LingBaW. Linguistics Beyond and Within 11 (2025), 7–22

Evolution in the Use of *Evolution*? An Overview of the Term in the Corpus of Historical American English

Shala Barczewska
Jan Kochanowski University, Poland

Abstract

While synchronous analyses of the interaction between language and society abound, e.g., in discourse studies, diachronic corpus-based studies are rare. This paper attempts to fill that gap with an investigation into the frequency and collocational preferences of *evolution* in COHA: The Corpus of Historical American English (Davies 2010). This lexeme was chosen for two reasons. First, the acceptance and teaching of biological evolution, especially in public schools, have been points of contention in American culture for 100, if not 150, years, comprising one of America's "culture wars." Hence, the topic is of contemporary as well as historical discourse interest. Second, a pattern between the frequency of *evolution* and the popularity of the theory in American history was noted by Barczewska (2017). This suggests a link between the use of the lexeme and the theory's reception. The current analysis investigates the frequency of *evolution* in COHA as a whole and according to genre. Collocational preferences within the corpus reveal changes in the way *evolution* is used over time. The paper also highlights the advantages and disadvantages of using COHA for similar research projects and suggests that the process applied here could be used to study the verbalization of other culture-shaping phenomena.

Keywords: corpus linguistics; COHA; evolution; cultural-linguistic analysis

1. Introduction

This paper represents a preliminary study of the collocates of *evolution* in American English as found in COHA: The Corpus of Historical American English (Davies 2010). The impetus for this study is a remark in Barczewska (2017) that the frequency of *evolution* in COHA drops significantly after the controversy over its teaching was ignited by the trial of John T. Scopes in 1925. This paper goes beyond Barczewska's observation by analyzing the instances of *evolution* in COHA and aligning them with events in US cultural and scientific history. It examines the corpus as a whole as well as sub-corpora divided according to genre and important phases in the reception of the theory. It also investigates the term's nearest collocates to identify patterns between the use of *evolution* and historical events. The results

LingBaW. Linguistics Beyond and Within, e-ISSN: 2450-5188
DOI: 10.31743/lingbaw.18366

were not separated into biological and non-biological senses of *evolution,* as we view the two senses as mutually influencing one another and expect cultural and historical events to impact both over time. This premise was validated by the data in this study and is evident in the examples from the corpus discussed in this paper.

Application of COHA for a cultural-linguistic study is something that, to date, has only been done on a limited scale. Moreover, those studies that have been done primarily focus on linguistic aspects, e.g., Jucker (2018), who analyzes apologies, or Ng et al. (2015), who study agism. Hence, this study also fills a gap in research by using a diachronic corpus, i.e., COHA, to analyze the intersection of language and society.

Section 2 provides the context of the study. It reviews the development of different views on biological origins and their subsequent influence on science standards and legal decisions regarding school curricula. Once the foundations of the historical and cultural context have been laid, the methodological approach used in this study is explained. This includes the choice of the corpus and the collocation span. The penultimate section presents the results and analysis. This section is divided into three subsections: the overall frequency of *evolution* in COHA and its distribution according to genre, the top ten 2L-2R collocates of *evolution* in COHA as a whole, and the top 2L-2R collocates of *evolution* according to time period. The conclusion draws together the findings, discusses limitations of using COHA for such a study, and provides suggestions for future research.

2. Teaching evolution: one of America's culture wars

Before beginning our linguistic analysis, it is important to understand the historical and cultural context of the reception of Darwinian evolution in the United States, particularly as it concerns teaching children in the public school system. Table 1 presents a general overview. Written vertically, the Theories of Origins column lists the period when theories about the origin of life gained prominence—their height in the graph roughly lines up with the corresponding Time Span in the middle section. The far-right column presents the focus of the debate over teaching evolution (and its alternatives) in the public school system during the corresponding time frame.

While it had its dissenters, acceptance of biological evolution and its cultural corollary, Social Darwinism, was widespread in the US at the turn of the twentieth century. School coursebooks in biology included words such as *civic* in their titles, and they taught the values of Social Darwinism, such as the now abhorrent notions of an evolution-based hierarchy of races and the need to sterilize the "feeble-minded" and "undesirables." It was not until the early 1920s that legislation regulating the teaching of biological evolution was proposed. Commentators suggest that this was, at least in part, a reaction to the horrors of scientific advancement in military capabilities during WWI and the fear that, if children were taught that they had evolved from animals, then such cruelty would only increase. The first law to include a punishment—a fine—was passed in Tennessee in 1925. It specifically outlawed teaching that man had evolved from a lower life form; teaching the evolution of other species was still permitted. Soon after, the American Civil Liberties Union advertised legal support for

Table 1: *Historical overview of the debate over teaching evolution (based on Barczewska 2017, p. 15, Larson 2003)*[1]

Theories of Origins and Their Popularization					Time Span	Public Education
Old Earth creationism	Darwinian Evolution (1859): theistic or naturalistic	Neo-Darwinian synthesis (1930-1940s)	Scientific creationism—young Earth (1960s)	Intelligent design (1984/1992)	1859–1919	Darwinian Evolution and Social Darwinism taught
					1920s	Bills passed outlawing teaching that man evolved from a lower life form
					1930–1959	Laws against teaching the evolution of man enforced
					1960s	Teaching of evolution legalized
					1970s	Bills passed legislating equal time for scientific creationism and biological evolution
					1980s	Scientific creationism declared religious by the courts and is removed from public school science curriculum
					1990s	Teaching evolution mandated in science standards
					2000s	Bills passed enabling teaching the strengths and weaknesses of evolution

any teacher willing to go to trial for breaking the law. City leaders in Dayton, TN, saw an opportunity to bring publicity to their town and convinced physical education teacher and substitute science teacher John T. Scopes to "confess" to having taught human evolution from *A Civic Biology* (G. Hunter 1914).[2] Prominent lawyers on both sides of the divide volunteered their services, and the event took on a circus-like atmosphere. Scopes lost and was fined, although the decision was later overturned on a technicality. As a result, schools and teachers nationwide decided to avoid the controversial topic of human evolution. Textbook publishers followed suit, and G. Hunter (1926) produced an updated version of *A Civic Biology* that reduced the frequency of *evolution* from 24 occurrences to one, which was in the reference section. Thus, in this book's first and second editions alone, the influence of the trial on textbook language is visible.

[1] As of 2014, a "Third Way" has been developing as an alternative to current theories in mainstream evolutionary biology. However, because their proposal does not presuppose a deity or intelligent designer and because they have restricted their work to academic studies, they have not encountered legal battles nor garnered much publicity (https://www.thethirdwayofevolution.com/).

[2] There is doubt as to whether or not he actually taught the evolution of man.

This was the status quo until the 1960s, when some science teachers took action to legalize the teaching of evolution.[3] *Epperson v. Arkansas* (1968) settled the matter at the national level, ensuring that the teaching of biological evolution in the public school system could not be banned. Around the same time, the notion of scientific creationism entered the mainstream with the publication of *The Genesis Flood: the Biblical record and its scientific implications* (Whitcomb and Morris 1961), revising and popularizing flood geology and young-earth creationism. As a result, evolution-related education legislation during the late 1960s through to the 1980s focused on providing equal time for scientific creationism and neo-Darwinian evolution: i.e., if one was taught, the other also had to be. However, this proved to only be a temporary solution.

The 1980s witnessed two cases resulting in the categorization of scientific creationism as "religion:" *McLean v. Arkansas Board of Education* (1982) and *Edwards v. Aguillard* (1987). This effectively removed the theory from the public-school science curriculum. The decision was based in part on *Lemon v. Kurtzman* (1971), which concluded that any school activity with religious content must have a secular purpose. In the 1990s, neo-Darwinian evolution became the only permissible explanation for the development of life, and science standards were rewritten to make it a central part of the biology program. However, not all scientists were content with the status quo.

During the mid-1980s and early 1990s, another opposition to purely random, materialistic evolution emerged: Intelligent Design (ID). Although attempts have been made to construe ID as *Creationism's Trojan Horse* (Forrest and Gross [2004] 2007), its research paradigm differs significantly. First, it makes no claims regarding the age of the earth or the reliability of the Bible, and fellows of the Discovery Institute (its main proponent) include Christians, Jews, Muslims, and agnostics. Second, the focus is not on geology and fossils as much as it is on genetics and information. For example, proponents argue that genes are irreducibly complex (Behe 2007), making progress via random mutation statistically improbable. Third, the Discovery Institute does not support teaching intelligent design in schools. Instead, they advocate for teaching more evolution, including the theory's strengths and weaknesses as expressed in scientific, peer-reviewed articles.[4] Despite its apparent scientific and secular motivation, this approach has not been met with approval by the National Centre for Science Education (NCSE), a leading organization dedicated to protecting evolution and climate change education. As a result, heated debates regarding science standards and curricula continue to crop up throughout the country.

Thus, the controversy is neither linear nor dichotomous. The debate carries across several platforms—from private to public, from social to legal—and represents one of America's many culture wars (see J. Hunter 1991).

---

[3] Teaching evolution was, in actuality, only outlawed in a small number of states.

[4] The policy of suggesting students read a book about intelligent design in the school library that lead to *Kitzmiller v. Dover Area School District* was not supported by Discovery Institute; they had actually advised against it.

3. Methodology

As mentioned in the introduction, the data for this study comes from COHA: The Corpus of Historical American English (Davies 2010). COHA contains over 475 million words from texts from 1820 to 2019. Compared to other corpora covering a similar time span, apart from the Google Books corpus, it is 50-100 times larger.[5] Although the Google Books Corpus contains an impressive 155 billion words of American English, it was not chosen for this study for three specific reasons. First, there is the issue of accuracy: often books appear in years they were not published, or, alternatively, they are over represented because they were released in different formats a few months apart. Second, the corpus is heavily weighted towards scientific books; thus, it is not an accurate representation of overall language use. Third, the interface does not allow the same depth of analysis as COHA. Specifically, the interface takes the user to a list of Google Books and not specific examples of the word/phrase in context.[6] Hence, the types of analyses conducted in this paper would not have been possible using Google Books.

The study was conducted in the following manner. The COHA interface was used first to identify the frequency of *evolution* according to decade and genre. Then, it was used to identify collocates in a 2L:2R span with a minimum MI score of 3. This was done for the corpus as a whole, as well as according to distinct phases according to the word's frequency. Our goal is to answer the following questions:

- Does the corpus provide support for the suggestion in Barczewska (2017) that it was the Scopes Trial and not, e.g., the Great Depression, that led to a decrease in the frequency of *evolution*?
- Does the use of *evolution* in COHA parallel events surrounding the theory in US history?
- How does text type impact the frequency of *evolution* in COHA?
- What do the top ten overall collocates of *evolution* in COHA reveal about the term, its senses, and/or its reception during these decades?
- Is there a visible difference in collocates of evolution during different phases of the theory's reception?

It is believed that the answers to these questions will also help establish the advantages and disadvantages of using COHA for similar research projects, thereby contributing to broader applications of the corpus.

Section 4, Results and analysis, is organized as follows. First (4.1), the presence of *evolution* in the corpus as a whole and according to genre is compared with the historical context discussed in Section 2. Then, the top ten collocates are listed and divided into three categories according to their distribution in COHA over time (4.2). In 4.3, the corpus is divided into three subcorpora according to the phases identified in 4.1, and 2L-2R collocates were identified for each. This enables a closer examination of each time period's unique relationship with the theory.

---

[5] Data come from english-corpora.org, where the mentioned corpora can be found.

[6] For an overview of using Google Books for linguistics study and its limitations, see Pechenick et al. 2015; Friginal et al. 2022.

4. Results and analysis

4.1. The frequency of evolution *in COHA*

The first step is to examine Barczewska's (2017) comment regarding the frequency of *evolution* in COHA. As Figure 1 demonstrates, its use appears to coincide with the historical events discussed in Section 2.

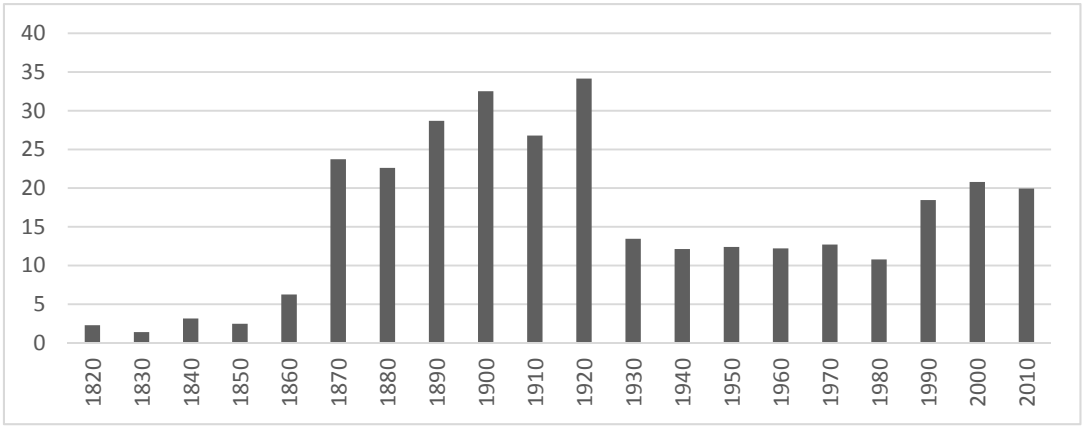

Figure 1: *Instances of* evolution *in COHA, words per million*

As the graph shows, occurrences of *evolution* in the corpus nearly quadrupled between 1860 and 1870 (6.26 and 23.74 tokens per million, respectively) when Darwin's theory was introduced and disseminated. The use of the word remained high, increasing to 34.16 words per million in 1920. This suggests that the theory was of increasing prominence and interest in American discourse. However, this changed drastically in the 1920s. Following the Scopes trial in 1925 and the subsequent removal of *evolution* from classroom science books, the frequency of *evolution* declined to 13.46 ww/mil. Although one could argue that the decline was due to the distractions of the Great Depression and international affairs in Europe, the data in Figure 2 seems to confirm that the verdict in the Scopes trial was the motivator, as we see a sharp drop after 1925. Word frequency does not begin increasing significantly until 1990, as the teaching of evolution becomes mandatory, but even then, it does not reach 1920 levels at just 18.46 words per million.

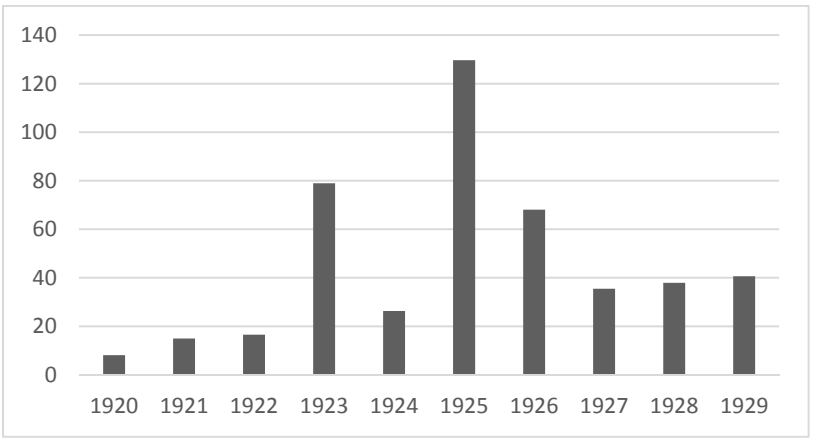

Figure 2: *Instances of* evolution *in COHA in the 1920s, words per million*

The data in Figure 1 also suggest three possible phases, which I would like to label *reception and fascination* (1860s-1920s), *trepidation* (1930s-1980s), and *embracing the conflict* (1990s-2010s). During the first period, there was an intensified interest in the theory and its implications. During the second, *evolution* appears to have become almost taboo in common discourse. During the final period, discussions about *evolution* again became relevant in the public sphere, albeit not as prominently as at the turn of the 20th century. This coincides with what we know about the intensity of the approval/disapproval of Darwin's theory during those time periods.

However, that is not the only way to divide the corpus data. We could also look at it in terms of genre, as presented in Figure 3.

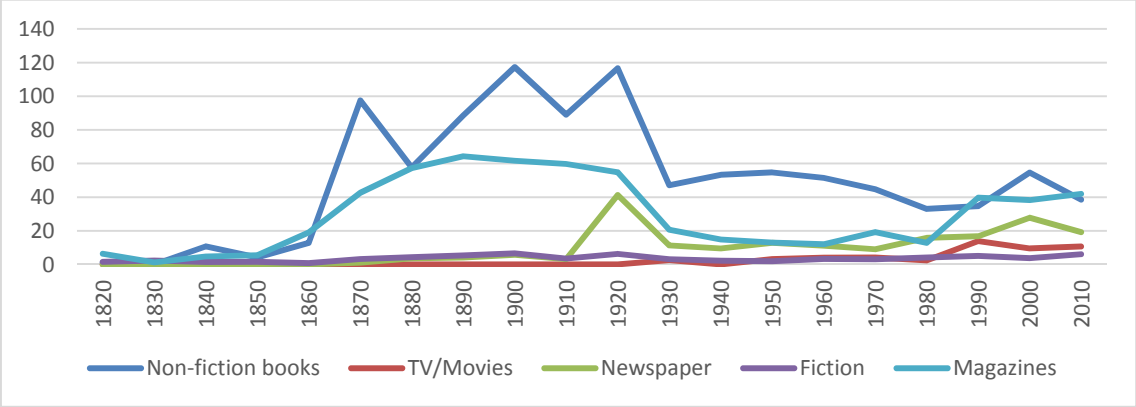

Figure 3: *Occurrences of* evolution *according to genre*

The results suggest that potential intersections between contemporary events and their lexical output are also constrained by genre. As we can see from the graph, newspapers tend to mention *evolution* primarily during periods of intense debate over the theory's teaching: the 1920s and again from the 1980s onward. Non-fiction books and magazines include the word *evolution* more frequently in their texts than any other genre, which is due, in part, to the nature of such material. Specifically, a book or article about a given topic, e.g., *evolution*, will mention it several times. An example of this phenomenon can be found in occurrences of *evolution* from the 1870s (according to COHA), in which over half come from the book *The History of the Warfare of Science with Theology* (White [1896] 1993).[7] Not surprisingly, *evolution* does not appear in movies until the 1930s, after sound has been introduced and more complex dialogues are possible. Five of the 1930s productions represented in COHA contain *evolution*, only one occurrence of which means *change* and not biological evolution. *Evolution* is not present in this sub-corpus during the 1940s, possibly due to concerns related to World War II, but it returns in the 1950s and continues to increase in frequency. *Evolution* even appears in a discussion between Ross and Phebe in the popular *Friends* TV sitcom, emphasizing the salience of the theory in contemporary culture. On the other hand, although written works of fiction discuss *evolution*, those represented in COHA do not seem to change

[7] It is not clear if the material in COHA is misdated or if the creators of the corpus had access to an older transcript. All searches for White's book date it 1986; however, Wikipedia does mention that it is based on a 1874 lecture that was published the same year.

in frequency according to contemporary events. There is only a slight peak in the 1920s and none in the 1980s, when TV/movie occurrences outrun those in fiction books.

4.2. The top ten collocates of evolution

Along with word frequency, we also see a shift in collocational preference. However, the results do not exhibit the expected level of variation. This is in part due to the fact that *evolution* is significantly more frequent between 1870 and 1920, and, as a result, MI scores and frequency counts inevitably favor collocations used during this period. The discussion in 4.3 attempts to correct for this and presents collocates according to the epochs identified in 4.1.

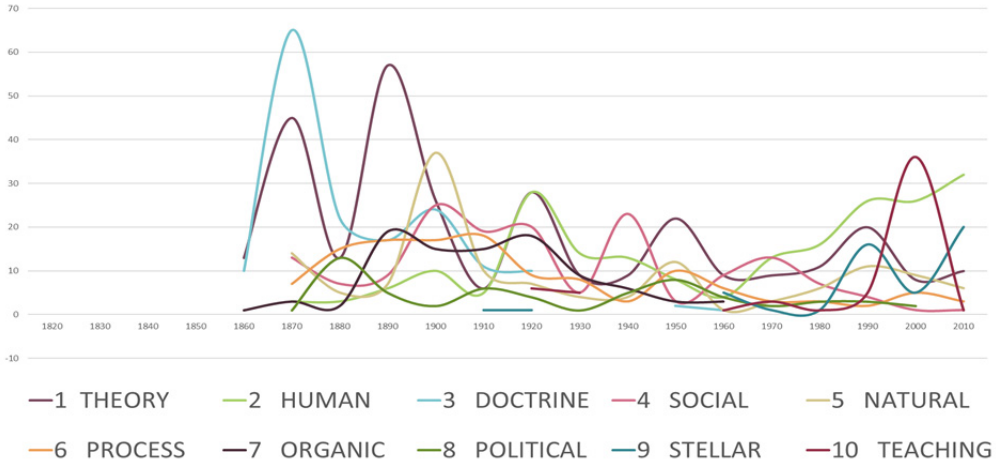

Figure 4: *Top ten collocates of* evolution *from 1820-2010*

There are two different ways to organize these collocates. If we categorize the collocates grammatically, there are four nouns—*theory*, *human*, *doctrine*, *process*; five adjectives—*social*, *natural*, *organic*, *political*, *stellar*; and one gerund—*teaching*. If we group the collocates according to dispersion, we also have three groups:

- collocates that co-occur with *evolution* in all decades: theory, human, social, natural, process
- collocates whose decline in use parallels the word's overall presence in COHA: *doctrine*
- collocates that vacillate in distribution: *organic*, *political*, *stellar*, *teaching*

As this study focuses on the intersection of history, culture, and language, the second grouping, which is done according to dispersion, has been chosen for analysis. The final group of collocates listed above, those that appear to fluctuate alongside contemporary events, receive the greatest attention.

4.2.1. Collocates that co-occur with *evolution* in all decades

Five of the top six collocates fit the first group; that is, they occur in each decade. *Theory*, despite visibly decreasing in frequency over the years, is a collocate across the decades. This is

not surprising, as it is integral to the description of evolution. It peaks in 1870 and 1890 with 45 and 57 co-occurrences, respectively, and four other decades also see twenty or greater co-occurrences. While not as frequent, *human* is also a consistent collocate, with peaks in the 1920s, 1990s, 2000s, and 2010s (28, 26, 26, and 32 co-occurrences, respectively). This could point to our focus, as humans, on our own anthropological and ancestral history. *Social,* with peaks in the 1900s, 1920s, and 1940s (25, 26, and 23 co-occurrences, respectively), often references Social Darwinism, a controversial notion related to Darwin's theory that seems to have declined in interest post-WWII. *Natural* (peaking between 1900 and 1909 with 37 co-occurrences) and *process* (peaking in the 1910s with 19 co-occurrences) are important attributes of biological evolution and are frequently discussed within both scientific and popular contexts. Because of their regularity, it can be argued that these five collocates have been, and continue to be, the most salient aspects of our shared knowledge about evolution. Despite their apparent benign nature, they are all contested, as the following excerpts demonstrate.

(1) California's decision to *teach evolution* as both fact and *theory* is a slippery equivocation that alloys science with dogma. (*Time Magazine* 1990)

(2) when we clearly see the results which follow from a rigid adherence to the doctrine of *natural evolution*, we can not help asking whether a grave mistake has not been made in attempting to explain intelligence and morality by a principle which necessarily excludes all freedom either in knowing or in willing. (*An Outline of Philosophy* 1901)

As suggested in (1) and (2), controversies involving these collocates include

- whether evolution is a *fact* or *theory*,
- if the *process* of evolution is fully *natural* or involves supernatural guidance or intervention,
- whether *human* evolution differs from that of animals, and
- if human society is an example of *social* Darwinism.

These questions have been debated by scientists, philosophers, and laymen since Darwin published his theory. They appeared in the Wilberforce and Huxley debate in 1860 and have found their way into legislation and legal trials from the 1920s to the present. Thus, they remain central to the concerns over the theory's scientific accuracy as well as applicability beyond biology.

4.2.2. Collocates whose decline in use parallels the word's overall presence in COHA

Doctrine, although statistically the third most frequent collocate overall, is only a strong collocate between the 1860s and 1920s. It co-occurs with *evolution* most frequently in the 1870s; however, it virtually disappears as a collocate in the 1920s, only to briefly reappear in the 1950s and 1960s. This coincides with the word's general usage, which, according to COHA, began declining in the latter half of the nineteenth century (Figure 5).

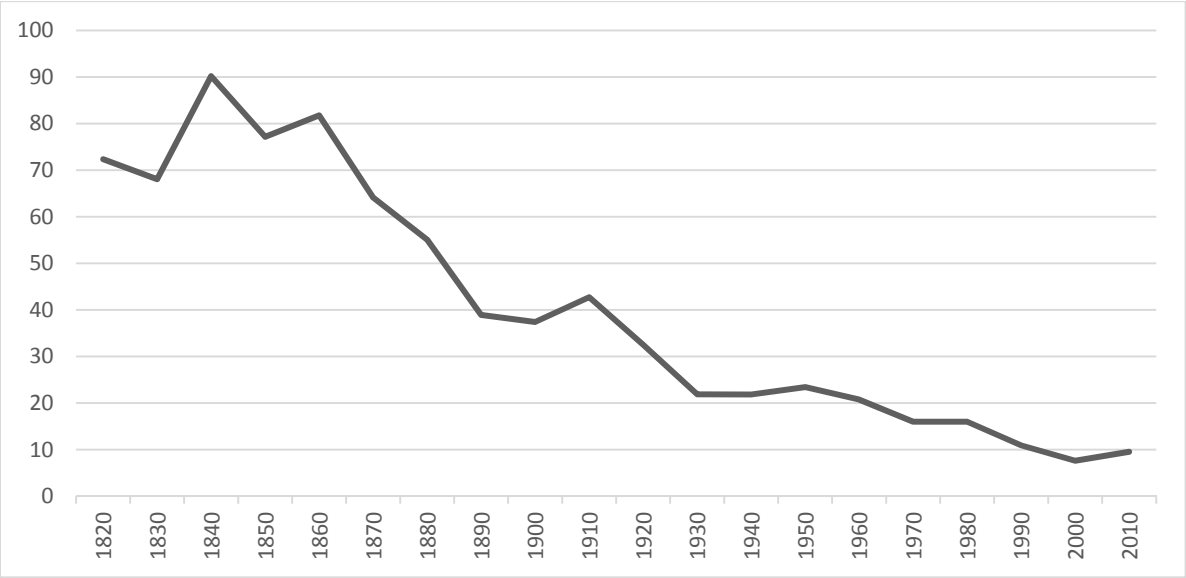

Figure 5: *Occurrences of* doctrine *in COHA according to decade, words per million*

These results could, in part, be due to the narrowing of the word's usage. Although contemporary dictionaries allow for *doctrine* to be used in science and politics, it is currently most often associated with the domain of religion, as noted in the Collins COBUILD Advanced Learner's Dictionary (2025): "A doctrine is a set of principles or beliefs, especially religious ones." Hence, at first glance, it would be tempting to see the late 19th-century phrase *doctrine of evolution* as metaphorical, perhaps suggesting that evolutionism is some sort of religion. However, a closer look at examples from the corpus suggests non-metaphorical uses are the most prevalent.

(3) and this is the *doctrine of evolution* which, if it be not proved conclusively, has great probability and great scientific (*Physics and Politics* 1872)

(4) contention that has declared itself between us. For on one hand the *doctrine of evolution* which Virchow attacks has already so far become a sure basis of biological science (*Freedom in Science and Teaching* 1879)

Instead, it appears as if we have an archaic use of the word, as these excerpts from the 1870s come closer to matching Webster's (1828) definition:

> In a general sense, whatever is taught. Hence, a principle or position in any science; whatever is laid down as true by an instructor or master. […] Hence a *doctrine* may be true or false; it may be a mere tenet or opinion.

As such, the use of *doctrine* at the turn of the century is neither metaphorical nor does it belong to the domain of religion. Thus, *doctrine of evolution* in (3)-(4) could be glossed as *teaching about evolution* or *theory of evolution.* This underscores the importance of not assigning contemporary meanings anachronistically, as there are often subtle changes or shifts in meaning (e.g., *organic*, discussed in 4.2.3, and *parent*, discussed in Person [2019]). Recognizing these shifts is important for discourse analyses and metaphor identification (cf. Steen et al. 2010) so as to avoid jumping to inaccurate conclusions.

4.2.3. Collocates that vacillate in distribution

The remaining four collocates—*organic*, *political*, *stellar*, *teaching*—appear, disappear, and reappear across the decades. Their inconsistent use should not be surprising, as they are less frequent collocates overall. Indeed, one of the reasons for setting the cut-off at ten is that the collocates were less and less consistent across decades. We will look at each of the final four collocates individually.

Organic collocates with *evolution* from the 1870s until the 1960s and only reappears in the 2000s. This is inconsistent with the overall frequency of *organic* in COHA, which has varied over the years and peaked in the 2010s. Moreover, the instances of *organic evolution* from the 2000s question whether or not *organic* could be viewed as a salient collocate of *evolution* for this time period. First, they only come from two sources: *Evolution: the remarkable history of a scientific theory* (three occurrences) and an article in *Psychology Today* (one occurrence). The former is a book by Edward Larson (2004), a respected scholar of the debate over evolution, who, incidentally, is also cited in Section 2 of this paper. Two instances of *organic evolution* from his book are given in (4).

(5) On the matter of *organic evolution* […] He studied it carefully (albeit not in the light of Darwin's later arguments for it) and found it wanting. […] Creation implies a creator, and so to dispense with the need for a biological creator, such ancient philosophers as Anaximander, Empedocles, the atomists, and the Epicureans advanced various crude notions of *organic evolution*.

In the excerpt, Larson describes early 19th-century scientist Georges Cuvier's unwillingness to accept *organic evolution* as scientifically viable (5). Thus, although the book was written in the 21st century, it borrows phraseology from 150 years prior.

The use of *organic evolution* in (7) could be viewed as metaphorical, as it does not refer to biological evolution but to spiritual or philosophical changes that the speaker has experienced due to her adaptation of Ayurveda. She tells her interviewer:

(6) It has been an *organic evolution*. Through my practice of yoga I was drawn to Ayurveda, a sister philosophy, and also to things that are generally better for me. My diet, and also the way that I respond to the world and the world to me, has gradually changed. (*Psychology Today* 2001)

Thus, these uses are either outdated, as in the case of Larson's book, or not related to biological evolution, as in the latter. Indeed, a comparison of the collocates of *organic* at the turn of the century and those of the more recent decades shows a shift from the domain of natural sciences to that of food and farming practices. As a result, we can say that *organic evolution* belongs to the earlier decades of the theory's development.

Political co-occurs with *evolution in* COHA three times before Darwin published his theory, all within the same article. However, a shift in meaning can be detected between the 1825 (7) and post-Darwin uses.

(7) Consider, again, the study of *political evolution*. It is true this is open to the most varied presentation, but its importance in the education of those who are to lead the people justifies it. (*North American Review* 1825)

(8) No doubt it may be urged that the *political evolution* has necessitated all this, and that is quite true. But it is also true that the *evolution* of events is aided or hindered by the wisdom or unwisdom of the individuals who direct public movements. (*New England Magazine* 1889)

(9) What is certain is that World War II compressed half a century of *political evolution* and economic change into not much more than a decade. (*Atlantic Monthly* 1959)

In (7) *evolution* could be replaced by *change*; however, in (8) and (9) it could not. The phrase *aided or hindered by the wisdom or unwisdom of the individuals who direct public movements,* by being set in opposition to *political evolution*, suggest two oppositional forces—natural and human-directed—and a conceptual comparison with the biological meaning. This comparison between biological evolution and political change seems to also motivate (9). Here, the conceptual link is the reference to the time that evolutionary processes usually require. The frequency of the co-occurrence of *political* and *evolution* peaks significantly in the 1880s (18 co-occurrences), and it appears in every decade up until the 2010s. This decline raises questions beyond the scope of this paper, e.g., is this decrease due to global political situations, or does it reflect a decreasing belief in the notion of unguided, natural political change in favor of some other mechanism?

The ninth-ranked collocate is *stellar*. Although the question of stellar evolution has been of interest to scientists since at least the 1890s, it seems to have peaked in common discourse in the 1990s. This could be a response to the publication of the first edition of the seminal *Stellar Structure and Evolution* (Kippenhahn and Weigert 1990). Instances of *stellar evolution* in the 1990s in COHA come from various sources; however, all but two of the 2000s and 2010s occurrences come from the same two journals, *Mercury* (2000s) and *Astronomy* (2010s), but not the same issues.

Finally, let us look at *teaching* and *evolution*. As much of the debate in the US revolves around education, the presence of this collocate is not surprising. What is surprising, perhaps, is that it did not rank higher. There are two possible reasons for this. One is the type of statistical calculation used. As *stellar* is less frequent in the corpus in general, its collocation strength with *evolution* is greater than that of *teaching*, which has other, more frequent collocates of its own. This hypothesis is confirmed by searching for both *stellar* and *teaching* independently in the COHA: s*tellar* occurs 3.06 times per million words, whereas *teaching* occurs 44.99 times. Moreover, *evolution* is the most frequent 2L:2R collocate of *stellar* but the 26th collocate of *teaching*. The other reason is that large collocate searches in COHA, such as this, do not lemmatize. If the lemma TEACH is sought within the 2L:2R span of *evolution*, 38 additional co-occurrences are found.[8] Co-occurrences of *teaching* and *evolution* underscore the controversy:

(10) "Why, yes, we're *teaching evolution*," said the young man, prepared to argue about it. (*Sons of the Puritans* 1939)

[8] When the lemma TEACH is the search term, only ten instances of *teaching* are retrieved instead of the expected 48. This is because the computer has only tagged those instances as VERB+ING. Any future study on the co-occurrences of *teach(ing)* and *evolution* would need to take this into account.

(11) Kraemer (1995) found that 40 percent of biology teachers in Minnesota spend little or no time *teaching evolution*, while 15 percent include creationism in their classes. (*BioScience* 2004)

Most co-occurrences of *teaching* and *evolution* in the 2000s, including (11), come from the article "How well do biology teachers understand the legal issues associated with the teaching of evolution?" The dialog in (10) comes from a work of fiction.

4.3. Comparison of collocates according to phases identified in 4.1

There are two problems with treating all of COHA as one whole for collocate identification. First, since *evolution* is more frequent at the turn of the century, collocates during that period have the greatest weight. Second, as a result, important collocates in later decades are less visible. Hence, a decision was made to identify the top fifteen 2L-2R collocates of evolution for the three phases identified in 4.1.[9] Of the ten collocates examined in 4.2, only three—*theory, natural, human*—occur in the top 15 of each time period and thus can truly be viewed as overall collocates of *evolution*. Another four—*doctrine*, *social*, *organic*, *process, political*—occur in the first two phases but not in the third. Of the remaining two, *stellar* only occurs in the latter two phases, and *teaching* only in the final phase, from 1990 to 2019.

Table 2: *Collocates of* evolution *according to the phase of the word's frequency in COHA*

	***Reception and fascination* 1860-1929**		***Trepidation* 1930-1989**		***Embracing the conflict* 1990-2019**	
	Collocate	**Hits**	**Collocate**	**Hits**	**Collocate**	**Hits**
1	THEORY	188	THEORY	69	HUMAN	84
2	DOCTRINE	159	HUMAN	68	TEACHING	42
3	SOCIAL	93	SOCIAL	60	STELLAR	41
4	PROCESS	83	PROCESS	33	THEORY	38
5	NATURAL	80	NATURAL	30	NATURAL	26
6	ORGANIC	73	POLITICAL	23	TEACH	18
7	HUMAN	55	GRADUAL	22	ECOLOGY	17
8	LAW	44	CULTURAL	22	BIOLOGICAL	16
9	STAGE	37	ORGANIC	21	CREATIONISM	15
10	POLITICAL	31	BIOLOGICAL	17	UNDERSTANDING	15
11	MENTAL	23	ECONOMIC	17	MICROSTRUCTURAL	13
12	GRADUAL	22	STELLAR	16	DIRECTED	13
13	CREATIVE	21	PARALLEL	13	BIOLOGY	11
14	HEAT	19	HISTORY	13	GALAXY	11
15	STAGES	18	STAGES	12	DARWINIAN	10

In the decades immediately following the publication of Darwin's theory, collocates suggest a focus on the theory and its scientific applications. However, extensions into other spheres are also visible, e.g., *social, political*, *mental, creative*. Below are examples of the latter two.

[9] Fifteen was used as the cut-off point instead of ten to increase the chances of finding shared collocates across time periods.

(12) For cases, see Animal Intelligence, in the chapters on Ants and Bees; and, for discussion of principles, *Mental Evolution* in Animals, in the chapters on Instinct. (*Darwin after Darwin* 1892)

(13) this explanation would place within comparatively recent times the *evolution of mental* powers which have distinguished the race from the most ancient times. (*Atlantic Monthly*, April 1896)

(14) He recognizes, as a truth for him, that theory of *creative evolution* which holds that in the ascending progress of the race each thinking person becomes a species by himself. (*The American Spirit in Literature* 1918)

The context suggests that both *mental evolution* in (11) and *evolution of mental powers* in (12) refer to the actual biological evolution of the mind and its abilities, rather than a metaphorical concept. The majority of the references to *creative evolution* discuss Henri Bergson's book by the same name, published in French in 1907 and English in 1911. From (13) it is unclear whether or not *creative evolution* is a biological phenomenon or metaphorical extension, but in either case it is connected with the biological notion.

Collocates from the 1930s to the 1970s suggest greater focus on evolution's social applications: *political, cultural, economic, history*.

(15) I would also suggest that in *cultural evolution* we might introduce the concept of the struggle for maintenance (*A Scientific Theory of Culture* 1944)

(16) But when Lenin became boss of Russia, people had to comprehend the logic of *economic evolution* the way he comprehended it and adapt their social consciousness to his interpretation (*Harpers* 1934)

(17) the critics are right in pointing out that new material needs also have been carried to the fore by *social* and *economic evolution*? (*Harpers* 1961)

Excerpts (15)-(17) also exemplify potential metaphorical uses of the term. They, along with the previous examples, also confirm the assumption made at the beginning of the study that many apparently non-biological uses of *evolution* are conceptually linked to Darwin's theory.

Collocates of recent decades—*biological, biology, teach, teaching, Darwinian, creationism*—underscore the national division regarding evolution.

(18) "I talk a lot about the holes in *evolution*; students need to know this information more than that they came from monkeys"; "I present *evolution* and *creationism* and let students make up their own minds" (*BioScience* 2004)

(19) issued a statement accusing Jones of trying "to stop the spread of a scientific idea and even to prevent criticism of *Darwinian evolution* through government-imposed censorship rather than open debate, and it won't work." (*San Francisco Chronicle* 2005)

The quotes in (19) come from the same article as (11) and, combined with (20), display the complexities of the viewpoint and the struggle between censorship and free speech.[10]

[10] The event referenced in (20) is Judge Jones' ruling on the *Dover* case mentioned in Section 2. See DeWitt et. al. (2006) for more context.

5. Conclusions and ideas for future research

The data presented above represent an introductory study of the presence of *evolution* in COHA. As a pilot study, this research was successful in its area of focus. Frequency data coincided with events relevant to the reception of the theory of evolution and the debate over its teaching in the US public school system. For example, a sharp decline in the word's usage was observed after 1925, the year of the (in)famous trial of John T. Scopes. Only when the theory of evolution became mandated in school curriculum did word frequency increase again. It also revealed differences in lexical output according to genre. Specifically, non-fiction works seemed to most closely reflect the intensity of the controversy surrounding evolution at a given time. The investigation into the top ten overall collocates of *evolution* gave both expected and surprising results. Expected collocates—theory, *process*, *natural,* and *teaching*—were interspersed with those whose co-occurrence with evolution appears archaic, e.g., *doctrine* and *organic*. This is because these words' main meanings have changed in the past 150 years. Finally, comparison of the collocates of *evolution* according to time period demonstrated different focal points at different phases of the theory's reception. Excerpts discussed in 4.2 and 4.3 suggest that non-biological collocates of *evolution*, e.g., *political, mental, economic,* often have some conceptual link to Darwin's theory.

This study also highlights the possible limitations of using COHA as a window to cultural-historical events. First, the frequency of a word or phrase may be the result of one text in the corpus. Second, COHA does not allow lemma searches at this level, and the collocates are limited by their word form. Third, quotations from the past may present a false indication of current use. Fourth, a collocate's frequency may not be as dependent on its association with the node word but more dependent on the uniqueness of the collocate in the corpus as a whole, e.g., *stellar*. While these potential pitfalls are not new to corpus linguists, they do reinforce the important role of the human researcher in corpus-based studies.

In conclusion, it is hoped that this study not only provides a better understanding of the shifting collocational patterns of *evolution* but also a possible methodology for studying aspects of culture (and culture wars) diachronically via their representation in corpora, thereby filling a gap in research on the intersection between language and society.

References

Barczewska, S. 2017. *Conceptualizing evolution education: A corpus-based analysis of US press discourse*. Newcastle upon Tyne: Cambridge Scholars Publishing.

Behe, M. J. 2007. *The edge of evolution: The search for the limits of Darwinism*. New York: Free Press.

Collins COBUILD Advanced Learner's Dictionary. 2025. HarperCollins Publishers. Available online at https://www.collinsdictionary.com/dictionary/english-word/doctrine, checked on 28 February 2025.

Davies, M. 2010. The Corpus of Historical American English (COHA). Available online at https://www.english-corpora.org/coha/.

DeWolf, D. K., J. G. West, C. Luskin, and J. Witt. 2006. *Traipsing into evolution: Intelligent design and the Kitzmiller v. Dover decision*. Seattle: Center for Science and Culture, Discovery Institute.

Forrest, B., and P. R. Gross. 2004/2007. *Creationism's Trojan horse: The wedge of intelligent design*. New York: Oxford University Press.

Friginal, E., M. Walker, and J. B. Randall. 2022. Exploring mega-corpora: Google Ngram Viewer and the Corpus of Historical American English. *E-JournALL* 1(1): 48–68.

Hunter, G. W. 1914. *A civic biology: Presented in problems*. New York: American Book Company.

Hunter, G. W. 1926. *New civic biology: Presented in problems*. New York: American Book Company.

Hunter, J. D. 1991. *Culture wars: The struggle to define America*. New York: Basic Books.

Jucker, A. H. 2018. Apologies in the history of English: Evidence from the Corpus of Historical American English (COHA). *Corpus Pragmatics* 2(4): 375–398.

Kippenhahn, R., and A. Weigert. 1990. *Stellar structure and evolution*. Berlin: Springer-Verlag.

Larson, E. J. 2003. *Trial and error: The American controversy over creation and evolution*. 3rd ed. Oxford: Oxford University Press.

Larson, E. J. 2006. *Evolution: The remarkable history of a scientific theory*. Modern Library Chronicles. New York: Modern Library.

Ng, R., H. G. Allore, M. Trentalange, J. K. Monin, and B. R. Levy. 2015. Increasing negativity of age stereotypes across 200 years: Evidence from a database of 400 million words. *PLOS One* 10(2): e0117086.

Pechenick, E. A., C. M. Danforth, and P. S. Dodds. 2015. Characterizing the Google Books Corpus: Strong limits to inferences of socio-cultural and linguistic evolution. *PLOS One* 10(10): e0137041.

Person, K. 2019. *What defines a parent?* BA thesis, Stockholms universitet, Stockholm.

Steen, G. J., E. Biernacka, A. G. Dorst, A. Kaal, C. I. López Rodríguez, and T. Pasma. 2010. Pragglejaz in practice. In G. Low, Z. Todd, A. Deignan, and L. Cameron (eds.), *Researching and applying metaphor in the real world*, 165–184. Amsterdam: John Benjamins Publishing Company.

Webster, N. 1828. *American dictionary of the English language*. Available online at https://webstersdictionary1828.com/Dictionary/doctrine, checked on 28 February 2025.

Whitcomb, J. C., and H. M. Morris. 1961. *The Genesis Flood: The biblical record and its scientific implications*. Philadelphia: Presbyterian and Reformed Publishing Co.

White, A. D. 1896/1993. *A history of the warfare of science with theology in Christendom*. Two volumes in one. Buffalo, NY: Prometheus Books.

Cite this article as:

Barczewska, S. (2025). Evolution in the Use of *Evolution*? An Overview of the Term in the Corpus of Historical American English. *LingBaW. Linguistics Beyond and Within, 11*, 7–22.