

# A big-data analysis of political rhetoric relating the developments of the United States, China, and global powers

Big-data  
analysis of  
political  
rhetoric

227

Patrick Carter  
*Appleby College, Oakville, Canada*  
Jeffrie Wang  
*The University of California, Berkeley, Berkeley, USA, and*  
Davis Chau  
*Innovative Model and Industry Development Institute, China*

Received 22 March 2020  
Revised 6 September 2020  
Accepted 15 September 2020

## Abstract

**Purpose** – The similarities between the developments of the United States (U.S.) and China into global powers (countries with global economic, military, and political influence) can be analyzed through big data analysis from both countries. The purpose of this paper is to examine whether or not China is on the same path to becoming a world power like what the U.S. did one hundred years ago.

**Design/methodology/approach** – The data of this study is drawn from political rhetoric and linguistic analysis by using “big data” technology to identify the most common words and political trends over time from speeches made by the U.S. and Chinese leaders from three periods, including 1905-1945 in U.S., 1977-2017 in U.S. and 1977-2017 in China.

**Findings** – Rhetoric relating to national identity was most common amongst Chinese and the U.S. leaders over time. The differences between the early-modern U.S. and the current U.S. showed the behavioral changes of countries as they become powerful. It is concluded that China is not a world power at this stage. Yet, it is currently on the path towards becoming one, and is already reflecting characteristics of present-day U.S., a current world power.

**Originality/value** – This paper presents a novel approach to analyze historical documents through big data text mining, a methodology scarcely used in historical studies. It highlights how China as of now is most likely in a transitional stage of becoming a world power.

**Keywords** Political rhetoric, Big data, United States, China, Global powers

**Paper type** Research paper

## Introduction

According to popular legend, the French Emperor Napoleon once said about China, “let her sleep, for when she wakes, she will shake the world.” Despite Napoleon having never actually uttered these famous words, they appear to have come true. In recent years, China has dominated the headlines of global news as the Asian giant has caught the attention of all world leaders. Academics debate and write on the possibilities of a Chinese superpower with global power projection while making their countless analyses of China’s situation. Journalists continue to produce an unending amount of reports and articles heralding the



© Patrick Carter, Jeffrie Wang and Davis Chau. Published in *Public Administration and Policy*. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) license. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this license may be seen at <http://creativecommons.org/licences/by/4.0/legalcode>

Public Administration and Policy  
Vol. 23 No. 3, 2020  
pp. 227-243  
Emerald Publishing Limited  
1727-2645  
DOI 10.1108/PAP-03-2020-0018

---

successes of the Chinese government in turning China into a world power, or criticizing a more assertive Chinese leadership, or warning of the rise of China and decline of the United States. And it is not hard to understand why. China's rapidly growing economy, expanding military, and greater participation in global affairs have shaken the world and shifted the balance of world power. However, while as spectacular as China's growth has been, this sort of progress is nothing new. The emergence of the United States onto the world stage also came about in a fashion similar to what China is doing today.

### Literature review

In the modern world, a country's power can be analyzed through two different lenses: hard power and soft power. Joseph Nye, the Harvard political scientist behind the theories of hard power and soft power, described hard power as coercion of other countries using military or economic power, and soft power as "[a country's] culture. . .its values. . .and its policies" being used to coerce foreign powers (Nye, 2009, p. 161). When a country maintains both a powerful military and significant soft power, they are more likely to be considered a global power. Through this, similarities between the rises of U.S. and China into superpowers can be drawn.

Most visibly, U.S.'s hard power developed in a manner that can also be seen in China today. In *The Rise and Fall of the Great Powers*, Kennedy (2017) attributed U.S.'s development to its "unique combination of having both a populous *and* a highly industrialized society" (p. 257). This would be proven in the late-nineteenth century as the American population dramatically increased from a population of 50,189,209 people in 1880 to 92,228,496 people in 1910, or a 184 percent increase in only thirty years (Bureau of the Census, 1975). This substantial population ensured the existence of a large labor force, which allowed U.S. to become the world's factory, much like China is today. American manufacturing increased from "20 percent of U.S. exports in 1890 to 35 percent by 1900 and nearly 50 percent by 1913" (Irwin, 2006). Then to build off of this huge manufacturing increase, U.S. established "a customs union for the Western Hemisphere, a free-trading sphere that would become an economic counterpart to the Monroe Doctrine" as said by UC Irvine history professor emerita Emily S. Rosenberg (Rosenberg, 1982, p. 53). By keeping mass-produced American goods cheaper than European counterparts, the economy soared and quickly became the world's largest economy.

Interestingly, this same story towards economic prosperity applies to China as well. In 1949, China already had a massive population of 541.67 million people, but by 1979 (30 years) the population had increased by 429.25 million to 970.92 million people, or an increase of approximately 180 percent (Jowett, 1984). This massive workforce was then utilized through the creation of special economic zones or SEZs where free trade was practiced with low tariffs, allowing for exports to increase through cheap Chinese products like past U.S. did with their goods (Lin and Shen, 2018). Chinese free markets thus allowed for industry to compose 29.303 percent of Chinese GDP by 2017 as reported by The World Bank (2019), an amount very similar to that seen in U.S. For example, in 2016, the World Intellectual Property Organization (WIPO) found that China produced the most patent applications in the world at more than double U.S. due to increased R&D funding from a developed economy (Santacreu and Zhu, 2018). Similarly, for U.S. in 1900, they were the drivers of "the marvelously convenient evolution of modern technology" as they, with inventors like Edison, rose to the top of patent-filing nations (Kennedy, 2017, p. 312). This relation is perhaps evidence of how similar the two countries' economic strategies are, as they both reached the same result of innovation. Clearly, the two countries have had very close paths of economic development.

Soft power has also seen similar developments. One major area of relation between the global powers is the emergences of distinctive national identities built on patriotic principles.

Recently, in China's case, the government has worked extensively to develop a unique Chinese identity for its citizens and Chinese people abroad, according to a professor of Chinese politics at Harvard University (Perry, 2015). China has also been developing and exporting the term "Chinese nation" or "Zhonghua minzu" and encouraging "popular sovereignty, or minzhu" by promoting a "growing pride [in] the Chinese nation" not only to "a domestic constituency, but also to overseas Chinese" (Perry, 2015, p. 910). As part of this global development of national identity, the establishment of hundreds of Confucian Institutes that teach Chinese culture around the world is China's investment in growing its image around the world and in the eyes of Chinese people (Ungerer *et al.*, 2010). Through these means, China intends to establish a growing pride amongst Chinese people towards their country, a clear indicator of how much the government values it for its development. All in all, "the three liberal pillars – unrestricted trade and investment, free enterprise, and free flow of cultural exchange – became the intellectual rationale for American expansion", and now quite possibly, of Chinese expansion (Rosenberg, 1982, p. 37). During the early twentieth century, the American identity became increasingly developed with events like Roosevelt's sending of the Great White Fleet around the world, the adopting of the current national anthem, and Buffalo Bill leading the cultural expansion (Rosenberg, 1982). From these notable examples, there are clearly many commonalities the actions undertaken by U.S. and China as they developed their soft power capabilities.

However, while these similarities are readily available and significant literature has already been dedicating towards considering these results, one area less studied is political rhetoric. China economy and culture may be growing in the modern world, but what do its leaders have to say about China's policies and direction? What did American leaders say about their own policies? This is worth considering specifically because political speeches have historically played a large role in deciding the policies of countries as this is indeed the case for U.S. A research report by Matthew Eshbaugh-Soha, a professor at the University of North Texas, concluded that through elements like policy signals in speeches, "presidential speeches are an effective means of influence over bureaucratic activity" and thus shape domestic policy (Eshbaugh-Soha, 2008). Further research also stresses the importance of political speeches in shaping not only domestic affairs, but also foreign policy. Michelle Murray Yang's research on President Nixon's speeches in China found that presidential speeches "played a vital role in cultivating a personal and professional relationship" between the U.S. and China (Yang, 2011). These speeches then in turn shape how future leaders decide to frame their speeches on both domestic and foreign policy and how they may convince their audiences (Yang, 2011; Osborne, 1980). Thus, taking into consideration the importance of speeches in American policy, it is worth considering that a similar case exists for China.

Speeches and the rhetoric used within them even hold significant importance to their effectiveness and in turn how they can affect policy. In a case of study of President Kennedy's speeches, Osborne found that "patterns of arrangement do exist in presidential speechmaking" and that "presidential speeches differ according to audience" with humorous words being especially with college speeches (Osborne, 1980). Other researchers have also noticed the importance of words in speeches and have based their research around it as well. In a study on metaphors in Taiwanese presidential speeches, Lu and Ahrens used character counts and word analyses to conclude that "linguistic studies of metaphors probe conceptual-level mappings as well as. . .[presidential] language via metaphor use for their own worldview" (Lu and Ahrens, 2008, p. 401). Or another example being Seider's research into business ideologies, which while not strictly political, made conclusions into business through content analysis of words that effectively show the significance of word analysis (Seider, 1974). Evidently, political rhetoric has been and is a source of significant insight into the actions of political leaders and why they choose to do the things they do. Knowing this, rhetoric serves as a base for significant insight into a country's policies.

---

While there is significant literature pointing to the relevance of linguistic analysis, its usage to specifically compare trends between countries has seen little usage among researchers. Therefore, this paper seeks to address this and conduct further research into this subject. Through analyzing the rhetoric and the types of words used in political speeches, trends can be identified between the two countries, and such trends can be used to compare and relate the unique histories of these two countries as they emerged as 21<sup>st</sup> century world powers.

While China does indeed have differences from U.S. like in ideology, they do not detract from the similar paths they took and are taking to world power. American and Chinese analysts and politicians today continue to try and lead their countries towards the path they believe to be correct. They each have their own attitudes towards their countries. They give speeches and make powerful oratories, but by continuously adding to the comprehension of the connections between the developments of U.S. and China, there can be a better understanding of China today and where it may choose to go tomorrow.

### **Research methodology**

To add to the investigation of how great powers develop and what similarities China and U.S. may have in becoming global powers, linguistic analysis was used. This method of analysis was conducted using “big data” software, or programs that allow for the efficient gathering and content analysis of text from large quantities of primary sources. Using programs such as RapidMiner, more than a thousand speeches from U.S. and Chinese leaders were downloaded from public online databases like the Miller Center. It is also worth noting that the data base from which the Chinese speeches were sourced had already translated the speeches to English. Each downloaded document was classified into their respective time period of 1905-1945 U.S. (its developmental period), 1977-2017 U.S. (its modern period), or 1977-2017 China (its current developmental period). These time periods were decided upon because 1905-1945 and 1977-2017 are approximately the ranges of time when U.S. and China developed (or are developing) into world powers, respectively. 1977-2017 U.S. was also added as a means by which to measure how U.S. has changed since its developmental phase and how similar modern U.S. is to present-day China within the same time frame.

RapidMiner scans all the documents within each time period without the need for manual examination by counting how many times every word appeared in a transcript. Trivial words that have no meaning of themselves such as “because” or “the” were ignored during this process. This process would continue until every speech was scanned by the program (in a time period). Once all downloaded documents for a time period were scanned, the program then produced the sixty-three most-common words in each time period alongside how many times those words appeared across all documents in a certain time period. The range of only the top sixty-three was chosen because beyond that number, words appeared in amounts too small and similar to each other to make observations.

The data were then manually analyzed by classifying each word under one of five categories: economy (E), military (M), national identity (N), societal development (S), and foreign affairs (F). Economy contains words that relate to economic or commercial affairs. Military contains words relating to a country’s strength and power. National identity covers words that reflect the national image the politician is trying to create, the morals that the country chooses to stand by, and the public image of the country. Societal development contains words that relate to the general development of its country and governmental functions. Foreign affairs contain terms that are related to how a country deals with foreign countries and its position in the international community.

For each of the three time periods (U.S. 1905-1945, U.S. 1977-2017 and China 1977-2017), the amount of words in each of the five classifications were totaled. Based on these numbers,

it can be seen what areas political leaders from U.S. and China placed most of their focus on. For example, if the topic of the economy made up the highest proportion of the content discussed in speeches, then it is clear that economy is the premier focus of the leaders. Furthermore, the average number of times each word category appeared per speech for each time period were analyzed. This offered another lens by which to consider the attitudes of global powers and developing powers. In addition, trend analysis was conducted to see how each time period changed over the course of time and if any comparison and conclusion could be drawn from this.

## Data and results

Table 1 shows the top twenty most common terms from each time period with the amount of times they appeared, ordered from the most frequent to the least frequent (Appendix). Each word is labeled under one of the five categories of national identity (N), societal development (S), foreign affairs (F), economy (E), and military (M). Tables 2-4 show the amount of words from Table 1 under each classification.

The most visible trend amongst these three tables is that national identity is by far the most frequent of the five categories, with it having approximately fifty percent of the top twenty most common words across all time periods as seen in Tables 2-4. Furthermore, the most common national identity words are all concentrated towards the top of the time periods in Table 1 as the most common words, with national identity terms rarely appearing towards the bottom of the chart in Table 1.

Furthermore, there appears to be a substantial gap in word-frequency between national identity words and words in other categories. For example, in China 1977-2017, the national identity word, “China”, appeared 25,912 times while the next non-national identity word, “economy”, appeared only 12,514 times (Appendix A). U.S. 1977-2017, however, appears to be the exception to this rule as economic terms seem also to have a similarly high usage compared to national identity words during that time period. Overall though, this trend of

U.S. 1905-1945 Word	Amount #	U.S. 1977-2017 Word	Amount #	China 1977-2017 Word	Amount #
state(s) - N	5190	america - N	4939	china - N	25912
nation - N	3964	work - E	4430	develop - N	17562
great(er) - N	3852	country - N	3835	party - N	14635
government - S	2980	economy - E	3452	development - N	14116
people - N	2005	nation - N	2544	people - N	13292
united - N	1685	world - F	2424	economy - E	12514
world - F	1497	presidential - S	2394	promote - S	10573
country - N	1483	life - N	2322	strength - M	10532
congress - S	1424	state(s) - N	2272	comrade(s) - N	8426
national - N	1026	freedom - N	2044	build - S	7992
business - E	1018	united - N	1949	central - S	7899
work - E	1012	strength - M	1933	work - E	7591
peace - S	950	economic - E	1880	chinese - N	7214
power - M	890	government - S	1783	economic - E	7058
free - N	877	people - N	1567	national - N	5779
nations - N	772	security - M	1553	committee - S	5763
service - N	765	americans - N	1426	country - N	5269
purpose(s) - N	749	effort(s) - N	1268	political - S	4968
force(s) - M	693	peace - S	1223	world - F	4961
system - S	675	Congress - S	1179	system - S	4806

**Table 1.**  
Top 20 most common  
words for each time  
period

high uses of national identity terms is indicative of the prevalence of national identity in American and Chinese politicians' speeches, and its importance in the development of a country. After national identity, societal development then composed the second most every time, followed by the military, economy, and foreign affairs categories, which all appeared in similarly low amounts. These trends can also be seen through a macro view when classifying all of the sixty-three terms in Appendix and including how many times each categorized word appeared across all speeches.

Figures 1-3 show what percentage each of the five classifications make up of the total amount of words analyzed by RapidMiner and are indicating similar trends to those seen in Tables 2-4. Like Tables 2-4 indicated, national identity evidently again composes by far the greatest amount of words used by American and Chinese politicians across all three time periods, followed by societal development and then the other three categories. Other than reinforcing trends seen in Tables 2-4; however, these graphs also highlight more detailed similarities and differences between the three time periods that Tables 2-4 did not show.

Firstly, it is evident that there are substantial differences between the U.S. 1905-1945 and the U.S. 1977-2017. Between the two time periods, societal development and foreign relations remained equal in frequency; however, there are very distinct differences in national identity, economy, and military. National identity words appeared to drop considerably in usage while economy or military terms became significantly more frequent from 1977 to 2017 compared to 1905 to 1945. This not only shows that modern U.S. leaders are much more diverse in what they focus on in their speeches, but also that they perhaps have placed significantly less importance on national identity. This could be indicative that modern U.S., as an established world power, has much less need for national identity than when they were becoming a global

**Table 2.**  
Classification of five  
categories of 20 words  
– U.S. 1905-1945

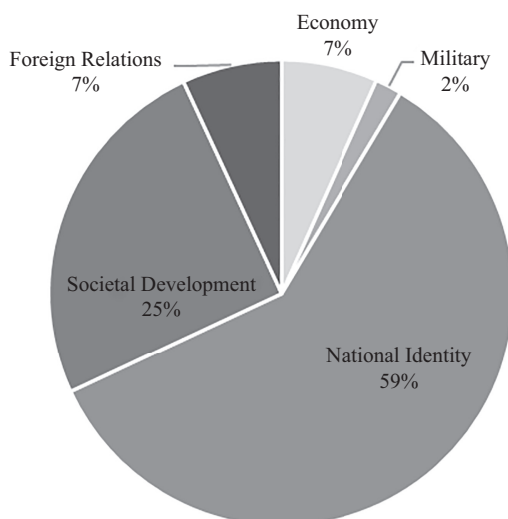
Type of classification	Amount #
N (national identity)	11
S (societal development)	4
F (foreign affairs)	1
E (economy)	2
M (military)	2

**Table 3.**  
Classification of five  
categories of 20 words  
– U.S. 1977-2017

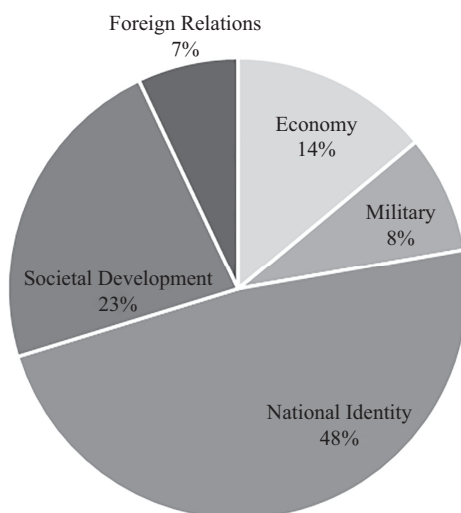
Type of classification	Amount #
N (national identity)	10
S (societal development)	4
F (foreign affairs)	1
E (economy)	3
M (military)	2

**Table 4.**  
Classification of five  
categories of 20 words  
– China 1977-2017

Type of classification	Amount #
N (national identity)	9
S (societal development)	6
F (foreign affairs)	1
E (economy)	3
M (military)	1



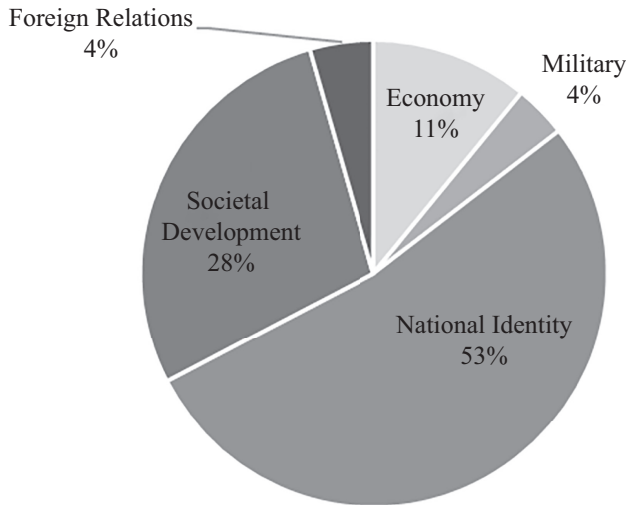
**Figure 1.**  
Classification of five  
categories by  
percentage – U.S.  
1905-1945



**Figure 2.**  
Classification of five  
categories by  
percentage – U.S.  
1977-2017

power. China from 1977 to 2017, generally accepted as a rising power, also supports this idea as they have similar degrees of societal development and foreign relations usage to both U.S. time periods. However, national identity, military, and economy appear in percentages in-between those of the U.S. 1905-1945 and of U.S. 1977-2017. China appears to place significant focuses on economy and military like modern U.S., but still utilizes much more national identity words than modern U.S. These trends are further continued when looking at how many words are used in each category for each time period, and how many times on average words of each category appeared in each speech.

Tables 5-7 show the average per speech usage of each word category. Similar trends as before are seen here; between U.S.'s developmental period (1905-1945) and its modern period



**Figure 3.**  
Classification of five categories by percentage - China 1977-2017

(1977-2017), national identity dropped from an average of 157 to 147 per speech. China sits in between these values at around 150 national identity terms per speech from 1977 to 2017. This further confirms the information presented by Figures 1-3 and highlights how China is possibly in a transitional stage, as its focuses are shifting to increased focuses on economy, military, and less on national identity like the modern U.S., a global power. This trend is further highlighted when each countries' political focuses on national identity, the most prevalent word category, is analyzed over time.

Figures 4-6 reflect each countries' usages of national identity in speeches over time, with the dotted lines being the overall average number of national identity words per speech as from Tables 5-7. It is clear that in Figure 4, U.S. speeches in 1905 used very high amounts of national identity words, but gradually made less use of national identity until around 150 words per speech in 1945. Figure 5 shows how U.S. speeches in 1977 largely continued this

**Table 5.**  
Per speech usage - U.S. 1905-1945

National Identity	Societal Development	Foreign Relations	Economy	Military
30477 ( <b>≈157/speech</b> )	12817 ( <b>≈66/speech</b> )	3553 ( <b>≈18 per speech</b> )	3449 ( <b>≈17/speech</b> )	947 ( <b>≈5/speech</b> )

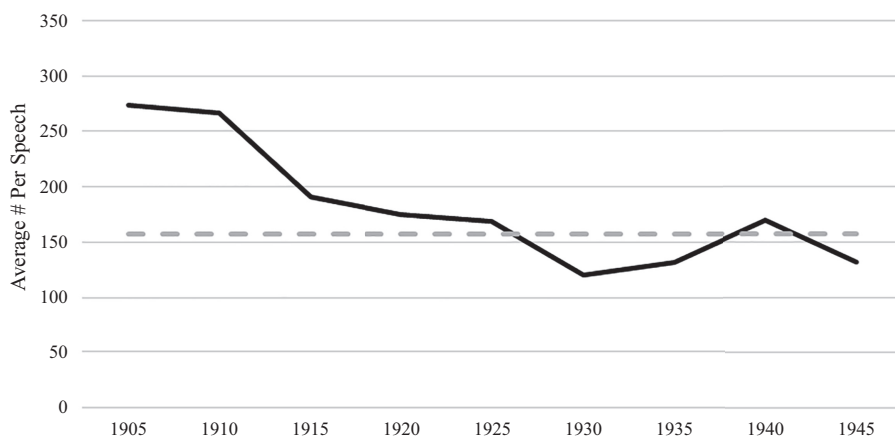
**Table 6.**  
Per speech usage - U.S. 1977-2017

National Identity	Societal Development	Foreign Relations	Economy	Military
34398 ( <b>≈147/speech</b> )	16244 ( <b>≈69/speech</b> )	5071 ( <b>≈22 per speech</b> )	10019 ( <b>≈43/speech</b> )	5916 ( <b>≈25/speech</b> )

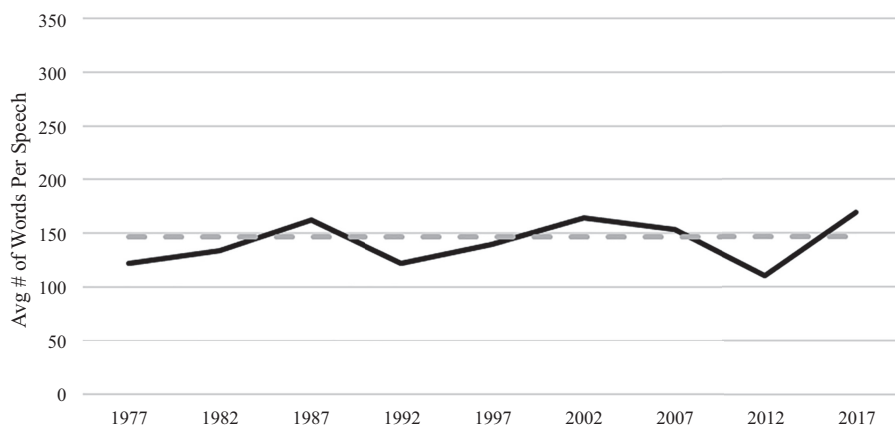
**Table 7.**  
Per speech usage - China 1977-2017

National Identity	Societal Development	Foreign Relations	Economy	Military
152991 ( <b>≈150/speech</b> )	81710 ( <b>≈80/speech</b> )	12880 ( <b>≈13 per speech</b> )	31572 ( <b>≈31/speech</b> )	10532 ( <b>≈10/speech</b> )

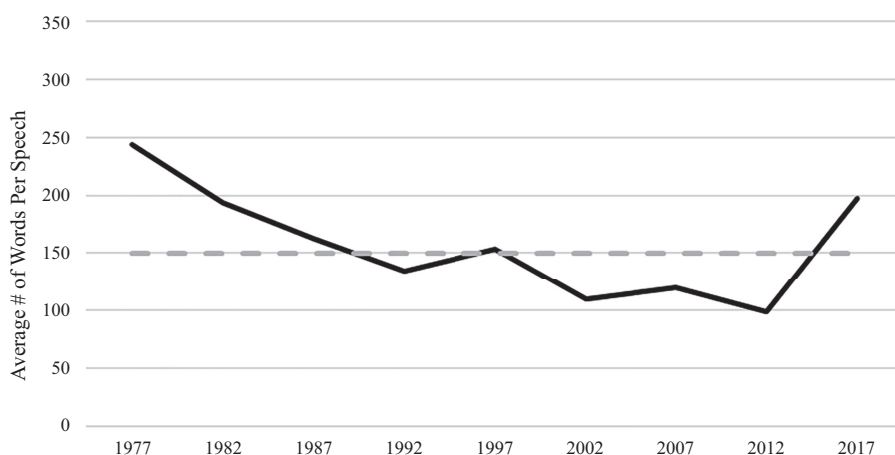




**Figure 4.**  
National identity usage  
– U.S. 1905 to 1945



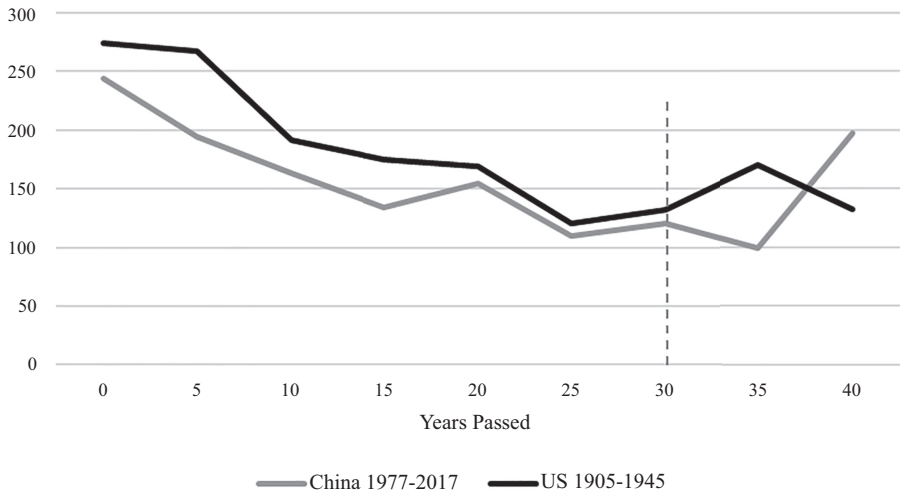
**Figure 5.**  
National identity usage  
– U.S. 1977 to 2017



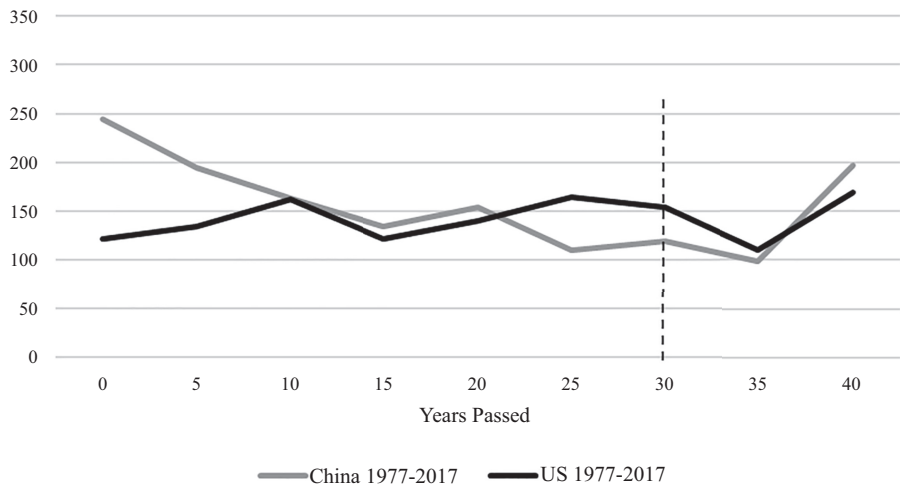
**Figure 6.**  
National identity usage  
– China 1977 to 2017

trend from 1945 with decreased importance placed upon national identity, and it remained largely below around 150 with a resurgence towards the present-day. It is noteworthy that the U.S. 2017 surge in national identity usage seems to contradict previous evidence that highlights how when a global power develops, national identity becomes less of a focus. Overall though, from these two figures, it is clear that there is a significant contrast between the importance placed upon national identity for presidents like Theodore Roosevelt in 1905 and for presidents like Barack Obama in 2012, showing the differences in attitudes between a growing power and an established one. China from 1977 to 2017 appears to be in a similar position to that of U.S. from 1905 to 1945, with political speeches in 1977 including high amounts of national identity words then steadily dropping until experiencing a resurgence in 2017.

When overlaying these trends over time onto each other as shown in Figures 7-8, multiple observations can be made. It is evident that the decreases in focuses on national identity



**Figure 7.**  
National identity over  
40 years – U.S. 1905-  
1945 and China  
1977-2017



**Figure 8.**  
National identity over  
40 years – U.S. 1977-  
2017 and China  
1977-2017

between U.S. 1905-1945 and China 1977-2017 match each other very closely in Figure 7, with them both following very similar trends until about 30 years, where they start to diverge. This shows how China's development until now, from a national identity usage standpoint, closely matches that of U.S.'s development and is perhaps an indication of where China stands in its current position in development. This point is further continued in Figure 8, where the trends for China 1977-2017 and for U.S. 1977-2017 do not appear to follow each other as closely as in Figure 7, and only match after 30 years. This indicates that as China developed over 40 years, it, like U.S. 1905-1945, decreased its focus on building a national identity until it possibly reached a position of power around 30 years. From here, its reversal to matching the U.S. 1977-2017 is evidence of China shifting its attitude towards that of a global power. While it would be assumed that both countries, as world powers, would continue to decrease in national identity usage, their common increases after 30 years are perhaps the two countries attempting to match with one another as China emerges as a global power. The more unique (compared to their predecessors) national policy actions of present-day's U.S. President Trump and Chinese President Xi could also have played a role in this sudden change in trends. Another interesting point is that China appears to be very quickly shifting attitudes towards that of a superpower while U.S. took much longer to do so, as only after 30 years their trends are looking more closely to modern U.S.

## Discussion

The results of this research offer multiple conclusions about the nature of the development of global powers, and the current position of China as an emerging world power.

Firstly, the data from the three time periods (U.S. 1905-1945, U.S. 1977-2017, and China 1977-2017) make it foremost clear that national identity is the most prevalent word category used in political speeches, which is understandable as "national identities change and redistribute the *motivation* to use power, just as national power changes and redistributes the *capacity* to use power", making them key to forming policy (Nau, 2001, p. 585). It also appears conclusive that no matter the stage of development for a country, whether a developing power or an established global power, the distribution of political focus is largely as follows: national identity being the most used, then societal development, then economy, and then foreign relations. The particularly high national identity usage for U.S. 1905-1945 may support the idea that national identity is key to starting and developing a global power. National identity was also consistently the most common focus in speeches as seen in the data. This trend proves how important national identity is to politicians, which seems to be as important as factors like a country's population or military size. The research goes further by indicating that national identity may be more important to politicians than the other traditionally regarded classifications like the economy. Aside from identifying the most prevalent political focuses of a country, it importantly highlights the behavior of global powers.

The data proves that there is a noticeable change in the attitudes of a country, or in this case U.S., as it transitions into the role of a world power. For example, the distinct shift towards a more military and economy focus from U.S. 1905-1945 to U.S. 1977-2017 shows the differences between how a developing power focuses its policy and how an already developed power does so. Or as notably, the dramatic drop in national identity word usage after 1977 shows how a global power may act differently from a developing one, and what could be an indicator for a developing global power. A global power may already have an established national identity that propelled it to a position of strength, and as such may find it more useful to focus politically on other matters. This makes sense given national identity's importance for early U.S. and decreased prevalence in later speeches of American leaders.

When considering how these indicators of a global power apply to China, similar developments can be seen. China's political focuses being overall "in-between" those of U.S.

1905-1945 and U.S. 1977-2017 indicate that it is possibly in a sort of transitional phase, a conclusion supported by its current position in world affairs. The changes that occurred between U.S. 1905-1945 and U.S. 1977-2017 could possibly be occurring in China right now, explaining why its distribution of political focuses, like national identity appearing on average 150 times per speech (U.S. 1905-1945 is 157 times, U.S. 1977-2017 is 147), appear in-between developmental U.S. and present U.S. While many may claim China's rapidly growing economy and global influence leverages it the title of a global power, its political focuses tend to indicate that it is still transitioning from a developing power to a developed one.

The trends over time as seen in the data also support this conclusion. U.S. 1905-1945 and China 1977-2017's both general decreases in national identity usage over time indicate that Chinese policies are much more reflective of a growing power than an already established global power. Both U.S. 1905-1945 and China beginning their developments with a very high prevalence of national identity in political speeches indicate that these two countries were working to establish their national identity as they began their development into global powers. This supports the previous conclusion that national identity is critical to starting the growth of an influential country; however, the trends over time add to this conclusion by also making it evident that national identity is possibly less important to political leaders in *maintaining* a global power. With this idea in mind, it makes sense that China today still has higher uses of national identity than U.S. because they are not yet at the point of maintaining their global influence. They are still trying to develop it. If China were already a global power, then it is likely that its political rhetoric like those regarding national identity would more closely resemble that of an already established world power – the present-day U.S. – rather than U.S. in 1905. Overall, from the point of view of the content of political speeches, it appears that China does not share the same mentality of the current global power.

These conclusions are not to say, however, that China is not currently in a position of power nor that it will not reach that position soon. Firstly, the data corroborates the idea that China is nearing a global power position because it is evident that China's political focuses are still generally diverse and less national identity focused like current U.S. Also, the trends of China's decreasing national identity indicate that it has long been on the path of development as it shares the same attitude towards national identity as U.S. did during its respective development from 1905 to 1945. In ten years from now, it is very possible that China's political focuses could shift much more to match present-day U.S. Indeed, the beginnings of that can be seen in the trends over time like in [Figures 7-8](#) at year 30, where China's national identity trendline diverges from U.S. 1905-1945 and instead switches to match that of U.S. 1977-2017. It is possible that China has reached a point where it is now approaching global power status, as its attitude towards national identity and policies seem to reflect that. It is interesting though that while these two countries match at year 30, they match in that they are both increasing their uses of national identity. This is perhaps indicative that the current leadership of U.S. and China today have very different interests from past leaders, or possibly that with the likely emergence of two global powers – U.S. and China – competing with each other, both countries are boosting their national identities to become stronger as world powers.

With these conclusions in mind, it is important to understand that U.S. and China still hold many differences that can make comparisons difficult. China is a country with a long prestigious history that it keeps in mind when conducting policy – as evidenced by its many Confucius Institutes around the world and President's Xi's many speeches promoting ancient Chinese art or the position of power that China will reclaim. U.S. is a relatively young country that has no history of being a global power until 1945, and it is not trying to reclaim any sort of past strength because U.S. is still the world's preeminent power. This differences in mentality and history between the two countries mean that something like national identity could possibly mean different things to both countries and politicians who could utilize rhetoric in different ways. However, while this is important to understand, it does not detract from this

study as it is clear from the data that there are indeed similarities and trends that can be identified. Further linguistic analyses into Chinese and U.S. political rhetoric could always be used to further understand what comparisons between China and U.S. can be made rhetorically sound.

Furthermore, despite this study, there remains a significant amount of research to be done to investigate why national identity plays such a significant role in political rhetoric, why national identity decreases in prevalence as countries become more globally influential, and if China is set to become a world power. This research's data identifies the existence of these patterns like national identity being important and it is decreasing in use over time, but why these trends exist requires further inquiry. Future research in this area could produce more insights into the nature of global powers. Likewise, to confirm the similarities between the respective rises of U.S. and China, big data analyses of other great powers like the United Kingdom or Germany could be done to find whether or not national identity is a unique focus for U.S. and China or simply important for all global powers.

Also, for this study, the amount of speeches analyzed differed between each time period. The data for China 1977-2017, for example, came from big data analysis of 1000 speeches, while the data for U.S. 1977-2017 came from 300 speeches. Equalizing these numbers could possibly make for more fair and accurate data sets. Nonetheless, this also reflects the relatively larger number of Chinese speeches, and the authors would like to incorporate all relevant data for analysis. Another limitation is that the data were derived from primary sources and organized through manual examination. Thus, even with set definitions for each of the five classifications, the amount of words under national identity or societal development, for example, could vary depending on the researcher. This could result in researchers coming upon differing conclusions, but this issue is offset somewhat through the publication of all data for anyone to analyze and use for research.

Therefore, the data and method of collecting itself is critical to the importance of this paper. The collection of data and use in analysis of countries highlights the capabilities for big data to contribute to social science research through the digital humanities. Big data has so far not been widely used to investigate the history or nature of global powers. Hence this paper can contribute by filling in this gap. Also, by proving the applicability of big data to the social sciences, it is showing how versatile the new technology is. In this research, big data analysis has been used to explore possible similarities between the developments of two great powers. It can also be used for research into other subjects. For example, one could investigate the specific diction used by U.S. and Chinese politicians based on the word list generated for this study. Also, further proving the usability of big data, it is possible to derive more analyses from the same data set in this paper, leaving a wide variety of options by which to use big data. Thus, this approach exemplifies the host of new possibilities to research social sciences.

By the same token, the results of this paper and the conclusions drawn from it add to the importance of political rhetoric when analyzing growing world powers. The works of historians like Paul Kennedy (2017) have been instrumental to global power research. The data of this study proves that more material can be used in the study of global powers than things like GDP figures or landmark foreign policy actions. Speeches and words themselves are significant sources of information and can be used for far more than just big data analysis. Past research using linguistic analysis have made significant discoveries like the study on speeches of Taiwanese presidents without the use of big data, showing the versatility and effectiveness of using political rhetoric for studies (Lu and Ahrens, 2008). This paper provides one method by which to use political rhetoric in drawing conclusions on world powers.

## Conclusion

This paper demonstrates the importance of national identity to leaders of great powers because of national identity words being used in consistently high amounts for both U.S. and

Chinese leaders. From this finding, new research opportunities into the role of national identity in global powers have been made and should be utilized in future studies. Also, many similarities between the political interests of U.S. and China can be seen in the data, making the case that China is perhaps on the same road that U.S. was when it began its development into a world power. While it may not provide a conclusive answer on whether or not China will become a superpower, it does indicate that China's current political focuses are much like those of the leaders of early-modern U.S.

This study also exemplifies the importance of linguistic analysis and how the interactions between political rhetoric and the progress of a country can be investigated. Big data is an effective means to do so and much research can be done under the same framework as this paper so as to encourage a growth in knowledge on the nature of global powers.

In conclusion, with the knowledge that the development of China can be seen in U.S.'s past as well, analyzing the current state of China and how it may develop in the future can be facilitated by recalling how U.S. became a world power a century ago.

### References

- Bureau of the Census (1975), "Historical statistics of the United States: colonial times to 1970 (Part 1)", available at: [https://fraser.stlouisfed.org/files/docs/publications/histstatus/hstat1970\\_cen\\_1975\\_v1.pdf](https://fraser.stlouisfed.org/files/docs/publications/histstatus/hstat1970_cen_1975_v1.pdf) (accessed 14 July 2019).
- Eshbaugh-Soha, M. (2008), "The impact of presidential speeches on the bureaucracy", *Social Science Quarterly*, Vol. 89 No. 1, pp. 116-132.
- Irwin, D.A. (2006), "Historical aspects of U.S. trade policy", *National Bureau of Economic Research*, available at: <https://www.nber.org/reporter/summer06/irwin.html> (accessed 15 July 2019).
- Jowett, A.J. (1984), "The growth of China's population, 1949-1982", *The Geographical Journal*, Vol. 150 No. 2, pp. 155-170.
- Kennedy, P.M. (2017), *The Rise and Fall of the Great Powers: Economic Change and Military Conflict from 1500 to 2000*, Vintage Books, New York, NY.
- Lin, J.Y. and Shen, Z. (2018), "Reform and development strategy", in Garnaut, R., Song, L. and Fang, C. (Eds.), *China's 40 Years of Reform and Development: 1978-2018*, ANU Press, Canberra, pp. 117-134.
- Lu, L. and Ahrens, K. (2008), "Ideological influence on building metaphors in Taiwanese presidential speeches", *Discourse and Society*, Vol. 19 No. 3, pp. 383-408.
- Nau, H. (2001), "Why 'The Rise and Fall of the Great Powers' was wrong", *Review of International Studies*, Vol. 27 No. 4, pp. 579-592.
- Nye, J.S. (2009), "Get smart: combining hard and soft power", *Foreign Affairs*, Vol. 88 No. 4, pp. 160-163.
- Osborne, L.L. (1980), "Rhetorical patterns in President Kennedy's major speeches: a case study", *Presidential Studies Quarterly*, Vol. 10 No. 3, pp. 332-335.
- Perry, E.J. (2015), "The populist dream of Chinese democracy", *The Journal of Asian Studies*, Vol. 74 No. 4, pp. 903-915.
- Rosenberg, E.S. (1982), *Spreading the American Dream: American Economic and Cultural Expansion, 1890-1945*, Hill and Wang, New York, NY.
- Santacreu, A.M. and Zhu, H. (2018), "What does China's rise in patents mean? A look at quality vs. quantity", *Federal Reserve Bank of St. Louis, Economic Synopses*, available at: <https://research.stlouisfed.org/publications/economic-synopses/2018/05/04/what-does-chinas-rise-in-patents-mean-a-look-at-quality-vs-quantity/>.

- 
- Seider, M. (1974), "American big business ideology: a content analysis of executive speeches", *American Sociological Review*, Vol. 39 No. 6, pp. 802-815.
- The World Bank (2019), "Manufacturing, value added (% of GDP)", available at: <https://data.worldbank.org/indicator/NV.IND.MANF.ZS?locations=CN> (accessed 15 July 2019).
- Ungerer, C., Storey, I. and Bateman, S. (2010), *Making Mischief: The Return of the South China Sea Dispute*, Australian Strategic Policy Institute, Canberra.
- Yang, M. (2011), "President Nixon's speeches and toasts during his 1972 trip to China: a study in diplomatic rhetoric", *Rhetoric and Public Affairs*, Vol. 14 No. 1, pp. 1-44.

(The Appendix follows overleaf)

Appendix A  
Raw data for 63 most common terms

US 1905-1945

Word	Amount #
state(s)	5190
nation	3964
great(er)	3852
government	2980
people	2005
united	1685
world	1497
country	1483
congress	1424
national	1026
business	1018
work	1012
peace	950
power	890
free	877
nations	772
service	765
purpose(s)	749
force(s)	693
system	675
action	669
life	608
foreign	574
policy	569
labor	558
order	546
america	543
freedom	506
economic	492
control	473
interest	472
duty	466
presidential	440
international	427
rights	423
future	421
believe	413
land	409
justice	400
progress	387
hope	380
further	371
industry	369
political	366
citizens	345
spirit	326
individual	321
common	316
self	295
become	286
help	283
opportunity	281
house	269
problems	264
security	254
continue	251
responsibility	251
human	246
support	245
liberty	240
history	238
wish	227
reason	209
confidence	192
<b>Total</b>	<b>51,243</b>

US 1977-2017

Word	Amount #
america	6684
work	4430
people	4200
nation	3944
country	3835
state(s)	3772
life	3322
freedom	2544
economy	2452
world	2424
presidential	2394
united	1949
government	1783
strength	1696
americans	1426
effort(s)	1268
peace	1223
congress	1179
help	1164
security	1053
believe	975
future	969
children	961
great	910
economic	770
nations	769
support	668
free	654
hope	651
families	642
history	627
national	625
system	619
military	615
change	598
human	590
citizens	568
power	562
continue	545
women	539
forces	534
rights	529
opportunity	494
policy	455
countries	429
build	420
progress	417
force	413
leaders	411
problems	402
forward	396
responsibility	396
house	380
action	356
political	344
protect	338
face	334
fellow	328
provide	324
lead	322
common	310
friends	285
share	285
order	241
<b>Total</b>	<b>75,760</b>

China 1977-2017

Word	Amount #
china	25912
develop	17562
party	14635
development	14116
people	13292
economy	12514
promote	10573
strength	10532
comrade(s)	8426
build	7992
central	7899
work	7591
chinese	7214
economic	7058
national	5779
committee	5763
country	5269
political	4968
world	4961
system	4806
countries	4519
reform	4409
achieve	3917
social	3752
construction	3692
socialist	3615
great	3569
international	3400
situation	3166
state	3029
government	2925
improve	2611
leadership	2333
community	2172
society	2078
continue	1982
spirit	1975
problems	1953
characteristics	1881
general	1855
overall	1847
issues	1838
support	1681
life	1668
study	1570
chairman	1551
forward	1549
stability	1508
policy	1478
further	1470
interests	1453
beijing	1428
progress	1424
increase	1399
actively	1321
order	1314
become	1287
strive	1249
focus	1219
attention	1215
efforts	1201
future	1193
create	1185
achievement	1131
<b>Total</b>	<b>289,685</b>



---

**About the authors**

Patrick Carter received his PhD in History from the University of Cambridge. He subsequently taught at both McMaster University (Canada) and the University of St Andrews (Scotland). He is a member of the Royal Historical Society (United Kingdom). Author of several respected Canadian secondary school texts including *American History* and *Canadian and World Politics*, his published historical scholarship includes many journal articles and chapters, as well as contributions to the *New Dictionary of National Biography* and other works. For the past twenty years he has taught history and politics in the Social Sciences Department at Appleby College in Canada, where he is currently Director of Moral and Ethical Education.

Jeffrie Wang is undergraduate student majoring in economics at the University of California, Berkeley. He is a recipient of the International Bronze Duke of Edinburgh Award and is a National AP Scholar. His research interests are international relations and history. Jeffrie Wang is the corresponding author and can be contacted at: [jwangx235@gmail.com](mailto:jwangx235@gmail.com)

Davis Chau is the founder and Director of Innovative Model and Industry Development Institute, a consulting firm specializing in China's regional, social and industrial development. He is a provincial "Leading Talent" of Shandong and advisor to several provincial and municipal cities in China. He is an acclaimed speaker and lecturer in Tsinghua University, Peking University, Shanghai Jiaotong University, state-owned teaching institutes and corporations. His research interests are big data, fintech, investment banking, business and industry innovations. He earned his BA in Anthropology and MBA at The Chinese University of Hong Kong, with finance specialization at New York University.

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgroupublishing.com/licensing/reprints.htm](http://www.emeraldgroupublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)