The Evolutionary Pattern of Language in English Fiction Over the Last Two Centuries: Insights From Linguistic Concreteness and Imageability

SAGE Open
January-March 2022: 1–13
© The Author(s) 2022
DOI: 10.1177/21582440211069386
journals.sagepub.com/home/sgo

Kun Sun¹ and Rong Wang²

Abstract

Fiction has come to play an essential part in human culture and life in recent centuries. Because of its importance, the language or discourse of fiction has been widely studied. Discerning the evolutionary pattern of language in fiction is greatly helpful in understanding the changes in human culture and society. However, most previous studies on literary language that made use of quantitative or computational methods restricted themselves to a few authors or groups and only took into account a relative short span of time. Furthermore, most of the quantitative analysis therein was based on a primary and rather primitive algorithm that summarizes the frequencies of linguistic units in corpora. To overcome these limitations, this study uses semantic similarity as computed by semi-supervised methods to examine diachronic changes. In this study, the three large-scale corpora representing the English language and English fiction were used to investigate the evolutionary patterns concerning diachronic concreteness/imageability. The data from the two measures are well mapped and shown to support the argument that English fiction, as a special genre, exhibits an evolutionary tendency toward increasing concreteness/imageability. This indicates that modern fiction may have become increasingly important in human society, but easier to read and process in words than 19th century English fiction. This study proposed that learnability, genre difference, the changes of population size, and the other factors might have caused the systemic change. The current study is thus an important contribution to understanding the evolutionary trend of language in fiction as well as understanding the development of the English language as a whole. This study creates a novel quantitative methodology and applies this to the examination of diachronic changes of language in literary works.

Keywords

semantic similarity, concreteness, semi-supervision, diachronic changes, evolutionary pattern

Introduction

We can take a key step toward gaining a better understanding of human nature and the changes and developments in the world by detecting cultural evolutionary trends and patterns (Shennan, 2009). Previous quantitative investigations into how cultural phenomena have evolved drew on both computing technology and the widespread and increasing availability of digital material (Carr et al., 2017; Hill et al., 2014; Hughes et al., 2012; Sun et al., 2021). Such research has contributed to creating a quantitative science focused on cultural changes as well as to giving us a better understanding of trends in the evolution of human culture.

Many studies find that languages evolve over time (Labov et al., 2011). It is hard to over-estimate the importance of diachronic studies of evolving language and related social

change. The knowledge of language changes probably helps in gaining significant insights into the origins of present-day culture and society. And such knowledge is also helpful in understanding how human communication and their communicative tools have evolved (Sneffjella et al., 2019). And it is clear that discerning the evolutionary pattern of language in

¹University of Tübingen, Germany

²University of Stuttgart, Germany

Corresponding Authors:

Kun Sun, The Department of Linguistics, University of Tübingen, Wilhelmstraße 19, 72074, Germany.
Email: kun.sun@uni-tuebingen.de

Rong Wang, The Institute of Natural Language Processing, University of Stuttgart, Stuttgart, Germany.
Email: rong4ivy@163.com

fiction also yields insights into changes in human language, culture and society. Past studies of such changes in language have often taken a number of linguistic features and have measured and examined the changes these linguistic features have undergone. However, because of the availability of massive digitized texts and recent progress of computational technologies, today we can obtain high-quality corpora, and more importantly, effective algorithms from computational linguistics that can be adopted to measure linguistic changes. This has important practical and research implications. It means that we need no longer be confined to small corpora and to a single algorithm (such as word frequency) in investigating the evolutionary patterns in the language over history.

As we know, fiction has become an essential part of human culture and life in recent centuries (Daly, 2000; Davis, 1997). Reading fiction can provide many surprising benefits, such as better understanding of our world and allowing us to be more productive (Kidd & Castano, 2013; Zunshine, 2006). Because of the cultural and social significance of fiction, its language or discourse has been the subject of much academic research (Fludernik, 2003). Understanding the evolutionary patterns of language in fiction is essential to understanding the changes in human culture and society. The application of quantitative analysis to the language in fiction or literature has grown significantly in recent years. One example of this is *stylometry*, which uses linguistic style to have a view to assign authorship to anonymous or disputed documents by using quantitative or computational methods (Eder et al., 2016; Grieve, 2007; Kenny, 2013). Most studies concerning language in literature use quantitative or computational methods focused on the synchronic perspectives. Although there have been a very limited number of diachronic studies, the majority of these studies confined themselves to a few authors or groups over a short period of time (Mahlberg, 2013; Morin & Acerbi, 2017; Peng & Hengartner, 2002). However, the absence of diachronic studies that use quantitative analysis to explore language in fiction has hitherto prevented us from gaining insight into the developmental patterns of language in fiction. In addition, most quantitative analysis is based on a primary and rather primitive algorithm that summarizes the frequencies of words (or grammatical units) in fiction corpora. However, with recent developments in computational models and computational linguistics, we now have multiple effective and sophisticated algorithms at our disposal that can be employed in quantitative analysis.

A frequent complaint about 19th century English fiction (e.g., Victorian novels) is that it is hard to read (Biber & Conrad, 2019, p. 155). One of the reasons for this difficulty is that the vocabulary is somewhat archaic. On the other hand, with the development of society and science, the English language has been widely perceived as having become more varied and abstract as a whole. Our question is

whether the words used in English fiction have become more abstract or more concrete since the 19th century. A number of studies have already examined this topic. Biber and Finegan (1989) and Biber and Conrad (2019, p. 155) held that English fiction from the 18th century has been becoming less abstract, that is to say, more concrete in its style. In other words, the linguistic concreteness in English fiction has decreased since 18th century. These studies supported this conclusion with statistics concerning frequency from a number of English fictional works taken from different periods. However, a broader question can be raised here: is it possible to detect lexical developments in large-scale fiction corpora over centuries using the algorithm that is not limited to merely detecting the frequencies of words?

One of convenient and credible ways to discerning the evolutionary patterns in the language of literary works is to examine diachronic linguistic changes. Lexical concreteness has proved to be an efficient means of detecting the semantic changes of words (Hamilton, Leskovec et al., 2016; Hills & Adelman, 2015). Concreteness has an extensive literature supporting its relevance to memory, learning, and word recognition, and this dimension has been investigated extensively in historical studies recently. Concreteness in a corpus (or text or document) can be estimated by semi-supervised methods, which could be also appropriate for application to documents (texts, corpora) (Hamilton, Clark et al., 2016; Hamilton, Leskovec et al., 2016). To date, however, we have known a little about diachronic changes in abstractness in language. The challenge in documenting such changes rests in identifying a clear linguistic marker of concreteness and tracking this marker in relevant historical corpora.

Everyday language has various genres. The language in these texts covers too broad a range and it is influenced by numerous factors. To date, several corpora on English language texts have been created that contain a balance of different genres. Previously, research on historical records has concentrated on how language is used and then examined how such usage can shed light on broader changes in general usage, style, and grammar. However, by focusing on a single genre, we can make comparisons with the English language as a whole. The specific concern of the present study, however, is to investigate whether English fiction has evolved toward increasing abstraction or not. Although past work has found a pattern of rising concreteness in American English (Hills & Adelman, 2015; Sneffjella et al., 2019), we would like to know whether this pattern is followed in other Englishes (i.e., British English) as well.

In order to address this concern, we adopt a new methodology that allows us to quantitatively investigate how language in English fiction evolved over the last two centuries. The following methods have recently been extensively used to detect linguistic concreteness. These methods are based on semi-supervised methods and are used to estimate the degree

concreteness in a corpus (text or document). Linguistic concreteness is closely associated with imageability (Scott et al., 2019). The dimension of imageability will be used for comparison and to evaluate the validity of the concreteness detected. This study will use concreteness and imageability for the purposes of cross-verification, thus creating a novel and effective methodology for detecting patterns and trends in the evolution of fictional language. A key issue in understanding the development of language in English fiction is seeing whether it, like British English generally, also moved over the course of time toward an increasing abstraction. This study will address the two questions:

- (1) How did language in English fiction evolve with respect to its concreteness and imageability over the course of the last two centuries?
- (2) What cognitive/social factors have influenced the evolutionary pattern of language in English fiction generally?

Background

In recent years, more and more large collections of text data (corpus) have become available in digital form, thus boosting computational approaches. The approaches used in Culturomics (Michel et al., 2011) and Sentiment Analysis (Liu, 2012) are closely related to the present study. The following will provide the background to the two types of measures and discuss how they were computed.

Concreteness and Imageability

Words can roughly be classified into two groups: concrete and abstract. Generally speaking, concrete words refer to particular concrete entities, actions, or events. However, notions, emotions, social categories, and terms used to describe introspective phenomena are typically referred to with abstract words. Previous research has been able to recognize diachronic changes in language by means of these two factors, concreteness and abstraction (Hills & Adelman, 2015; Sneffjella et al., 2019). Changes of concreteness will be helpful in understanding a tendency toward becoming more abstract/concrete in words of English fiction. Such changes, or the lack therefore, can also show whether words in English fiction became more concrete and imageable or not. There are a number of factors that influence changes in linguistic concreteness. One example is word learnability. Numerous accounts of language evolution have suggested that language should evolve so as to be learnable (Christiansen & Chater, 2008; Smith & Kirby, 2008). Hills and Adelman (2015) found that changes of concreteness are associated with word learnability. That is to say, people could use more concrete words in order to facilitate word learning. Hills and Adelman (2015) associated with learnability driving by the language market.

Furthermore, although concreteness is the main measure, we will also measure another dimension, that of “imageability,” so as to allow comparison with the first measure and thus to test its validity. “Imageability” represents the degree of effort involved in generating a mental image of something (imageable, unimageable). Many studies have examined the effect of imageability on psycholinguistics and cognitive studies (McMullen & Bryden, 1987; Strain & Herdman, 1999). Despite the strong connection between imageability and concreteness (Paivio et al., 1968; Richardson, 1975), these two factors have rarely been employed as a means to detect a word’s different semantic aspects (Scott et al., 2019). A lower imageability degree shows that the word being used is less imageable and hence more difficult. We can find out whether or not English fiction has become more concrete and imageable over a given time-frame by looking at the changes that have taken place in concreteness and imageability.

Although often correlated with concreteness, imageability is not a redundant property. Most abstract things are hard to visualize; however, some still can easily evoke images. For instance, *pain* summons up an emotional and even visual image. However, there are also concrete things that are difficult to visualize, for example, *monastery* is harder to visualize than *monkey*. In this sense, concreteness and imageability overlap but they are not synonymous constructs (Connell & Lynott, 2012; Soares et al., 2017). Connell and Lynott (2012) suggest that a contributing factor to the concreteness rating of a word is the degree to which it is a typical exemplar of a concrete category (objects, materials, etc.). For these reasons, imageability can be used to cross-verify the reliability of changes in concreteness. That is to say, when concreteness and imageability have similar diachronic change trends, this could well indicate that such trends are reliable. In such a case, imageability can be treated as the baseline.

Quantitative use is reliant upon norms of imageability and concreteness. Recently, large-scale concreteness norms were collected using concreteness rankings as made by online participants (Brysbaert et al., 2014). The data on concreteness ratings in Brysbaert et al. (2014)’s database was drawn from more than 4,000 participants with the Amazon Mechanical Turk collecting all the information. An in-house experimental platform was used to collect the imageability data database (Scott et al., 2019). The participants were English speakers from the University of Glasgow.

Computational Methods

The method used to evaluate an individual word’s concreteness is not the same as that used to assess a corpus’ concreteness. The three principal methods used up to now to work out the degree of concreteness in a corpus are (1) to search a corpus for words (*lemmatized tokens*), then to gather their rating norms using the concreteness database and subsequently to determine what the mean of all these rating norms

is (Graesser et al., 2004). The second method, proposed by Hills and Adelman (2015), assesses the concreteness of a word in terms of a word's concreteness ratings as weighted by the word's frequency of occurrence in a corpus. The third method that was proposed by Hamilton, Leskovec et al. (2016) and Sneffjella et al. (2019) used semantic similarity as derived from the word embeddings (that are themselves obtained by transforming the corpus) in order to calculate the degree of concreteness of each word in a corpus. This was done by measuring the semantic similarity between "seed words" and all words in this corpus. The following discussion compares the three methods.

The Coh-Metrix (Graesser et al., 2004) tool used the first method to calculate the degree of concreteness for the content words in a corpus. Another tool, the Automatic Analysis of Lexical Sophistication (TAALES) (Kyle et al., 2018), was used to compute lexical features related to word concreteness. This tool employed an algorithm similar to the Coh-Metrix. However, there is a potential problem with the mean obtained here: the degree of concreteness for this corpus is not equal to the mean of the summation of the scores of all (unique) content words in this corpus. However, despite the fact that all the words were considered, this method of summation is still problematic because the summing up of the scores of all the words totally ignores the frequency of words, the usage of a lemma and its context. The problem with ignoring word frequency was discussed in the preceding. In the following, we want to illustrate the problem that arises from ignoring word usage and context. For instance, "taking" has the lemma "take" just as "took" has the lemma "taken." However, "taking" can be also used as a noun. If we calculate the degree of concreteness of "taking" (noun), should we also treat "taking" (noun) in the same way as "take" (verb)? It is clear that we should not assign the same score in an absolute fashion to various word forms with the same lemma. This means that this algorithm is incapable of precisely measuring the concreteness of a corpus.

The assumption that a word is perceived to have the same degree of concreteness/imageability for English users over all decades or used with the same degree of concreteness/imageability across may be problematic and this approach falls short of capturing such potentially different perceptions and usages. As discussed in the 2.1, numerous studies have explored the semantic changes words have undergone in history (Garg et al., 2018; Hamilton, Clark et al., 2016; Hamilton, Leskovec et al., 2016; Kozłowski et al., 2019). It is also possible that the same word has a different degree of concreteness/imageability for English users in different genres. For example, "chaos" may be an abstract word in fiction, but a concrete word in a scientific text. In the light of such cases, we believe that concreteness/imageability changes over time in different genres. That is to say, the trends in changes of concreteness/imageability in English fiction might differ from those in the English language generally.

The second method discussed by Hills and Adelman (2015) and Hills et al. (2016) considers the word's frequency of occurrence in a corpus as weight. The algorithm is potentially better than the first method because the algorithm in Hills and Adelman (2015) solved the problem of neglecting words that occur twice or more.

However, the effective calculation of linguistic concreteness can be carried out based on word embeddings (mainly computed by word2vec, GloVe, or FastText). Before reviewing the third method, we need to introduce one crucial concept. The meanings of words can be represented using vectors as part of a high-dimensional "semantic space." These vectors are called "word embeddings." Word embeddings have been widely applied in natural language processing, cognitive sciences, language studies, and diachronic changes in language and culture (ref. Bakarov, 2018; Garg et al., 2018; Hamilton, Leskovec et al., 2016). However, these studies are based on word association or semantic similarity. A key problem that must be considered in the current context is that of ascertaining which words are crucial in sciences that have a variety of subfields. It is obviously very hard to ascertain what the crucial words are when, as in this study, a wide variety of different kinds of fiction spanning two centuries are the object of study. However, instead of using word association, we can examine changes in word concreteness. Concreteness is a specific psychological construct with a rich theoretical literature, and a reliable means of detecting linguistic changes.

Hamilton, Leskovec et al. (2016) subsequently proposed a new algorithm for quantifying semantic change by evaluating the word embeddings. Sneffjella et al. (2019) borrowed the algorithm of Hamilton, Leskovec et al. (2016) to measure the changes of diachronic concreteness. If the algorithm of Hills and Adelman (2015) is "static," the method of Hamilton, Leskovec et al. (2016) is "dynamic." We will not use the static approach to evaluate historical changes (Hills & Adelman, 2015) where the sum of a word's concreteness ratings is weighted by the word's frequency of occurrence in a given decade. The use of contemporary norms would not be able to account for an abrupt change in the meanings of many words or a periodic fluctuation in meanings of many words over time.

A few remarks suffice to illustrate how this "dynamic" algorithm works. We want to calculate the concreteness of a given text. Using this "dynamic" algorithm, the first step is training the text into a database of word embeddings and thereby building a weighted lexical graph by connecting each word. Second, we choose the positive and negative "seed words" that must occur in this text. After that, random walks are run from the seed words and polarity scores (shown in Figure 1) are then assigned in terms of the frequency of random walk visits. Finally, we walk from both the positive and negative seed sets, resulting in positive ($\text{score} + (w_i)$) and negative ($\text{score} - (w_i)$) label scores. The mean of all such scores can be obtained. The mean is this text's degree of concreteness. For further details about this, see the section on Methods.

Figure 1. Visual summary of the *dynamic* algorithm of semantic similarity (Hamilton, Clark et al., 2016). In this graph, “love” and “hate” are “seed words” representing positive and negative dimensions. Based on the semantic similarity between the other words and “seed words,” the other words obtain their positive or negative scores. (a) Run random walks from seed words and (b) Assign polarity scores based on frequency of random walk visits.

Table 1. The Three Corpora Used in the Present Study.

Corpora	Years	Genre(s)	Tokens
English Fiction of Google Books (EngFic)	1820s–1990s	Fiction	32,851 million
CLMET	1740s–1920s	Literary works (50%) + treatises + non-fiction(?treaties)	34 million
COHA	1800s–2000s	Four genres: fiction (50%) + magazines + newspaper + non-fiction	400 million

Materials and Methods

Materials

The current study will use the three different types of corpora in order to make comparisons. Two of the corpora are comprehensive and contain various genres. The third one is the corpus of English fiction from the Google Books corpus.

COHA is a good balanced historical corpus; however, it only goes back to 1800s. To our knowledge, there are no large and balanced historical corpora currently available. The Corpus of Late Modern English (CLMET) can work as an alternative if comparisons are also made with other genres. The two corpora can thus represent the English language as a whole. Sneffjella et al. (2019) once trained the COHA into the 20 sub-corpora of FastText-style word embeddings according to one decade =.

The Corpus of Late Modern English Texts is a corpus of roughly 35 million words of British English from 1710 to 1920 (De Smet, 2005). As we know, COHA is balanced for different genres in each decade. By contrast, CLMET is not balanced, because there is a relatively significant portion of literary works. The annotation is XML style and similar methods to those used in RSC were employed to extract text and yield word embeddings. Comparing three corpora with different genres yields new findings. Google Books Corpus

contains different genres of books. The current study makes use of the selection of English fiction. The word embeddings in the selection were trained by Hamilton, Leskovec et al. (2016) (<https://nlp.stanford.edu/projects/histwords/>). This corpus represents the fiction genre and includes complete literary works. Table 1 provides the basic information on the four corpora used in the current study, including their duration, size, material genres.

The Google Fiction word embeddings are assigned a greater importance than the others in the present study. The dimension of word embeddings for each decade is 300. The calculation method used in the present study needs to find the “seed words” among the shared words for the sub-corpora that we want to process. If we include all 20 sub-corpora, it would mean that the shared word number is the minimum of all (686 words). Further, we need to choose “seed words” from the 686 words. When the shared word number is small, it will have a negative influence on the selection of “seed words.” The number of word embeddings in 1800s and 1810s is too small, so these were not included. In order to verify our methods, we formed two groups: the first starting from 1820 and the second starting from 1850. The first group has the 1,750 shared words and the second group has 4,499 shared words. Each group will have different “seed words” due to the different shared words. Further details can

be found in the following section. The size of word embeddings for the three corpora and “seed words” of concreteness and imageability can be seen in section 1 of the supplementary material (abbreviated as: **SM**). Here the division between fiction and non-fiction (Table 1) is possible in COHA and CLMET, but such a temporary division is not necessarily clear-cut. This is because one of the two points of focus in the present study is the examination of language changes in fiction. The other genres can be taken as non-fiction so as to enable comparison.

Methods

Dynamic algorithm. The “dynamic” algorithm of (Sneffjella et al., 2019) is potentially better than the “static” one. The “dynamic” algorithm, derived from Hamilton, Leskovec et al. (2016), can measure the historical concreteness of English words. The “dynamic” algorithm applies semantic similarity using word embeddings, as repeatedly mentioned above. We will now provide a brief introduction to this algorithm.

First, we divided a corpus into several small sub-corpora according to their historical time span (i.e., one decade), and then transform each sub-corpus into word embeddings using FastText (Bojanowski et al., 2017). Here we used dimension 300 with subword information ranging from character bigrams up to character 6-grams in order to maintain consistency with Sneffjella et al. (2019). The word embeddings of this sub-corpus actually contain the information on the frequency of each word in this sub-corpus. All target words we examined for concreteness/imageability are a set of words that occurred in each sub-corpus (words with frequency ≤ 2 were excluded). We then selected a small set of seed words using the criteria of concreteness and imageability and chosen by considering the factors that are discussed below.

“SentProp,” which belongs to the SocialSent package (Hamilton, Clark et al., 2016), is an essential tool at this stage of the investigation. We used SentProp to make two random walks through a semantic similarity graph and between each word (i.e., word tokens) contained in this sub-corpus (the word embeddings database in this corpus) and the “seed words.” Words and edges constitute the nodes of this graph. This is established through linking the individual words to 25 of their closest neighbors, semantically speaking. The *cosine distance* is used to detect these semantic neighbors. In this way, SentProp can work out the individual words’ semantic dimension (that is, their concreteness) *s*. Randomly samples from seed word subsets are made by SentProp in each sub-corpus. It then works out the scores for the seed words sampled by employing word embeddings. These semantic similarities become probabilities when the graph has been constructed, thus allowing us to create a transition matrix that shows the probability involved in a random movement

from one word to any of its semantic neighbors (Hamilton, Leskovec et al. (2016) provides a more detailed explanation of this point). For each individual word and for a random subset of seeds, we calculate the polarity scores with respect to both abstraction (*a*) and concreteness(*c*), the final score being $c/(c \pm a)$ and lying somewhere between 0 and 1. The word’s concreteness is high if the word score is close to one (and the other way around). To get the mean concreteness score of the sub-corpus, we need to divide the sum total of all the word scores from each sub-corpus by the number of words in this sub-corpus. In this way we establish the score for mean concreteness or the concreteness degree. A mean concreteness score that is close to 0 suggests that there is a low degree of concreteness (and the other way around). The algorithm can produce a concreteness value that reflects the usage within a specific corpus, that is the historical period of use or genre.

With regard to the setup of “SentProp,” we basically followed the settings of Sneffjella et al. (2019) in processing the word embeddings. However, as the number of “seed words” (20 positive and 20 negative) was a little larger than that in Sneffjella et al. (2019) (15 positive and 15 negative), we made small changes to the setups in the SentProp, namely, the 30 nearest neighbors, $\beta = .9$, a tolerance of $1-e8$, and 100 bootstrap iterations using 15 random seed words out of the 20 at each pole of the concreteness continuum to process CLMET and EngFic respectively. Note that Sneffjella et al. (2019) used 10 random seed words out of the 15 at each pole of the concreteness continuum (some details can be seen in the subsection of “seed words”). In spite of this, the setup difference does not make a great difference in detecting diachronic changes in the CLMET and EngFic. To make convenient comparisons, we used the same settings as Sneffjella et al. (2019) in processing the imageability in COHA with 15 positive and 15 negative words.

The core of this algorithm lies in the use of semantic similarity derived from word embeddings in order to calculate the approximate concreteness score of each word according to the “seed words” of concreteness. The co-occurrences of each individual word together with the other words contained in a corpus can be denoted using word embeddings, where one word’s context is identified by the different vectors. The following factors, which have been neglected or unsolved in the first and second algorithms, need to be fully considered: word frequency, word meaning and word context. Because it includes these factors, this computational algorithm is able to precisely estimate the concreteness degree of a corpus. What this means is that when it comes to assessing how concrete a corpus is, this algorithm enjoys considerable advantages over the previously discussed two algorithms. In the corpora, the highest point and the lowest point in the concreteness and imageability scores is denoted using high frequency “seed words.” We acquired the scores

employed for the concreteness of the “seed words” in our algorithm by drawing on the database created by Brysbaert et al. (2014), which has 40,000 lemmas for English words. The Glasgow database is most extensive of the available databases on imageability subjective norms (Scott et al., 2019). We will now examine how we selected the “seed words.”

“Seed words.” Sneffjella et al. (2019) took fifteen concrete and abstract words as “seed words.” They chose seed words that (i) occurred with high frequency in each decade and were associated with (ii) extreme concrete or abstract ratings. Similarly, we applied such criteria in choosing seed words. We set up the word frequency in each decade and chose words with high concreteness/abstractness or imageability/less imageability ratings.

We need to use the same “seed words” in each of the sub-corpora we are investigating and the number of sub corpora investigated determines the “seed words.” For example, we looked for 20 words that scored highest with respect to their concreteness scores and twenty with the lowest scores in concreteness scores (as identified by the concreteness database). Unqualified words were then excluded by taking into account those with a frequency of >10 . We continued until we had found the 40 words that scored the highest and the lowest with respect to their concreteness. The “seed words” of imageability are selected using the same method. In contrast to Sneffjella et al. (2019), we used 40 “seed words” (20 positive and 20 negative) to process CLMET and EngFic. Sneffjella et al. (2019) only used 30 “seed words” (15 positive and 15 negative) to compute concreteness in COHA. However, we believe that an appropriate increase in the number of “seed words” contributes to the validity of computed scores and robustness of the calculation. Despite this, our new setup and the original setup in Sneffjella et al. (2019) yielded the same trend in detecting diachronic changes. The “seed words” can be seen in section 2 of the SM.

Results

Validity of the Computed Scores

First, we need to test the reliability and validity of the computed concreteness and imageability in the three corpora. Figure 2 compares human concreteness and imageability ratings to the computed scores of concreteness and imageability aggregated across all the decades for the three corpora. The aggregated data is the mean of the same set of words across all the decades in one corpus. It turned out that all correlations were strong and significant.

Specifically, with regard to CLMET, our unsupervised model’s concreteness scores correlate with Brysbaert et al.

(2014) at $\rho = .56$ ($p < .001$) and imageability scores correlate with Scott et al. (2019) at $\rho = .53$ ($p < .001$). As for EngFic, our unsupervised model’s concreteness scores correlate with Brysbaert et al. (2014) at $\rho = .63$ ($p < .001$) and the imageability scores correlate with Scott et al. (2019) at $\rho = .68$ ($p < .001$). Sneffjella et al. (2019) have computed concreteness in COHA with $\rho = .70$ ($p < .001$). Our model’s imageability in COHA correlates with Scott et al. (2019) at $\rho = .66$. The p value in all correlations is smaller than .001. Note that the correlation coefficients ($\rho = .66, .56, .53, .63, .68$) would be described as “strong” positive associations because all rho values are greater than .5 (Cohen, 1992). The associations are clearly statistically significant ($p < .001$). This strongly suggests our computed scores are reliable and valid.

The Historical Changes in Concreteness/Imageability in the Three Corpora

COHA was used to estimate the diachronic change of concreteness and imageability for the English language, as is shown in Figure 3. When linear regression analysis is performed, different regression lines can be compared to see if their constants and slope coefficients are different. The dotted lines located in the center of a see-saw line are linear regression ones. The parameters of linear regression are displayed at the right-top. The negative value of slope coefficient indicates a decline trend, or vice versa. If the p value is smaller than .001, it will be followed by three stars, that is it has a strong statistical significance. If the p value is between .001 and .01, it will be followed by two stars, that is its statistical significance is not stronger than three stars. When the p value is greater than .05, there will be no stars, that is there is no statistical significance.

We found the concreteness in the COHA rose slightly since 1850s. By contrast, imageability shows a clear decline in the same time frame. The absolute value of the change in the slope coefficients of imageability (linear regression) is much larger than that of the change of concreteness. We can find that the imageability data does not converge with that of concreteness.

The data on concreteness and imageability from the CLMET does converge: both have negative slopes coefficients for linear regression, as shown in Figure 4. The p value in either concreteness or imageability is of significance here. It suggests that both concreteness and imageability in this corpus have markedly declined in the period from 1740 to 1920. They are both taken to be representative of the English language, yet both of these corpora exhibit contrary trends.

The CLMET exhibits an opposite trend with respect to linguistic concreteness. It indicates that the English language in this corpus has become more abstract during the period under consideration. However, the trends concerning

Figure 2. The correlations between human ratings and computed scores for the three corpora.

Figure 3. The diachronic change of linguistic concreteness and imageability in COHA (1850–2000).

Figure 4. The diachronic change of linguistic concreteness and imageability in CLMET (1740–1920).

changes in imageability from these two corpora are consistent with one another.

Figure 5 shows that concreteness and imageability in the Google Books Corpus of English Fiction increased from 1820 to 1990. The p value in both measures is very significant here. More important still, the two types of data converge. This indicates that English fiction has become increasingly concrete and imageable.

Discussion

Comparison of the Developmental Trends

We thus obtained the data concerning the slopes coefficient of linear regression of the data on concreteness/imageability and the p -values in each corpus. The value of slope coefficient and p -value potentially reflects the direction and degree of changes. The relevant data is summarized in Table 2.

As we know, CLMET and COHA are comprehensive corpora that contain texts from multiple genres. What also

Figure 5. Google Books Corpus of English Fiction (1820–1990).

must be considered here is that there is one negative slope coefficient for the two corpora, in other words, the positive slope coefficient of concreteness is in the minority. This means we found that concreteness has opposed trends in the two corpora. By contrast, the data on imageability in the two corpora converge. However, this does not indicate that English as a whole has become less imageable over the period of 250 years. The dimension of imageability has been not extensively explored, so the database size of its state-of-art rating norm is small (5,500 words, Scott et al., 2019). The seed words of imageability selected from this database seems not to be comprehensive. The imageability seed words are useful to detect changes in small-scale corpora like CLMET and EnglishFiction. However, detecting the changes in so large COHA may not work very well given we use these imageability seed words. We choose to believe that the decline of imageability is not realistic (i.e., imageability should have increased). The problem cannot be solved well until another larger database of imageability will become available. This is the limit of the dynamic algorithm, that is, the success of this algorithm depends on the quality of seed words to some degree. The selection of seed words correspondingly depends on the large-scale database of rating norm (more details can be seen in Section 4 of SM).

Additionally, Hills and Adelman (2015) showed using multiple corpora (Google ngrams, COHA, and presidential speeches) that concreteness has risen in American English over the last 200 years. The findings of Hills and Adelman (2015) are basically consistent with our findings. The discrepancy in concreteness between COHA and CLMET could be mainly due to the different backgrounds of the two corpora. The CLMET is made up of British English but the COHA represents American English. It is not difficult to conceive that British and American English could well have undergone different diachronic changes. We want to further analyze what caused such changes in the English language and in English fiction.

Table 2. Comparison Between the Development of Concreteness/Imageability in the Three Corpora.

Corpora	Slope coefficients of linear regression of the data on diachronic concreteness	Slope coefficients of linear regression of the data on diachronic imageability
Fiction in Google Books (1820-1990)	↗ 0.0002956 (<i>p</i> -value <.001 ***)	↗ 0.0003047 (<i>p</i> -value <.001 ***)
CLMET (1740–1920) [literary works (50%) + treatises + non-fiction(? treatises)]	↘ -0.002734 (<i>p</i> -value <.001 ***)	↘ -0.001814 (<i>p</i> -value <.01 **)
COHA (1850–2000)[four genres: fiction (50%) + magazines + newspaper + non-fictions]	↗ 0.0006 (<i>p</i> -value <.01 **)	↘ -0.003715 (<i>p</i> -value <.001 ***)

p*-value was under 0.05. *p*-value was under 0.01.

“Whether a word is acquired or processed with ease or with difficulty depends on a number of factors such as pronounceability, orthography, length, morphology, grammar, and various semantic features” (Schmitt & McCarthy, 1997, pp. 143–153). All these factors can be defined as “word learnability.” The semantic properties, including abstractness, idiomaticity and multiple meanings, may interfere with the learning process. It is generally assumed that abstract nouns are more difficult to learn than concrete nouns due to the more complex nature of the former. As discussed in the subsection 2.1, the concreteness changes trend could tally with the direction of the evolution of learnability. In order to learn language more easily and quickly, language users are likely to use more concrete words as time goes by. However, this general trend could vary in different genres or styles. We found that American English (COHA) may have become more concrete, but British English (CLMET) does not seem to follow this trend. There is a clear trend toward concreteness in English fiction. These findings are partly consistent with the ones of Hills and Adelman (2015). All of this potentially supports the hypotheses that language adapts to the cognitive demands of its users. However, we believe that language changes must intersect with more complex factors such as genre differences as well as social and cultural factors. Every genre has been created to serve a particular group of readers. In this sense, learnability in English fiction may be a little different from the English language as a whole.

The hypothesis in Sneffjella et al. (2019) could provide a better interpretation of the discrepancy in concreteness between COHA and CLMET. Sneffjella et al. (2019) hold that there are languages where “the population of its speakers (or of non-native learners of that language) did not grow over time or where the language marketplace became less crowded.” This argument should be inspired by the language market hypothesis proposed by Hills et al. (2016). The potential for demographic factors has extensively been investigated to influence language evolution (Bromham et al., 2015). Given the CLMET represents British English, its changes in concreteness/imageability could reflect the changes in the population of its speakers, particularly in the late 19th century with decline of the British empire. By contrast, the increasing population in and national power of the US from mid-19th century parallels the increase of concreteness in American English (COHA) (more details can be

found in section 5 in **SM**). Although in theory the English language should increase its degree of concreteness, following the demands of language learnability, historical developments, and population are also important factors that impact the evolution of concreteness. We can predict the increase of concreteness as a general (universal) phenomenon but this general phenomenon could be justified under some conditions because the change of concreteness in the CLMET does not support this general phenomenon.

Both concreteness and imageability in English fiction have positive slopes. The data on the two measures converges. This strongly suggests that English fiction has become more concrete (that is less abstract) and more imageable. It is for this reason that we assume that English fiction has become increasingly easy to read. It could also indicate that 19th century fiction is more difficult to read than modern fiction. The impact of changes in population can also explain such changes. With the rise of the English language on a world scale, English fiction also acquired a wider market. That means that English novelists would be likely to consider concrete and imageable words as a means for attracting more readers. The concreteness and imageability changes in English fiction are consistent with the past studies (Biber & Conrad, 2019; Biber & Finegan, 1989). These studies used frequency evidence from a number of English fiction texts taken from different periods in order to observe the patterns in its development. They found that English fiction from 18th century onward became less abstract and more concrete in its style.

Contrary to this trend in English fiction, imageability in the English language has decreased. This suggests that at the least the English language has become less imageable over the last two centuries. More discussions on this can be found in section 3 of **SM**. In contrast to this, English fiction has become more imageable. They thus exhibit two opposed trends as regards in imageability. We can also note that concreteness/imageability is likely to be subject to different changes over time in different genres. As mentioned in 2.2, the same word can be perceived by English language users as having a different degree of concreteness/imageability in different genres. English fiction, as a distinct genre to the English language itself, may have followed a different diachronic law because it has been influenced by very unique factors which will be discussed in the following subsection.

Other Causes of the Changes in English Fiction

In 5.1, we have discussed how learnability, genre difference and the changes of population size influenced the changes of concreteness/imageability. The following section analyzes what other factors could have caused the increase of concreteness/imageability in English fiction over the last two centuries. As discussed in the Introduction, fiction as a special style of writing plays a crucial role in human society. As an essential part of human culture, the language of fiction must also have been greatly impacted by social and cultural changes.

In the 1800s, people had no television, radio, movies or internet to distract them and they spent their leisure time on books and theater. In view of this situation, writers of fiction tended to engage in a lot of scene setting in their works much as one would in a play. For instance, they would typically give their readers a great deal of background on the characters, often add a bit of general history and were likely to tell their readers about the weather too. These descriptions were used to make readers feel as though they were watching a play. However, for modern readers, such scene-setting could seem rather unnecessary, redundant and tedious because such readers have more channels for the receipt of visual information due to the influence of movies, TV, color books etc. Consequently, the words used the 19th century authors seem unfamiliar to modern readers. For example, the pastures were not “green”; they were instead “verdant,” a term that describes the lushness of thick new grass. Such terms are clearly likely to confuse modern readers. Another factor here is that much that was familiar to the readers in Victorian era is not familiar to modern readers. The words used in Victorian era had specific connotations. However, these words have undergone drastic changes in their meaning or have been abandoned or used less frequently. Modern readers feel uncomfortable in reading such unfamiliar words. All these factors could make the words used in the 19th century fiction seem more abstract and difficult than the words employed in 20th century works.

Not only do the words in English fiction themselves change over time, but also the styles of writing do as well. For instance, Charles Dickens wrote in a style that was very popular among English authors in the 19th century. This style is very flowery, descriptive, and ornate, with a great deal of what modern critics often deride as “purple prose.” This style is the exact opposite of what most readers today look for in their novels. Modern readers’ expectations for novels have been heavily shaped by the styles of early 20th-century modernist writers such as Ernest Hemingway and John Steinbeck. The thesis that style in fiction has undergone linguistic changes over time has been much discussed (Leech & Short, 2007, pp. 298–303). Although the aforementioned writers’ respective styles were in many respects different, all wrote in a style that is deliberately much simpler, plainer, and less elaborate than the styles of Dickens and his contemporaries in

the 19th century. This could be a second reason why the 19th century fiction seems less concrete and imageable than that of the 20th century.

A final factor that needs to be considered here is the influence of social changes with respect to the readers’ background and the impact of this on the diachronic changes in the language of English fiction. As mentioned in 5.1, Biber and Finegan (1989) and Biber and Conrad (2019) concluded that language in English fiction from the 18th century onward down to the present has become increasingly concrete. They also provide some reasons for this in their analysis of these trends.

According to Biber and Conrad (2019, p. 156), the increase in literacy brought about a rise in demand for fiction that was simpler and easier to read. This could provide a reason why the language used in fiction became easier to read, namely such language encourages a larger readership and this in turn has commercial benefits. Fiction seems to have become an essential component of mass (popular) culture, although this thesis is controversial (Fluck, 1988). The evidence for this is that children’s fiction, science fiction, fantasy were all created in 19th century (or earlier) but became popular in the 20th century (Leavis, 2011; McCracken, 1998). These diversified fictional genres were written in plain language in order to attract more readers.

Therefore, Biber and Conrad (2019, p. 155) concluded: “18th century novels used an elaborated linguistic style, with long sentences and complex noun phrases, while 20th-century novels have changed so that they typically rely on a simpler style with more verbs, short clauses, and adverbials. The linguistic changes between 18th- and 20th-century novels reflect these changing attitudes about language. Thus, there is a fairly steady progression toward simpler, more colloquial styles in novels across these periods..”

Based on an analysis of population increase and changes in concreteness, we see that when the population of speakers of a given language did in fact grow or the language marketplace became crowded, the concreteness of this language is likely to diachronically increase. The increase in concreteness in English fiction indicates that the number of readers of English fiction has increased. This supports our argument that fiction plays a crucial role in human society. In comparison to fiction, the English language overall became more complicated in this period. We have already mentioned that the two main factors (word learnability and the rate of increase in population) potentially had a great impact on diachronic changes in concreteness/imageability. However, other social factors could also have influenced such changes. For example, numerous scientific writings were published as modern science developed. The greater numbers of people in higher education encouraged the creation of a more complicated language overall. By contrast, several social and cultural reasons probably lay behind the trend toward in concreteness in fiction, but these certainly did not impact in the entire English language, which was itself influenced by other unknown factors.

Conclusion

This study used a semi-supervised computational method to detect the evolutionary patterns in linguistic concreteness and imageability in the three corpora. Linguistic concreteness and imageability based on word embeddings were used to quantify the diachronic changes in word concreteness and imageability over a number of decades using the data contained in three corpora. We found that the COHA representing American English exhibited a marked increase in concreteness. By contrast, the CLMET representing British English exhibited a clear downward trend in both concreteness and imageability in this same period. The overall trends in diachronic concreteness and imageability in English fiction show a little difference from that of the British English as a whole. From the perspective of concreteness and imageability, English fiction has become more concrete and imageable over the last two centuries. This also indicates that English fiction has taken on an increasingly prominent role in human society. Moreover, the current study put forward what caused the systemic change of concreteness/imageability in English language and English fiction, for instance cognitive factors, genre difference, the changes of population size, social, and cultural factors etc. Our findings, which are based on “big-data” analysis, are confirmed by previous studies. Overall, this study clearly demonstrates how language in English fiction followed a general evolutionary pattern. The current study revealed how the evolutionary pattern of language in English fiction differed from that of the English language generally. This study created a novel methodology for examining evolutionary patterns in literary works and thus provided a new outlook on the cultural changes, which was described using a qualitative historical analysis.

Acknowledgments

We acknowledge support by Deutsche Forschungsgemeinschaft and Open Access Publishing Fund of University of Tübingen. We would also like to show our gratitude to three anonymous reviewers for valuable comments on an earlier version of the manuscript. We are also immensely grateful to the editor, Ms. Avanti Dhiman for her comments and great patience.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was also supported by ERC (project wide, no. 742545). The current study was also supported by Humanities and Social Sciences Project funded by Ministry of Education in China (no. 17YJC752032).

ORCID iD

Kun Sun <https://orcid.org/0000-0001-9766-269X>

Supplemental Material

Supplemental material for this article is available online.

References

- Bakarov, A. (2018). A survey of word embeddings evaluation methods. ArXiv preprint arXiv:1801.09536.
- Biber, D., & Conrad, S. (2019). *Register, genre, and style*. Cambridge University Press.
- Biber, D., & Finegan, E. (1989). Drift and the evolution of English style: A history of three genres. *Language*, 65, 487–517.
- Bojanowski, P., Grave, E., Joulin, A., & Mikolov, T. (2017). Enriching word vectors with subword information. *Transactions of the Association for Computational Linguistics*, 5, 135–146.
- Bromham, L., Hua, X., Fitzpatrick, T. G., & Greenhill, S. J. (2015). Rate of language evolution is affected by population size. *Proceedings of the National Academy of Sciences of the United States of America*, 112(7), 2097–2102.
- Brysbaert, M., Warriner, A. B., & Kuperman, V. (2014). Concreteness ratings for 40 thousand generally known English word lemmas. *Behavior Research Methods*, 46(3), 904–911.
- Carr, J. W., Smith, K., Cornish, H., & Kirby, S. (2017). The cultural evolution of structured languages in an open-ended, continuous world. *Cognitive Science*, 41(4), 892–923.
- Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31, 489–509.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159.
- Connell, L., & Lynott, D. (2012). Strength of perceptual experience predicts word processing performance better than concreteness or imageability. *Cognition*, 125(3), 452–465.
- Daly, N. (2000). *Modernism, romance and the fin de siècle: Popular fiction and British culture*. Cambridge University Press.
- Davis, L. J. (1997). *Factual fictions: The origins of the English novel*. University of Pennsylvania Press.
- De Smet, H. (2005). A corpus of Late Modern English texts. *Iceme Journal*, 29(29), 69–82.
- Eder, M., Rybicki, J., & Kestemont, M. (2016). Stylometry with R: A package for computational text analysis. *The R Journal*, 8(1).
- Fluck, W. (1988). Fiction and fictionality in popular culture: Some observations on the aesthetics of popular culture. *Journal of Popular Culture*, 21(4), 49–62.
- Fludernik, M. (2003). *The fictions of language and the languages of fiction*. Routledge.
- Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences of the United States of America*, 115(16), E3635–E3644.
- Graesser, A. C., McNamara, D. S., Louwerse, M. M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods Instruments & Computers*, 36(2), 193–202.
- Grieve, J. (2007). Quantitative authorship attribution: An evaluation of techniques. *Literary and Linguistic Computing*, 22(3), 251–270.
- Hamilton, W. L., Clark, K., Leskovec, J., & Jurafsky, D. (2016, November 1–5). *Inducing domain-specific sentiment lexicons from unlabeled corpora*. In Proceedings of the conference on

- empirical methods in natural language processing [Conference session]. Austin, TX, United States (pp. 595).
- Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016, August 7–12). *Diachronic word embeddings reveal statistical laws of semantic change* [Conference session]. Proceedings of the 54th annual meeting of the association for computational linguistics (Vol. 1: Long Papers, pp. 1489–1501). Berlin, Germany.
- Hill, F., Korhonen, A., & Bentz, C. (2014). A quantitative empirical analysis of the abstract/concrete distinction. *Cognitive Science*, 38(1), 162–177.
- Hills, T. T., Adelman, J., & Noguchi, T. (2016). Attention economies, information crowding, and language change. In M. Jones (Ed.), *Big data in cognitive science* (pp. 270–293). Psychology Press.
- Hills, T. T., & Adelman, J. S. (2015). Recent evolution of learnability in American English from 1800 to 2000. *Cognition*, 143, 87–92.
- Hughes, J. M., Foti, N. J., Krakauer, D. C., & Rockmore, D. N. (2012). Quantitative patterns of stylistic influence in the evolution of literature. *Proceedings of the National Academy of Sciences of the United States of America*, 109(20), 7682–7686.
- Kenny, A. (2013). *The computation of style: An introduction to statistics for students of literature and humanities*. Elsevier.
- Kidd, D. C., & Castano, E. (2013). Reading literary fiction improves theory of mind. *Science*, 342, 377–380.
- Kozlowski, A. C., Taddy, M., & Evans, J. A. (2019). The geometry of culture: Analyzing the meanings of class through word embeddings. *American Sociological Review*, 84(5), 905–949.
- Kyle, K., Crossley, S., & Berger, C. (2018). The tool for the automatic analysis of lexical sophistication (taales): Version 2.0. *Behavior Research Methods*, 50(3), 1030–1046.
- Labov, W., Ash, S., Ravindranath, M., Weldon, T., Baranowski, M., & Nagy, N. (2011). Properties of the sociolinguistic monitor. *Journal of Sociolinguistics*, 15, 431–463.
- Leavis, L. (2011). *Fiction and the Reading public*. Random House.
- Leech, G. N., & Short, M. (2007). *Style in fiction: A linguistic introduction to English fictional prose*. Pearson Education.
- Liu, B. (2012). *Sentiment analysis and opinion mining*. Morgan & Claypool.
- Mahlberg, M. (2013). *Corpus Stylistics and Dickens's fiction*. Routledge.
- McCracken, S. (1998). *Pulp: Reading popular fiction*. Manchester University Press.
- McMullen, P. A., & Bryden, M. P. (1987). The effects of word imageability and frequency on hemispheric asymmetry in lexical decisions. *Brain and Language*, 31(1), 11–25.
- Michel, J.-B., Shen, Y. K., Aiden, A. P., Veres, A., Gray, M. K., Pickett, J. P., Hoiberg, D., Clancy, D., Norvig, P., Orwant, J., Pinker, S., Nowak, M. A., & Aiden, E. L. (2011). Quantitative analysis of culture using millions of digitized books. *Science*, 331(6014), 176–182.
- Morin, O., & Acerbi, A. (2017). Birth of the cool: A two-centuries decline in emotional expression in Anglophone fiction. *Cognition & Emotion*, 31(8), 1663–1675.
- Paivio, A., Yuille, J. C., & Madigan, S. A. (1968). Concreteness, imagery, and meaningfulness values for 925 nouns. *Journal of Experimental Psychology*, 76(1), 1–25.
- Peng, R. D., & Hengartner, N. W. (2002). Quantitative analysis of literary styles. *The American Statistician*, 56(3), 175–185.
- Richardson, J. T. E. (1975). Concreteness and imageability. *Quarterly Journal of Experimental Psychology*, 27(2), 235–249.
- Schmitt, N., & McCarthy, M. (1997). *Vocabulary: Description, acquisition and pedagogy*. Cambridge University Press.
- Scott, G. G., Keitel, A., Becirspahic, M., Yao, B., & Sereno, S. C. (2019). The Glasgow norms: Ratings of 5,500 words on nine scales. *Behavior Research Methods*, 51(3), 1258–1270.
- Shennan, S. (2009). *Pattern and process in cultural evolution*. University of California Press.
- Smith, K., & Kirby, S. (2008). Cultural evolution: Implications for understanding the human language faculty and its evolution. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 363, 3591–3603.
- Sneffjella, B., Gnreux, M., & Kuperman, V. (2019). Historical evolution of concrete and abstract language revisited. *Behavior Research Methods*, 51, 1693–1705.
- Soares, A. P., Costa, A. S., Machado, J., Comesaa, M., & Oliveira, H. M. (2017). The Minho word pool: Norms for imageability, concreteness, and subjective frequency for 3,800 Portuguese words. *Behavior Research Methods*, 49, 1065–1081.
- Strain, E., & Herdman, C. M. (1999). Imageability effects in word naming: An individual differences analysis. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie exprimentale*, 53(4), 347–359.
- Sun, K., Liu, H., & Xiong, W. (2021). The evolutionary pattern of language in scientific writings: A case study of philosophical Transactions of Royal Society (1665–1869). *Scientometrics*, 126, 1695–1724.
- Zunshine, L. (2006). *Why we read fiction: Theory of mind and the novel*. Ohio State University Press.