Culturomics: The Diachronic Spreading of Confucianism and Taoism in the UK and the USA, 1800s–2000s

Siyu Lei[1] & Huayong Li[1]

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

Correspondence: School of Foreign Languages, Sichuan University of Arts and Science, Dazhou, China. E-mail: 723166@qq.com

Abstract

After being introduced to the USA and the UK, Confucianism and Taoism are more and more popular and have more influence, which are reflected in the rising trend of the keywords related to Confucianism and Taoism in the Google English books from 1800s to 2000s. Based on corpora GBAE and GBBE, the author studies the diachronic spreading of Confucianism and Taoism in the UK and the USA. It is found that: (1) Confucianism and Taoism are more welcomed in the UK than in the USA; (2) compared with Taoism, Confucianism is more popular in the UK and the USA; (3) Confucius is more popular both in the UK and in the USA than Lao Tzu; and (4) the general spreading trend of Confucianism and Taoism is rising in 1810s–2000s in the UK and the USA. The author quantitatively answers the frequently asked questions related to the spreading of Confucianism and Taoism.

1. Introduction

Culturomics is the portmanteau combination of culture and genomics, which is defined as the application of high-throughput data collection and analysis to study human culture (Michel et al., 2011). Culturomics extends the boundaries of scientific inquiry to a range of new phenomena from the social sciences to the humanities, and further advance digital humanities. After Michel et al. (2011), there are studies based on Google books (e.g., Twenge et al., 2012a; Twenge et al., 2012b; Kesebir & Kesebir, 2012; Greenfield, 2013). Juola (2013) applies 2-gram data based on Google books to measure cultural complexity. Caruana-Galizia (2016) uses 1-gram data from Google Books from 1870–1945 to test George Orwell's hypothesis that everyday language deteriorates under dictatorships. The aforementioned studies are all used the native language data to research 'native' culture. Google Books are great important cultural products, which can be quantitatively used to study the culture changes. Lamoreasu and Morling (2012) point out that there are three reasons to study cultural phenomena through cultural products: (1) cultural products 'capture' outside-the-head culture; (2) cultural products are not subject to the self-report-measure biases; and (3) cultural products can further shape the culture norms. Google Books are crystallization of wisdom of elites and ordinary people, and the diachronic data of Google books N-gram can reflect the diachronic change of the native culture or the foreign culture spreading in the UK and the USA. So, Google Books N-gram is powerful sources to study native culture changes. The data of Google books published in the UK and the USA may reflect the diachronic spreading status of certain Chinese culture in the two countries. So, the present research tries to research the diachronic spreading of Confucianism and Taoism in the UK and the USA based on data of Google books N-gram.

Confucianism and Taoism were introduced to the UK and the USA in the 17th century. Works of Confucius were introduced to European by translation through the agency of Jesuit missionary stationed in China (Fung, 2008). Matteo Ricci was among the very earliest to introduce the thought of Confucius, and father Prospero Intorcetta wrote about the life and works of Confucius in Latin in 1687. Confucius has been introduced to the United States for more than 300 years and had great influence on U.S media, U.S president, U.S judiciary, and so on (Note 1). Confucius with Moses and Solon are engraved in the east gate of the Federal Supreme Court of the United States, see Figure 1, which shows the influence of the three figures to the USA. Zhang (2011) writes a monograph to study the images of Confucius in the American newspapers from 1849 to now, pointing out that there are different images in American newspapers in different period. Lao Tzu and his thought Taoism have great

influence on the United Kingdom (Note 2). There is Taoist & Tai Chi Society of Great Britain (Note 3), which means that British are interested in Lao Tzu, and Taoism nowadays. *Tao Te Ching* written by Lao Tzu is among the most translated works in world literature (Note 4).
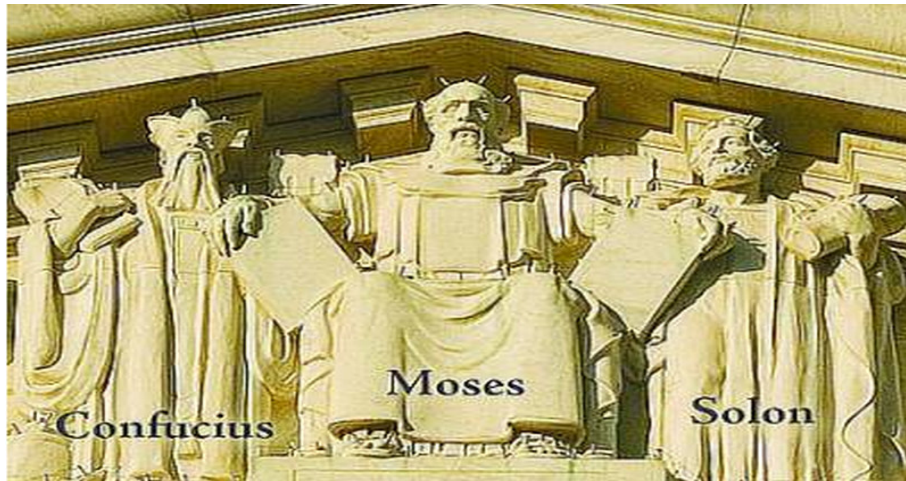

Figure 1. Confucius, Moses, and Salon in the east gate of the Federal Supreme Court of the United States

It is not uncommon to hear assertions that Confucius is more popular in the UK than in the USA, that Lao Tzu is more welcomed in the USA than Confucius, and that Confucianism has more influence on the USA than on the UK. These assertions are usually based on subjective judgments or intuition. Though most of scholars and researchers have realized the importance and influence of Confucius and Lao Tzu, and their thought to the UK and the USA, there are some questions related to the Confucius and Lao Tzu without answers: (1) who is more popular in the UK and the USA; (2) whose thought has much more powerful influence on the UK and the USA; (3) Does America much more appreciate Confucius or Lao Tzu? Based on the personal experience of the author, these questions are also frequently asked by Chinese netizens or scholars (Notes 5 & 6) and given different answers by netizens and scholars based on their intuition or personal experience.

There is a turn toward digital humanities and computational humanities. Lazer et al. (2009) argue that 'computational social science is emerging that leverages the capacity to collect and analyse data with an unprecedented breadth and depth and scale'. Berry (2011) elaborates the computational turn of digital humanities and believes that the questions raised in the humanities today can be researched computationally. There are substantial humanities research based on Google N-gram or big data (Michel et al., 2011; Juola, 2013; Caruanagalizia, 2016; Aiden & Michel, 2013). Consequently, aforementioned questions might be quantitatively and objectively addressed in the waves of digital humanities and computational humanities.

Based on culturomics, the present research studies the diachronic spreading of Confucianism and Taoism in the UK and the USA based on the data from Google Books N-gram.

2. Methodology

2.1 Object of Research

The object of this research is to study the diachronic spreading of Confucianism and Taoism in the UK and the USA, which is carried by the cultural keywords in the books published by the UK and the USA in 1800s–2000s. By observing the diachronic frequency change of cultural keywords (Confucianism, Confucius, Taoism and Lao Tzu) in the books published by America or Britain, it can, to some degree, diachronically indicate the spreading status of Confucianism and Taoism in the two countries. These books are digitalized by Google company, which are the data sources of the two corpora called 'Google Books: American English' (hereafter named GBAE) and 'Google Books British English' (hereafter named GBBE). Based on data of Google Books N-gram, GBAE and GBBE are created and released by Mark Davies (Note 7). Consequently, by retrieving the keywords related to Confucianism, Taoism, Confucius, and Lao Tzu in the two corpora, the diachronic spreading status of Confucianism and Taoism can be observed.

2.2 Research Questions

Based on GBAE and GBBE, this research studies the following three questions related to the diachronic spreading of Confucianism and Taoism in the UK and the USA: (1) whose thought has much more popular; (2) who is more welcomed; and (3) what is the general diachronic spreading status of Confucianism and Taoism in the UK and the USA.

2.3 Research Tool

This research is based on the two corpora GBAE and GBBE. Quantitative analysis has massive potential as a research tool (Gooding, 2013). The volume of GBAE is 150 billion words, and GBBE is 30 billion words (Note 8). GBAE and GBBE are created by Mark Davies in 2011 based on the data public available from Google books N-gram. GBAE and GBBE have the same 'data' as the regular Google Books, but it allows for many of the same searches as COHA (Note 9). The most powerful advantage of Google N-gram or similar corpora is large-scale data, which is necessary to investigate some social phenomena and cultural phenomena. For example, though the volume of corpus COHA is 560 million words, which are large enough to conduct most of diachronic linguistic-related research, the author can not retrieve the concordance contained 'Confucianism', 'Taoism', or 'Lao Tzu in COHA. It is impossible to conduct the diachronic research related to the spreading of Confucianism and Taoism in this corpus because the volume of corpus COHA is not large enough. Fortunately, the volume of GBAE and GBBE is large enough and can provide substantial data related to the spreading of Confucianism in America and Britain. Computational analysis enables us to observe cultural trends and subject them to quantitative investigation (Michel et al., 2011). So, it is suitable to use GBAE and GBBE to investigate the three previous research questions.

2.4 Procedures

(1) Retrieve the diachronic frequency change of cultural keywords *Confucianism* in corpus GBAE and GBBE respectively. After entering the website (Note 10). of corpus GBAE, on the left-hand column of the interface, the author chooses 'CHART' in the 'DISPLAY' section and put 'Confucianism' into 'WORD(S) BOX'. Then kick the bottom 'SEARCH' and on the right-hand column the retrieving result of *Confucianism* can be showed. There are three ways to display the retrieving results: TOKENS, PER MIL and CHART. In order to accurately compare the results of all the cultural keywords in different periods and different corpora (GBAE and GBBE), the author chooses the normalized frequency (PER MIL) in each decade from 1810s to 2000s with time scale 10 years. PER MIL means the occurrences of a token per million words. By the previous steps, the author gets the data of the diachronic normalized frequency of *Confucianism* in GBAE. With the same steps, the author obtains the data of the diachronic normalized frequency of *Confucianism* in GBBE, *Taoism* in GBAE, and Taoism in GBBE.

(2) Use the excel of MS Office to generate Figure 2 based on the collected data of the diachronic normalized frequency of *Confucianism* and *Taoism* in GBAE and GBBE. Figure 2 also lists the collected data of normalized frequency of *Confucianism* and *Taoism* at the bottom.

(3) With the similar steps in the above (1) and (2), the author gets data of the diachronic normalized frequency related to Confucius, and Lao Tzu in GBAE and GBBE, and then generates Figure 3 and Figure 4.

3. Results

Based on the collected data of Confucianism, Taoism, Confucius, and Lao Tzu in the corpora GBAE and GBBE, the author generates Figure 2, Figure 3 and Figure 4, which can provide answers to the previous three research questions.

Figure 2 indicates the diachronic change of Confucianism and Taoism in the UK and the USA from 1810s to 2000s. From Figure 2, it can be found that: (1) the frequency of Confucianism in the UK is generally higher than that in the USA, which means that Confucianism is received more attention in the UK than in the USA in the 200 years; (2) the general trend of Confucianism and Taoism in the UK and in the USA is rising, which means that Confucianism and Taoism are more and more popular in the two countries and have more powerful influence; (3) comparatively speaking, the frequency of Confucianism is higher than the frequency of Taoism both in the UK and in the USA, which means that Confucianism receives more attention in the two countries than Taoism; and (4) the spreading of Confucianism both in the UK and in the USA is much earlier than Taoism: Confucianism starts from 1810s while Taoism starts from 1860s.
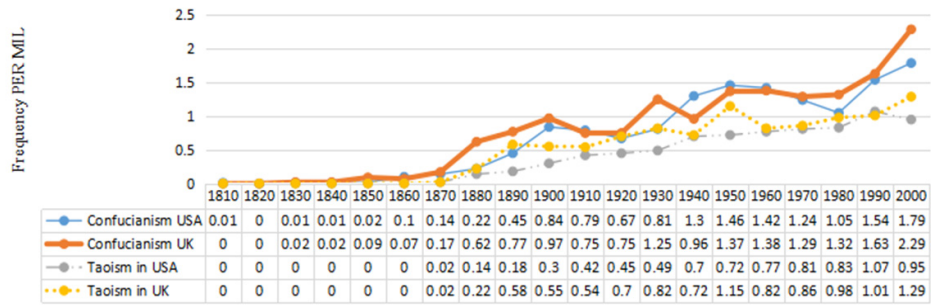
Figure 2. Diachronic change of Confucianism and Taoism in UAS and UK

Figure 3 indicates the diachronic frequency change of the two Chinese sages called Confucius and Lao Tzu in the UK and the USA. It can be found that: (1) the diachronic frequency of Confucius is dramatically higher than that of Lao Tzu, which means the ancient Chinese sage, Confucius is much more welcomed and has more influence in the USA and the UK than Lao Tzu; (2) the diachronic frequency of Confucius is much higher in the UK than that in the USA, which means Confucius is more popular and welcomed in the UK than in the USA; and (3) The spreading of Confucius is much earlier than Lao Tzu both in the USA and in the UK.
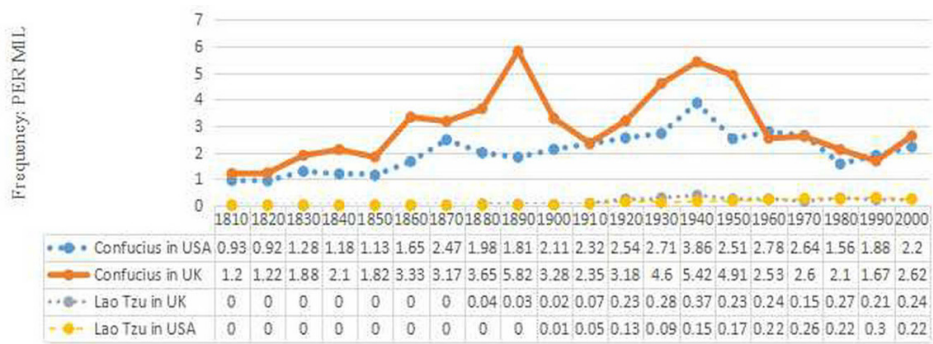

Figure 3. Diachronic change of Confucius and Lao Tzu in UAS and UK

Generally speaking, the diachronic frequency of Lao Tzu in the UK and the USA is lower than Confucius. In order to observe the spreading of Lao Tzu in the USA and the UK in details, the author generates the Figure 4 based on the collected data from corpora GBAE and GBBE. From Figure 4, it can be found that: (1) from 1900s to 1960s, the frequency of Lao Tzu in the UK is much higher than in the USA, which means Lao Tzu is more popular and receives more attention in the UK than in the USA; (2) in 1970s and 1990s, the frequency of Lao Tzu in the USA is higher than in the UK, which means Lao Tzu receives more attention in the USA in these periods than in the UK; and (3) in 1980s and 2000s, Lao Tzu receives more attention in the UK than in the USA.
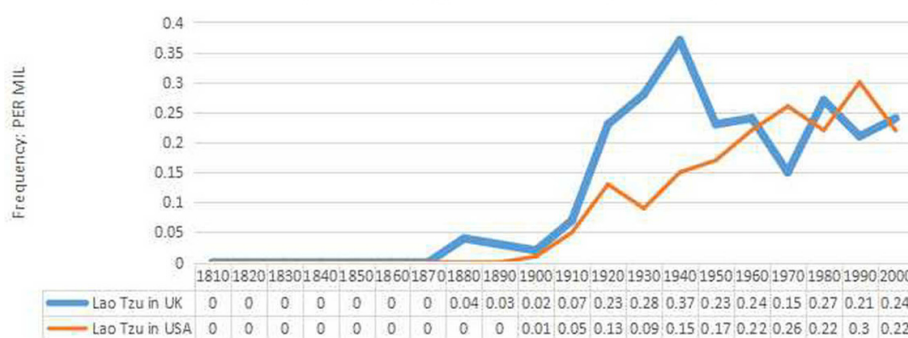

Figure 4. Diachronic change of Lao Tzu in UAS and UK

4. Discussion

4.1 Confucianism is More Popular than Taoism in the USA and in the UK

To the Chinese ancient philosophic thought, Confucianism and Taoism, they are introduced into the UK and the USA in the 17th century by Jesuit missionary through translation. In the following 300 years, they are studied in details in the two countries and published a myriad of English books related Confucianism and Taoism. However, which thought is more popular in the two countries? Many Chinese netizens, and even some scholars are interested in knowing the answers. Some of Chinese scholars try to answer this question based on their intuition and make the conclusion that Taoism is 'hotter' and 'more popular' than Confucianism in the USA (Note 11). To this question, it has not been quantitatively studied. Fortunately, there are big data of Google books N-gram that can objectively and quantitatively provide answers to the previous questions. In this research, it is quantitatively revealed that Confucianism is more popular that Taoism in the USA and the UK in periods from 1800s to 2000s, see Figure 2. One reason for explaining the much more popularity of Confucianism both in the USA and in the UK is that the study of Confucianism is much earlier than Taoism in the two countries: Confucianism was discussed and studied in the two countries, and published English books before 1810s while Taoism was discussed and studied, and published English books after 1880s in the UK and after 1890s in the USA, see Figure 2. Another reason is that Confucianism is much more concrete while Taoism is much more abstract and difficult to to understand for the ordinary people in the USA and the UK. Taoism refers to **ying and yang** (Note 12), and **five phases** (Note 13), which are quite difficult for ordinary Chinese to understand, let alone the British or Americans, so that Taoism is comparatively difficult to spread in the two countries. Fore the same reason, within the USA, Confucianism is more popular than Taoism and within the UK, Confucianism also is more popular than Taoism.

Consequently, the so-called assertion (Note 14) that Taoism is more popular in the UK, the USA or Westerns than Confucianism is not right.

4.2 Confucius is More Popular than Lao Tzu in the UK and the USA

It is revealed that Confucius is more popular than Lao Tzu in the UK and the USA, see Figure 3. Confucius represents the Confucianism while Lao Tzu represents Taoism. Confucianism is more popular than Taoism in the UK and the USA, see Section 4.1. Consequently, Confucius is more popular than Lao Tzu both in the UK and the USA. Another reason is that the thought of Confucius is much easier to be understood than that of Lao Tzu, by the ordinary British and Americans. Hence, the so-called assertion (Note 15) that Lao Tzu is more welcomed by British or Americans than Confucius is not correct.

4.3 Confucianism and Taoism Are More Welcomed in the UK than in the USA

It is quantitatively revealed that Confucianism and Taoism are more welcomed, see Figure 2 and Figure 3. There are two reasons for the phenomena: (1) the UK is an ancient country and has much longer history than the USA so that the UK has more traditional convention to study, borrow and absorb foreign culture and ideas than the USA; (2) the UK has the most powerful economic and military strength in the 17th–c18th century, which provides the solid foundation for absorbing and spreading Confucianism and Taoism in the UK. Britain became a world power where the sun never sets, which further reinforces the spreading of Confucianism and Taoism in the UK and British Commonwealth. This is revealed in the higher frequency of Confucianism and Taoism in the English books published in the UK than in the USA.

Though there is difference in the diachronic spreading of Confucianism and Taoism between the UK and the USA, the general spreading trend of them is rising, which means the influence of Confucianism and Taoism is rising in the two countries.

5. Conclusion

Based on the collected data from GBAE and GBBE, it is found that: (1) Confucianism and Taoism are more welcomed both in the UK than in the USA; (2) compared with Taoism, Confucianism is more popular both in the UK and in the USA than Taoism; (3) Confucius is more popular both in the UK and the USA than Lao Tzu; and (4) the general spreading trend of Confucianism and Taoism is rising in 1810s–2000s both in the UK and the USA. The findings can quantitatively answer the previous research questions and prove the in-correction of aforementioned so-called assertions based on intuition. The present research also provides the detailed data for observing the diachronic spreading status and future trend of Confucianism and Taoism in the UK and the USA, from 1800s to 2000s.

Acknowledgments

References

Aiden, E., & Michel, J. B. (2013). *Uncharted: Big Data as a Lens on Human Culture*. New York: Riverhead Books.

Berry, D. M. (2011). The computational turn: thinking about the digital humanities. *Culture Machine*, *12*, 1–22. https://doi.org/10.3917/presa.185.0209

Caruanagalizia, P. (2016). Politics and the German language: testing Orwell's hypothesis using the Google N-gram corpus. *Digital Scholarship in the Humanities*, *31*(3), 441–456. https://doi.org/10.1093/llc/fqv011

Fung, Y. M. (2008). Problematizing Contemporary Confucianism in East Asia. In J. Richey (Ed.), *Teaching Confucianism*. Oxford: Oxford University Press.

Gooding, P. (2013). Mass digitization and the garbage dump: the conflicting needs of quantitative and qualitative methods. *Literary & Linguistic Computing*, *28*(3), 425–431. https://doi.org/10.1093/llc/fqs054

Greenfield, P. M. (2013). The changing psychology of culture from 1800 through 2000. *Pshchological Science*, *24*(9), 1722–1731. https://doi.org/10.1177/0956797613479387

Juola, P. (2013). Using the Google N-gram corpus to measure cultural complexity. *Literary & Linguistic Computing*, *28*(4), 668–675. https://doi.org/10.1093/llc/fqt017

Kesebir, S., & Kesebir, P. (2012). The cultural salience of moral character and virtue declined in twentieth century. *The Journal of Positive Psychology*, *7*(6), 471–480. https://doi.org/10.1080/17439760.2012.715182

Lamoreaux, M., & Morling, B. (2012). Outside the head and outside individualism-collectivism: Further meta-analyses of cultural products. *Journal of Cross-Cultural Psychology*, *43*(2), 299–327. https://doi.org/10.1177/0022022110385234

Michel, J. B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., The Google Books Team, Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M., Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, *331*(6014), 176–182. https://doi.org/10.1126/science.1199644

Twenge, J. M., Campbell, K. W., & Gentile, B. (2012a). Increases in Individualistic words and phrases in American Books, 1960–2008. *PLoS ONE*, *7*(7), 1–5. https://doi.org/10.1371/journal.pone.0040181

Twenge, J. M., Campbell, K. W., & Gentile, B. (2012b). Changes in Pronoun Use in American books and the Rise of Individualism, 1960–2008. *Journal of Cross-Cultural Psychology*, *44*(3), 406–415. https://doi.org10.1177/0022022112455100

Zhang, T. (2011). *Confucius in the United State of America: the image of Confucius in American Newspapers from 1984*. Beijing: Beijing University Press.

Notes

Note 1. See http://en.people.cn/102775/205267/index.html#

Note 2. See http://www.taoist.org/the UK/

Note 3. See http://www.taoist.org/the UK/

Note 4. See https://en.wikipedia.org/wiki/Tao_Te_Ching#CITEREFChan2013

Note 5. See https://bbs.hupu.com/14588008.html

Note 6. See http://3g.163.com/dy/article/D9M2H3J80525DM59.html

Note 7. See https://googlebooks.byu.edu/compare-googleBooks.asp

Note 8. See https://googlebooks.byu.edu/

Note 9. See https://googlebooks.byu.edu/compare-googleBooks.asp

Note 10. See https://googlebooks.byu.edu/x.asp

Note 11. See http://www.360doc.com/content/12/0226/02/5079158_189672707.shtml

Note 12. https://en.wikipedia.org/wiki/Yin_and_yang

Note 13. https://en.wikipedia.org/wiki/Wu_Xing

Note 14. See https://zhidao.baidu.com/question/497348776.html?fr=iks&word=%C8%E5%BC%D2%BA%CD%B5%C0%BC%D2%D4%DA%CE%F7%B7%BD&ie=gbk

Note 15. See https://www.douban.com/group/topic/10618548/

Copyrights