

The evolution of romantic love in Chinese fiction in the very long run (618 - 2022): A quantitative approach^{*}

Ying Zhong, Valentin Thouzeau and Nicolas Baumard

Institut Jean Nicod, Département d'études cognitives, École normale supérieure, Université PSL, EHESS, CNRS, 75005 Paris, France

Abstract

Literary scholars have long observed the fluctuating popularity of romantic love in Chinese fiction, and the existence of periods when romantic love was particularly central to Chinese fiction: The High Tang Dynasty (short stories), the Yuan Dynasty (plays), the High Qing Dynasty (scholar-beauty novels), and the most recent period (modern series, web novels). However, these observations have remained qualitative, and the overall cultural history of China is still unknown. Is there a pattern in the history of love in China? Can we describe it quantitatively? And can we explain it? Here, we present a new database of summary plots of Chinese fiction (N = 3496) from the Tang Dynasty (618 AD) to the modern era (2022). Using the method of linguistic inquiry, we first confirm that the evolutionary pattern of romantic love in the long run corresponds to qualitative observations reported by literary scholars and cultural historians, with an increase during the ancient and the contemporary phase (post-1978). We then test whether these patterns can be explained by a change in people's preferences in response to increasing economic development. Consistent with previous work, we show that the rise of romantic love correlates with the ups and downs of economic development in Chinese history.

Keywords

Cultural evolution, Chinese Fiction, Economic development, Romantic love, Bag of words

1. Introduction

Chinese culture is rich in romantic content, ranging from ancient myths and classical literary works to modern popular dramas. However, it is well known that the importance of romantic love to people has varied throughout history. In particular, while the elite literary tradition prior to the Tang was largely silent on the topic of romantic love, there was an explosion of interest in and writing about love after the period following the An Lushan Rebellion (after 763 AD) [13]. The importance of love continued into the Ming and Qing Dynasties [26][21][3]. For instance, companionate love occasionally became a major quality for a happy marriage during the Qing dynasty [21]. In the early twentieth century, the May Fourth Movement brought a

CHR 2023: Computational Humanities Research Conference, December 6 – 8, 2023, Paris, France


^{*}All code and data can be downloaded in https://github.com/silverzoe/CHR2023_193

✉ yingzhong196@gmail.com (Y. Zhong); valentin.thouzeau@gmail.com (V. Thouzeau); nbaumard@gmail.com (N. Baumard)

🌐 <https://valentinthouzeau.wordpress.com/> (V. Thouzeau); <https://nicolasbaumards.org/> (N. Baumard)

🆔 0009-0003-7051-9609 (Y. Zhong); 0000-0002-2096-2675 (V. Thouzeau); 0000-0002-1439-9150 (N. Baumard)

© 2023 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

 CEUR Workshop Proceedings (CEUR-WS.org)

revolutionary view of "love" (*aiqing*) as a symbol of freedom against Confucian family institution and the practice of love letters were especially popular among young people back then [32]. This progressive outlook didn't last long, however, and after the late 1920s gave way to a more conservative ethos. From the 1950s to the 1970s, love took on a socialist style, emphasizing expressions of love based on class [18]. After the reform and opening up period, love experienced a revival with the emergence of more diversified media and content, including romantic dramas or web novels.

Despite the important qualitative data that scholars have provided, many questions remain unanswered. Are qualitative observations based on close reading supported by studies based on distant reading, using much larger samples? Is there an overall increase in romantic love in the long run? Why does love emerge as a significant popular topic in fiction while remaining silent in another period? Can we quantify the historical development of love in fiction? How can we understand the evolution of love in Chinese fiction? In this paper, we create a new database of Chinese fictions from 618 to 2022 AD. Using text mining and natural language processing tools, we quantify the cultural history of love in Chinese fiction. Finally, we aim to explain this history. Recent works in historical psychology[4] have shown a systematic relationship between economic development and the importance of romantic love in fictional literature. However, this relationship is mostly based on Western literature, and included only a very small sample of Chinese literature. Here, we used our much larger sample to test this relationship further. We thus use a general linear model to examine the relationship between living standards (take GDP per capita as a proxy) and the evolution of love throughout history. Our data consistently support the idea that an increase in love in Chinese fiction is associated with economic development over the long term.

2. Data

Web platforms such as Wikipedia or Baidu Baike are powerful sources for studying the content of fiction [4][22][19][10]. The data provided by online encyclopedias include year of publication, content keywords, genres, or character information, which makes it easier to track the presence of specific content through word searches. It also makes it possible to avoid shifts in language usage over time and facilitates bag-of-words approaches. For example, a modern Chinese reader will annotate a love story with the Chinese characters for love (爱情), which is sufficient to track the presence of romantic love.

We collected descriptions and years of publication of fictions from several sources, including literary scholars, Wikipedia and Baidu Baike. These multiple sources allow us to control the influence of platforms and trace further back into ancient times.

2.1. Sampling

Acquiring comprehensive descriptions of Chinese fiction is a significant challenge, primarily due to the limited use of Wikipedia, the world's most comprehensive open access platform, in mainland China. In contrast, Baidu Baike, a widely used resource in China, is not open access and lacks a structured data framework. To overcome these limitations, a systematic data collection methodology is of great importance.

We addressed this problem with a three-pronged approach. First, we targeted contemporary novels and video-type fictions that can be automatically queried from Wikipedia. This step was conducted using Wikidata query services¹, where we automatically obtained all Wikipedia URL links and the earliest publication year of the works. Then, based on the URL links, we automatically extracted descriptions and summaries from the Wikipedia pages.

Second, we complemented the dataset with Baidu Baike descriptions. It is not possible to automatically sample Baidu Baike data with a query service. For written work, we relied on manual collection based on multiple indexes of names, including (1) literary bibliography and (2) popular online book website Douban. For the ancient fiction, we used the work of Lu Xun [14] and works of JiangSu Provincial Academy of Social Sciences [30]. For the Republic of China 1917 to 1949, we used the work of Liqun Qian [20]. For novels between 1949 and 2000, we used work of Hong Zicheng [12]. To sample popular novels, we used the Douban 250 top list². The ancient samples obtained by this method on Baidu Baike basically cover the ancient samples that exist in Wikipedia.

Third, we collected a sub-sample, which are adaptations of all ancient works mentioned by Lu Xun [14]. For this part, we manually collected all descriptions of adaptations (novels, series, films or video games) in Wikipedia and Baidu Baike, as well as their adaptation relationship.

2.2. Description

We ended up with a total number of 3496 fictions (580 adaptations within). Some works have multiple sources of descriptions (N = 4193 descriptions). In the whole data set, 1895 fictions are textual, 1601 fictions are video. Figure 1 shows the distribution of description by sources and eras.

3. Method: Measuring romantic love in Chinese fiction

So far, existing tools for natural language processing have some limitations. WordNet [8], dealing with the meaning of words by making use of taxonomic resources, could not find similar words, and has limited usage in Chinese language. LIWC22 [35], mainly used in measuring culture, sentiment, meta-cognition studies, is less applicable in the domain of romantic love. Synonyms are great as a resource but miss nuances to compute accurate word similarity.

Here, we developed a systematic methodology to conduct linguistic measure of love in texts. As inspired by works in English [4][24][23][16], we started from seed words, then expanded the dictionary with *Synonyms* tools and bag-of-words.

3.1. Building Inquiry Dictionaries for romantic love

We conducted a literature review on terminology around love in both psychology and literary literature to select seed words. First we choose two narrow connotation of words describing

¹<https://query.wikidata.org/>

²250 Top popular books in Dou Ban: <https://book.douban.com/top250>

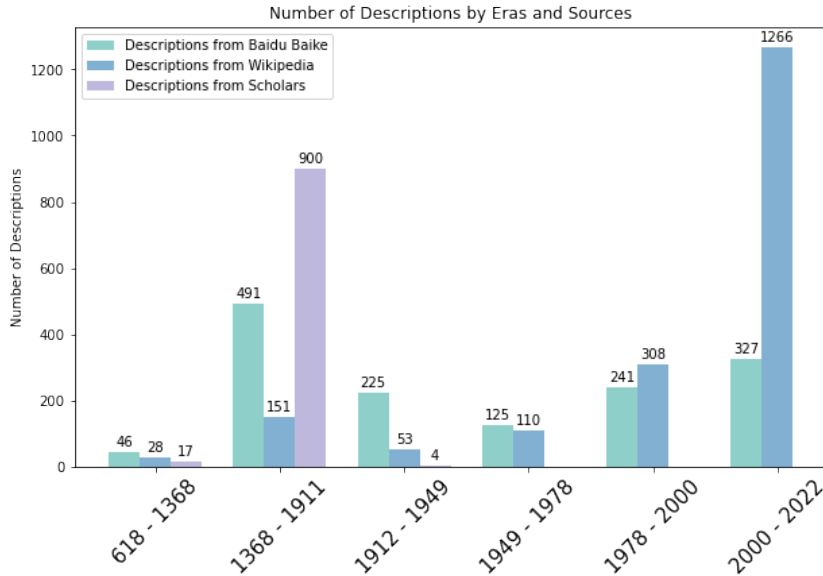


Figure 1: Number of Descriptions by Eras and Sources

”romantic love” [26][18][33]: 爱情 (love), 恋爱 (fall in love). We also consider the wider connotation of words around love components in behavioral sense: 亲密 (intimacy), 激情 (passion), 承诺 (commitment).

The workflow of generating dictionaries is described in Figure 2. In a first method, based on our seed words, we created a dictionary using WantWords to get the most relevant synonyms³ (see [28]). In a second method, we used Python packages *Synonyms* to automatically generate lists of words for each seed word. We then train a vector space using Baidu Baike corpus (124MB), based on the model of Word2vec [29][11]. Taking the lists of words generated by *Synonyms* as input, we get output of the most similar words (based on Cosine Similarity) for each input word. In the end, we convert the two dictionaries into one large and non-repeated dictionary⁴. We manually check the meaning of each words and remove the words that are not related to romantic love.

To test the validity of the dictionaries, we used descriptions of films from different genres. We verified that romantic comedy films contain significantly more words occurrences of our dictionary than historical films. The mean ratio of love words in romantic comedy films is 0.0007, significantly larger than that in historical films 0.0002 ($t = 4.45$, $p < 0.001^{***}$).

3.2. Data preprocessing

We did the following steps for preprocessing texts and compute the variables used in the model analysis:

³<https://github.com/thunlp/WantWords>

⁴Final dictionary could be checked in: https://github.com/silverzoe/CHR2023_193

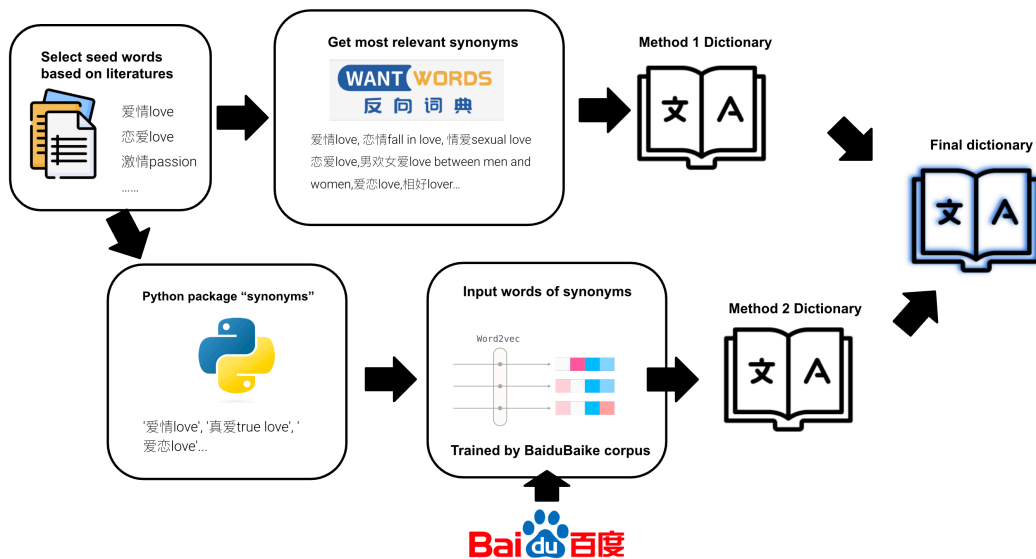


Figure 2: Workflow of generating dictionaries for romantic love in Chinese

- Tokenize description from sentences to words, using 'Jieba' Python package⁵
- Remove punctuation, using 'zhon.hanzi' Python package⁶
- Use string length function to calculate the length of each descriptions
- Calculate the occurrences of love words according to the final dictionary in each description
- Merge GDP data⁷ [5] with our Chinese fiction data by Year
- Compute type of fiction: Textual or Video
- Compute source of description: Wikipedia, Baidu Baikē or Scholar
- Compute the differences of occurrences of love words between adaptations and their original works

3.3. Correlation between love in full texts and descriptions

Another way to evaluate the validity of our methodology is to check the correlation with the presence of love in the description and in the full texts. We acquired 802 full texts of Chinese fiction from open source websites, including Wikisource, Handian, Open literature, and Gushiwenwang. There are 334 fictions with both descriptions and full texts in our database.

⁵<https://github.com/fxsjy/jieba>

⁶<https://pypi.org/project/zhon/>

⁷The Maddison 2020 project data has a lot of missing values in its historical data, and we don't want to lose ancient samples, so we fill in the gaps with interval data. We estimate that this does not introduce much bias, as there is relatively little variation within the ancient GDP intervals. For data from 2018 onwards, we use World Bank data, which are converted to the same units as the rest of the data.

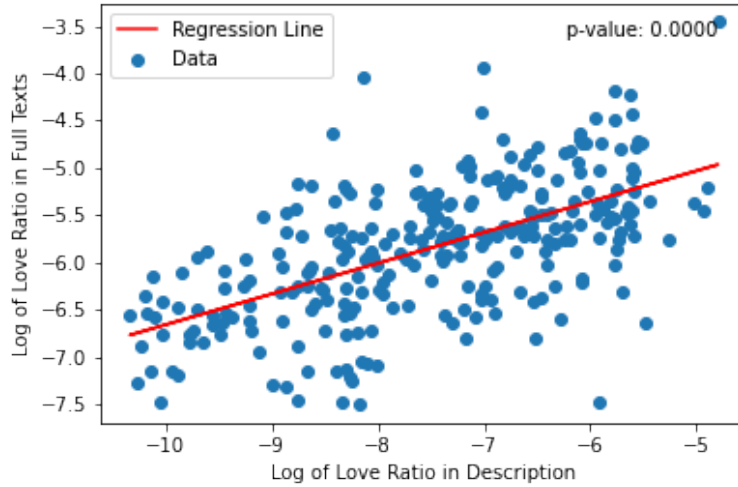


Figure 3: Correlation between love words ratio in descriptions and full texts (logged version)

We used the dictionaries for description as seed words, then found all the neighbour words of the seed words in our 802 full texts. We then used these full texts to train a word2vec vector space and expand the neighbour words to a larger dictionary.

Then, we followed the same preprocessing method presented above and computed the ratio of love (number of words related to love divided by text length) for full texts. We find a positive correlation between the love ratio in full texts and descriptions as shown in Figure 3 ($r = 0.53$, $p < 0.001^{***}$).

4. Results

4.1. Love in the very long run

In large-scale historical analyses, the considerable disparity in the number of available data between the distant past and the present, as well as China’s complex historical development, pose a challenge to linear modelling.

In Figure 4, a plot of the Generalized Additive Model (GAM) demonstrates the patterns of evolution of love in Chinese fiction over time (the black line). The result of the model show the significant non-linear relationship between year and frequency of love words in the long run history, controlling with length, type and source of descriptions. ($edf = 8.118$, $chi^2 = 168.5$, $p < 0.0001^{***}$).

Figure 4 shows the Love was a relatively high in Tang (800s), followed by a gradual decline and then a continuous increase until the mid of eighteenth century. Then, love seems to decrease until revive at the the Republic of China, when love was especially main topic in Mandarin ducks and butterflies school and martial art stories [26][18]. And lastly, love becomes more and more important with time since 1970s followed by a drop after 2010s.

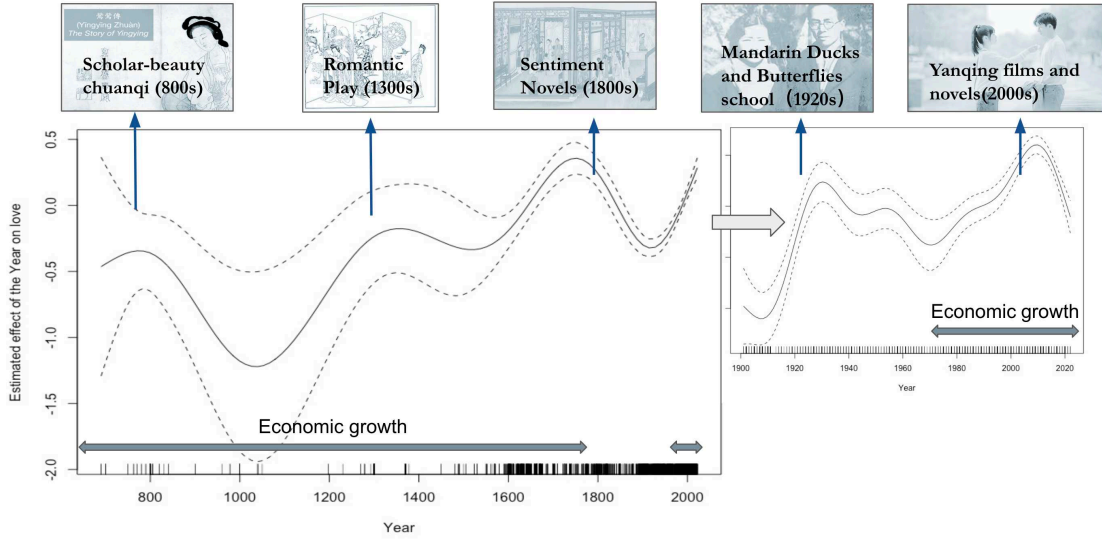


Figure 4: a) The evolution of love in Chinese fiction from 618 to 2022, b) The evolution of love in Chinese fiction from 1900 to 2022. The black curve shows the relationship between the predictor love and the response variable year. The shaded regions around the curve represent confidence intervals for the GAM model.

4.2. Love increased during ancient and contemporary China

To further understand the evolution of love in history, we used a General Linear Model (GLM) with Poisson error to examine the relationship between the frequency of love words in descriptions and year during periods of economic growth, while controlling for factors such as fiction type, description source, and description length.

We fitted two models separately for (1) from 618 AD Tang dynasty until 1799 AD High Qing and (2) recent 44 years since 1978 Chinese economic reform. Because the remarkable economic growth experienced during these two historical periods and valid observation of love made by historians for these two periods makes them ideal for understanding the evolution of love by time and evaluating our methodology.

Model 1 focuses on the ancient period, spanning from 618 AD to 1799 AD, characterized by minimal historical changes in the political and social structure with some evidence of continuous increase of living standards [37][1] and expression of love [3]. Model 2 focuses on the last 44 years, a period marked by significant changes in Chinese history. Notably, GDP per capita was highly positively correlated with the year for recent 44 years ($r = 0.97$, $p < 0.0001^{***}$).

$$\text{love} = \beta_0 + \beta_1 \cdot \text{Year} + \beta_2 \cdot \text{length} + \beta_3 \cdot \text{Source} + \epsilon_{\text{Pois}} \quad (1)$$

$$\text{love} = \beta_0 + \beta_1 \cdot \text{Year} + \beta_2 \cdot \text{length} + \beta_3 \cdot \text{Source} + \beta_4 \cdot \text{Type} + \epsilon_{\text{Pois}} \quad (2)$$

The result shows that Love is positively correlated with time from 618 to 1799 AD (N = 742 descriptions, $\beta_1 = 7.318 \times 10^{-4}$, $p < 0.001^{***}$, AIC = 2378, Model 1). We also confirm that over

the recent last 44 years, love has continued to increase significantly ($N = 2,058$ descriptions, $\beta_1 = 7.651 \times 10^{-3}$, $p < 0.001^{***}$, $AIC = 7370$, Model 2). Overall, the results indicate that during economically prosperous times, love becomes more and more important in Chinese fiction.

4.3. Economic development is positively associated with love

To model the relationship of economic development and prevalence of love in Chinese fiction in the long run, we model the relationship of Gdp per capita and love words frequencies, while controlling for time, length, source, and type of fictions, in Model 3.

$$\text{love} = \beta_0 + \beta_1 \cdot \text{Gdp} + \beta_2 \cdot \text{length} + \beta_3 \cdot \text{Source} + \beta_4 \cdot \text{Type} + \beta_5 \cdot \text{Year} + \epsilon_{\text{Pois}} \quad (3)$$

4,193 descriptions of fiction from the beginning of the Tang dynasty in 618 to the year 2022 were examined. Our findings show that, in line with previous works[4], increases in GDP per capita can explain the rise of frequency of love in fiction ($\beta_1 = 3.366 \times 10^{-5}$, $p < 0.001^{***}$, Model 3⁸), while we do not find significant correlation with years and love ($\beta_5 = 1.231 \times 10^{-4}$, $p = 0.17$, Model 3), which means when GDP per capita rises, people express more love in fiction.

4.4. Adaptations contain more Love than original works

Finally, we tested the robustness of this result using a different sample. In Chinese fictions, many works are adaptations of more ancient works. In our data, there are around 580 fictions adapted from other works. Figure 5 shows the patterns of adaptations. The largest adaptation network are around *Journey to the west* ($N = 166$), *Water Margin* ($N = 51$), *Dream of the Red Chamber* ($N = 41$), *Monkey king: hero is back* ($N = 36$). A descriptive Word Cloud analysis (Figure 6) reveals prevalent themes in popular adaptations (where there are at least 10 adaptations). Notably, "love" emerged as a significant topic, with the presence of words "couple," "love," "marriage," "happy end," and "yuan yang" (mandarin ducks, a metaphor of couple).

We hypothesized that if economic development has an impact on people's preferences and does increase the interest for romantic love in fiction, then when a work from a period of lower development is adapted in a period of higher development, the importance of love should be more important in the new version than in the older version. In line with this idea, the heroic warrior Monkey King of Journey to the West was transformed into a romantic love character in Hong Kong and Japanese films in the 20th century, before becoming successful in mainland China in the early 21st century. A similar phenomenon has recently occurred when the stories of immortal heroes are transformed into romantic love stories among new generations in China [34]. These seem to show that people living in more developed times or regions are more interested in making love adaptations, even if the original work is "love-free".

We use 580 adaptations to test this idea. Overall, we find that adaptations contain a proportion of 0.0001 more love words compared to the original works ($t = 3.5202$, $p < 0.001^{***}$). In order to explain this phenomenon, we calculated the difference in occurrences between the adaptation and their original works. Then we model the relationship between difference in love frequency and difference in GDP per capita, finding a positive correlation ($\beta = 4.731 \times 10^{-1}$,

⁸Here we did not provide a robustness score because we are not predicting the precise effect of GDP on love in fiction. Rather, it matters in if it fit into our prediction and how estimate is in line with the direction we predicted

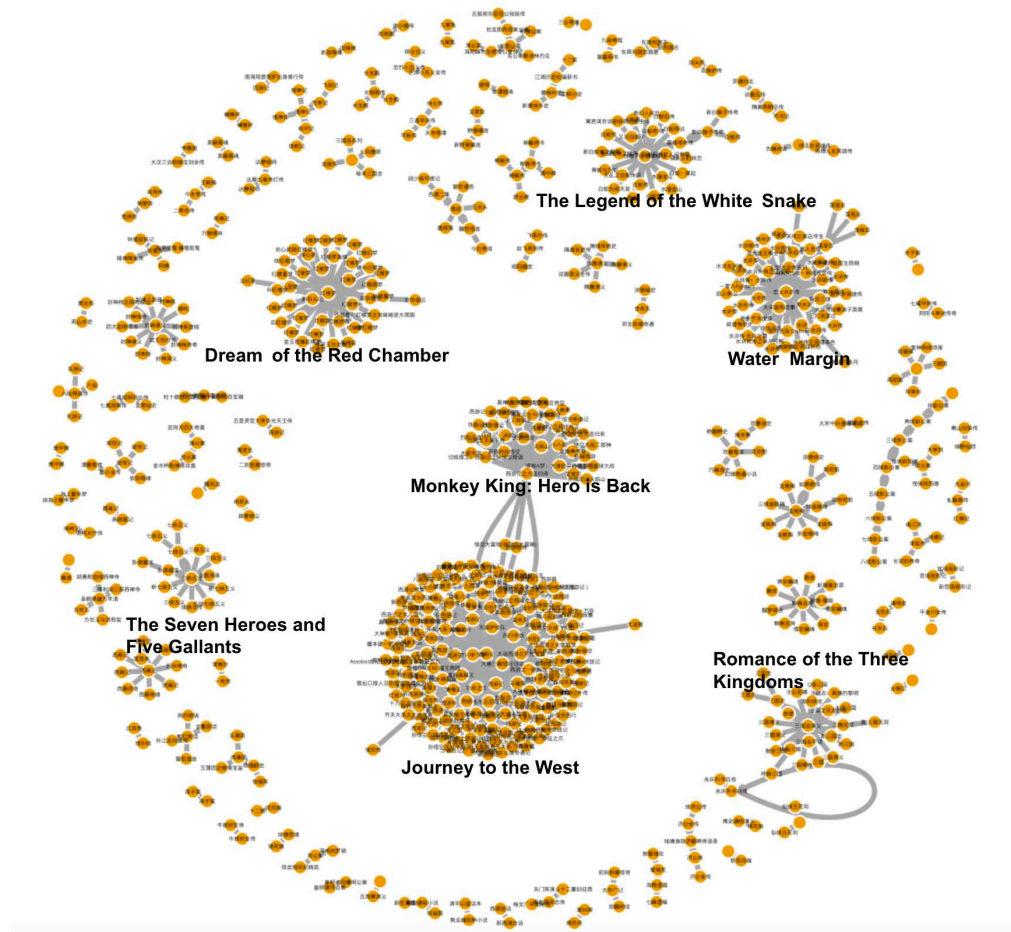


Figure 5: Network of adaptations and original fiction (N = 580 adaptations)

$p < 0.001^{***}$). However, this effect is not significant when we control for changes in type of fiction ($\beta = 1.531 \times 10^{-1}$, $p = 0.20$). In particular, changing from a textual fiction to a video fiction significantly increases the love words frequencies ($\beta = 9.14 \times 10^{-1}$, $p < 0.001^{***}$).

5. Conclusion

This short paper is the first to quantify romantic love in Chinese fiction in the long run. We built a new database and used a systematic method to quantify love in Chinese fiction. We show that there are numerous variations in the prevalence of love in fiction over time. Specifically, love increased from the Tang to the High Qing dynasties, as well as contemporary China. These patterns of development in ancient fiction are remarkably consistent with the historical and empirical records of the importance of love in Chinese society. In the ninth century Tang, romantic feelings were valued as a defining quality of an individual [25]. In the Ming dynasty,



Figure 6: Word Cloud with descriptions that have at least ten adaptations

love was a central value for marriage [6]. In the Qing Dynasty, the rise of love was particularly evident in women’s love poetry [31]. Moving forward, the popularity of love in fiction from the 1920s to the post-2000s is closely related to the fact that love is an aspect of modernization packages, with people valuing love in reality and placing a high value on commitment to an intimate other [7][17] and advocating free choice of love [36] in early 20th century. Nowadays, love has permeated the internet, becoming a prominent means of expression for young people [9].

Then, how can we explain the emergence of love in Chinese fiction? How can we explain the increasing importance of love in Chinese fiction in both ancient and modern times. We find that love in fiction is positively related to GDP per capita in China. This is consistent with behavioral hypotheses that changes in the environment, especially in living standards, lead to the rise of romantic pair bonding [4][15][2][27]. This psychological trait should be favoured when long-term orientation is manifested, as individuals benefit from the formation of long-term pair bonds and higher levels of parental care.

We also discovered a unique pattern of adaptation in which people add more love than in the original works. The direction is consistent with economic development, but mainly because love is more represented in a video-type adaptation than in text-format fiction.

In future work, we aim to replicate our analysis with full texts to open up more possibilities on specific love contexts, for example, passionate love versus long-term love. We should also consider our sample selection biases, for example, the explosion in the publication of web-novels, is so vast that the samples recorded on Baidu and Wikipedia only represent a small fraction of the most popular ones. Further work is needed to look more closely at the cultural evolution of love in specific periods where social constraints may play a role (such as why Republic of China went through a modernization revolution despite the economic development without leapfrogging).

References

- [1] J. Baten, D. Ma, S. Morgan, and Q. Wang. “Evolution of living standards and human capital in China in the 18–20th centuries: Evidences from real wages, age-heaping, and anthropometrics”. In: *Explorations in Economic History* 47.3 (2010), pp. 347–359.
- [2] N. Baumard. “Psychological origins of the industrial revolution”. In: *Behavioral and Brain Sciences* 42 (2019), e189.
- [3] N. Baumard. “The Ancient Literary Fictions Values Survey”. In: (2021).
- [4] N. Baumard, E. Huillery, A. Hyafil, and L. Safra. “The cultural evolution of love in literary history”. In: *Nature Human Behaviour* 6.4 (2022), pp. 506–522.
- [5] J. Bolt and J. L. Van Zanden. “Maddison style estimates of the evolution of the world economy. A new 2020 update”. In: *Maddison-Project Working Paper WP-15, University of Groningen, Groningen, The Netherlands* (2020).
- [6] K. Carlitz. “10. Lovers, Talkers, Monsters, and Good Women”. In: *Beyond Exemplar Tales: Women’s Biography in Chinese History* (2011), p. 175.
- [7] H. Donner and G. Santos. “Love, marriage, and intimate citizenship in contemporary China and India: an introduction”. In: *Modern Asian Studies* 50.4 (2016), pp. 1123–1146.
- [8] C. Fellbaum. “WordNet”. In: *Theory and applications of ontology: computer applications*. Springer, 2010, pp. 231–243.
- [9] J. Feng. *Romancing the internet: Producing and consuming Chinese web romance*. Vol. 5. Brill, 2013.
- [10] M. Figlerowicz and L. Mertehikian. “An Ever-Expanding World Literary Genre: Defining Magic Realism on Wikipedia”. In: *Journal of Cultural Analytics* 8.2 (2023).
- [11] Y. Goldberg and O. Levy. “word2vec Explained: deriving Mikolov et al.’s negative-sampling word-embedding method”. In: *arXiv preprint arXiv:1402.3722* (2014).
- [12] Z. Hong. *A history of contemporary Chinese literature*. Vol. 1. Brill, 2007.
- [13] D. Hsieh. *Love and women in early Chinese fiction*. The Chinese University of Hong Kong Press, 2009.
- [14] L. Hsun. *A brief history of Chinese fiction*. The Minerva Group, Inc., 2000.
- [15] R. F. Inglehart. *Cultural evolution: People’s motivations are changing, and reshaping the world*. 2020.
- [16] J. C. Jackson, M. Gelfand, S. De, and A. Fox. “The loosening of American culture over 200 years is associated with a creativity–order trade-off”. In: *Nature human behaviour* 3.3 (2019), pp. 244–250.
- [17] N. Karjanto, Z. Zhang, and Y. R. Park. “Young love in contemporary China: Attitudes and behaviors in romantic relationships”. In: *Qeios* (2023).
- [18] H. Lee. *Revolution of the heart: A genealogy of love in China, 1900-1950*. Stanford University Press, 2006.

- [19] A. S. Lippolis. “Italian Nostalgia: National and Global Identities of the Italian Novel”. In: *Journal of Cultural Analytics* 8.2 (2023).
- [20] Q. Liqun, W. Rumin, and W. Fuhui. *Thirty years in Modern Chinese Literature*. 1998.
- [21] W. Lu. *Arranged companions: Marriage and intimacy in Qing China*. University of Washington Press, 2021.
- [22] D. Marengo, C. Montag, et al. “Digital phenotyping of big five personality via facebook data mining: a meta-analysis”. In: *Digital Psychology* 1.1 (2020), pp. 52–64.
- [23] M. d. J. D. Martins and N. Baumard. “The rise of prosociality in fiction preceded democratic revolutions in Early Modern Europe”. In: *Proceedings of the National Academy of Sciences* 117.46 (2020), pp. 28684–28691.
- [24] O. Morin and A. Acerbi. “Birth of the cool: a two-centuries decline in emotional expression in Anglophone fiction”. In: *Cognition and emotion* 31.8 (2017), pp. 1663–1675.
- [25] S. Owen. *The End of the Chinese ‘Middle Ages’: Essays in Mid-Tang Literary Culture*. Stanford University Press, 1996.
- [26] L. Pan. *When true love came to China*. Hong Kong University Press, 2015.
- [27] G. V. Pepper and D. Nettle. “The behavioural constellation of deprivation: Causes and consequences”. In: *Behavioral and Brain Sciences* 40 (2017), e314.
- [28] F. Qi, L. Zhang, Y. Yang, Z. Liu, and M. Sun. “Wantwords: An open-source online reverse dictionary system”. In: *Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations*. 2020, pp. 175–181.
- [29] X. Rong. “word2vec parameter learning explained”. In: *arXiv preprint arXiv:1411.2738* (2014).
- [30] J. P. A. of Social Sciences. *A bibliography of vernacular novels Zhongguo tongsu xiaoshuo-zongmu tiyao*. China Federation of Literary and Art Circles Publishing Corporation, 1990.
- [31] P. N. Stearns and M. Knapp. “Men and romantic love: Pinpointing a 20th-century change”. In: *Journal of Social History* (1993), pp. 769–795.
- [32] H. Su. “Infrastructures of romantic love: Chinese love letters in the early 1900s”. In: *Interactions: Studies in Communication & Culture* 7.3 (2016), pp. 265–280.
- [33] S. R. Sumter, P. M. Valkenburg, and J. Peter. “Perceptions of love across the lifespan: Differences in passion, intimacy, and commitment”. In: *International Journal of Behavioral Development* 37.5 (2013), pp. 417–427.
- [34] W. Sun and L. Yang. “Love stories in contemporary China: Cultural production in the new millennium”. In: *Love Stories in China The Politics of Intimacy in the Twenty-First Century* (2020).
- [35] Y. R. Tausczik and J. W. Pennebaker. “The psychological meaning of words: LIWC and computerized text analysis methods”. In: *Journal of language and social psychology* 29.1 (2010), pp. 24–54.
- [36] Y. Yan. “Courtship, love and premarital sex in a north China village”. In: *The China Journal* 48 (2002), pp. 29–53.

- [37] X. Yi, B. van Leeuwen, and J. L. van Zanden. “Urbanization in China, ca. 1100–1900”. In: *Frontiers of Economics in China* 13.3 (2018), pp. 322–368.