helper/?tpl=help&word=#wmean>

⁶ <http://tech.sina.com.cn/i/2017-01-23/doc-ixzutkf2340692.shtml>

⁷ All the example sentences are from the concordances collected by Web Corp Live

⁸ <http://baijiahao.baidu.com/s?id=1580694960384056346&wfr=spider&for=pc>

⁹ <https://baike.baidu.com/item/%E5%9B%BE%E6%A0%B7%E5%9B%BE%E6%A3%AE%E7%A0%B4/9830824?fr=aladdin>

¹⁰ <https://corpus.byu.edu/now/>

REFERENCES

- [1] Altmann, E. G., Whichard, Z. L. & Motter, A. E. (2013). Identifying Trends in Word Frequency Dynamics. *Journal of Statistical Physics*, 151(1-2), 277-288.
- [2] Altmann, E. G., Pierrehumbert, J. B. & Motter, A. E. (2010). Niche as a Determinant of Word Fate in Online Groups. *Plos One*, 6(5), e19009.
- [3] China Internet Information Centre. (2017). China Statistical Report on Internet Development. <http://www.cnnic.net.cn/hlwzzyj/hlwzbg/hlwjbg/201701/P020170123364672657408.pdf>. Accessed 25 December 2017.
- [4] Cao, J. (2012). Introduction to Internet Linguistic Communication. Beijing: Tsinghua University Press.
- [5] Ellis, N.C. and Larsen-Freeman, D. (2009). *Language as a Complex Adaptive System*. New Jersey: Wiley-Blackwell.
- [6] Gove, Philip B. & the Merriam-Webster editorial staff. (1993). *Webster's Third International Dictionary of the English Language Unabridged*. Cologne: K ö n e m a n n.
- [7] He, Z. R. (2014). Memetic Understanding of Language in Fashion. *Shandong Foreign Language Teaching*, (2), 8-13.
- [8] He, Z. R. (2016). Linguistic Meme and its Variant Application. *Journal of Xinjiang Normal University (Philosophy and Social Sciences)*, (2), 132-139.
- [9] Kilgariff, A. & Grefenstette, G. (2003). Introduction to the Special Issue on the Web as Corpus. *Computational Linguistics*, (3), 333-347.
- [10] Kretzschmar, W. (2015). *Language and Complex Systems*. Cambridge: Cambridge University Press.
- [11] Liu, H.T & Lin, Y. N. (2018). Methodology and Trends of Linguistic Research in the Era of Big Data. *Journal of Xinjiang Normal University (Philosophy and Social Sciences)*, (1), 72-83.
- [12] Liu, T. J., Shu, K. H. & Laurent, P. (2013). Observing Features of PTT Neologisms: a Corpus-Driven Study with N-gram Model. In: *Proceedings of The 25th Conference on Computational Linguistics and Speech Processing (ROCLING)*. <https://www.aclweb.org/anthology/O13-1025>. Accessed 28 December 2016.
- [13] Schmid, H. J. (2008). New Words in the Mind: Concept-formation and Entrenchment of Neologisms. *Anglia-Zeitschrift für englische Philologie*, 126 (1), 1-36.
- [14] Wang, D. & Zhong, M. (2008). Internet language: A revolution in the history of language. *Journal of University of Electronic Science and Technology of China (Social Sciences Edition)*, (6), 33-36.
- [15] Wu, Y. (2009). A memetic Account of Internet Linguistic Variation. *Journal of Guangdong University of Foreign Languages and Trade*, (3), 75-78.
- [16] Zhuo, J. S. & Shu, K. (2018). Chinese Neologism. In Huang, Chu-Ren, Z. Jing-Schmidt and B. Meisterernst (eds.), *The Routledge Handbook of Chinese Applied Linguistics(to appear)*. London/New York: Routledge.
- [17] Zhang, J. H. (2016). Internet Corpora-based Diachronic Research on the Diffusion of Shanzhai through the Internet. *Journal of Quantitative Linguistics*, 24(2-3), 1-16.
- [18] Zhou, R. (2007). Chinese Netspeak: Now, that's a Different Language. http://www.chinadaily.com.cn/cndy/2007-02/08/content_804036.htm. Accessed 25 December 2017.

Zongwei Song was born in Chongqing, China in 1978. He received his M.A. degree in English Linguistics and Literature from Sichuan International Studies University, China in 2012.

He is currently a lecturer in the School of Foreign Languages, Sichuan University of Arts and Science, Dazhou, China. His research interests include corpus linguistics, cognitive linguistics and second language acquisition.