

## I. Introduction

It seems time again I found myself being asked, mostly by friends and family (who not understanding what exactly it is my studies entail, mistakenly believe me to be an expert on prescriptive English grammar) what the difference is between "toward" and "towards" and which form is grammatical (or more grammatical in a given situation). A quick search allowed me to readily verify that both forms were acceptable, at least from a descriptive point of view, but the questions of what distinguished the two forms in use, and where the distinction originates, were always beyond my grasp. This situation continued for many years with the question unresolved, before the accumulated curiosity and social necessity of finding the answer ultimately resulted in the study now before you.

Due to the size restrictions on this study, however, it is only possible to answer certain aspects of this question. A full examination of the collocations of each form of TOWARD, for example, will not be attempted here, nor will any attempt be made at identifying a pattern whereby one might predict the use of *toward* or *towards* in any particular context. Rather, this study will examine the variance between different registers of the two word forms, as well as the general usage trends for the two best-studied English varieties, namely British and American English. Following this, some work will be undertaken to identify the history, in as far as it is possible, in the usage of each form, in order to get a clearer picture of how they have come to reach the distribution we see today.

## II. Corpora and Methodology

Due to the rather extensive scope of this study and the broad comparisons being made, it becomes necessary to employ four different corpora.

For our analysis of contemporary American English, we turn to the Corpus of Contemporary American English (COCA). The COCA is a 425 million word corpus, containing both spoken and written texts. The material within dates from 1990 to 2011, and the texts provided come exclusively from American English sources. Given its composition, it is easily the least problematic corpus employed in the study. Data from this corpus is compared internally, examining variation within American English across different registers, as well as being compared to data obtained from the British English data, as appropriate.

The data employed for the analysis of the British English uses of *TOWARD* comes from the British National Corpus (BNC). The BNC is a 100 million word corpus, smaller than the COCA. Like the COCA, it also contains a variety of spoken and written texts, but the material within dates from the 1980's up to 1993, so it is not perfectly contemporary with the data obtained from the COCA, and this difference, small though it may be, must be taken into account in considering the reliability of any results. Particularly small differences cannot be considered as significant as they otherwise might be, as small variations may be the result of the different time frames of the two corpora, rather than simply as a result of the different dialects. Additionally, while the data obtained from the BNC is primarily focused on British English, it is not exclusively so. Some data exists from American English and other sources, mingled with the other data of the BNC. While this can be problematic, it should be recognized that these differences are still relatively small. A search of written texts in the BNC, for example, limited to sources from the UK and Ireland, for the form *towards*, yielded 11,355 hits, while restricting the search to written texts originating in other Commonwealth countries, continental Europe, USA and elsewhere (all other categories) combined returned only 496 hits. No attempt, however, has been made to eliminate non-British text examples from the data, since most of the written texts in the BNC have not been tagged in this way at all. This distinction must still be accounted for before any conclusions can be successfully drawn from the results. As with the COCA, data from the BNC will be compared internally, examining the uses of each word form across different registers, as well as being

compared with the data obtained from COCA, to compare the variations in use between the two dialects. The categories used to sort the data different registers, and the number of registers identified, vary between the two corpora, but no attempt will be made to compare non-equivalent registers.

For an examination of historical American English use, data will be taken from the Corpus of Historical American English (COHA). The COHA is a 400 million word corpus, comparable in size to the COCA. It contains exclusively written texts and spans a period from 1810 to 2009. While the texts come from purely American sources, that cannot be considered equivalent to being written in American English, and some small portion of the corpus is made up of texts written in other dialects, even in Scots, in a few examples which can be found. The use of American English, however, clearly makes up an overwhelming majority of the texts, so any strong tendencies found within the data can still be said to be representative of American English as a whole. Data from the corpus will be compared internally, in order to conduct a diachronic analysis of the uses of the two forms within the American context, and the results of that analysis will be compared to the results of an analysis of historical data for British English, in order to get a comparison of the two dialects in a historical perspective, that thereby more thorough and convincing arguments can be made about the history and evolution of the usage of each form.

The historical data for British English will be taken from the Old Bailey Corpus (OBC). Easily the most problematic corpus employed, the OBC contains text documenting 197,745 criminal trials held at London's central criminal court. Because the interface and corpus design having been intended for very different types of study, much relevant data is not immediately available from the corpus, such as the total word count of the corpus itself, or even a words per million count for the data. Indeed, searches generally only produce the total number of hits for any given word, and the distribution of texts across time periods is uneven, sometimes drastically so, making direct diachronic analysis of any single word use virtually impossible. For this study, however, it has been possible to compare the ratio of one word form with another (e.g. total *towards* tokens/total *toward* tokens), rather than using the raw

count of occurrences, providing data which may be compared regardless of the imbalances in text quantities within different time frames. For added security, samples from the beginning and end of the corpus time-line have been omitted, due to their small selection of texts, producing results which are less reliable from the insufficient sample size. Care has been taken not to compare incompatible types of data, such as words per million data with word form ratios. Because the texts come only from the London criminal court system, it can be assured with a fair level of certainty that all the information gathered is representative of British English, albeit a more specific variety than that gained from the BNC. Data gained may be considered representative of historical London English specifically, more so than historical British English as a whole, but one may be assured of a relative freedom from non-British material. Data used from the OBC for this study is limited to written texts taken from 1710 to 1900. Data gained from this corpus will be compared internally, in order to perform a diachronic analysis of the use of each word form across the time period within the associated dialect. Results gained from that analysis will be compared to results gained from analysis of the data from the COHA, in order to get a more far-reaching historical perspective on the evolution of the use of each word form.

### **III. Data and Analysis**

The choice of where to begin such a discussion is essentially arbitrary. One could as easily begin in the past, then trace the patterns of usage toward<sup>1</sup> the present, following the natural course of history, or just as easily begin in the present, then take a look back to discover the way that current state of affairs came into being. Given, however, the relative sizes and levels of reliability of each corpus I have deemed it most prudent to begin where my data may be presented with the least qualifiers, so the analysis will first be focused on modern American English usage of the two word forms.

---

<sup>1</sup> My use of the form *toward* in this case is, as will be demonstrated, consistent with the American English usage for academic prose, in keeping with the majority of my wording and punctuation choices for this work.

Figure 3.1

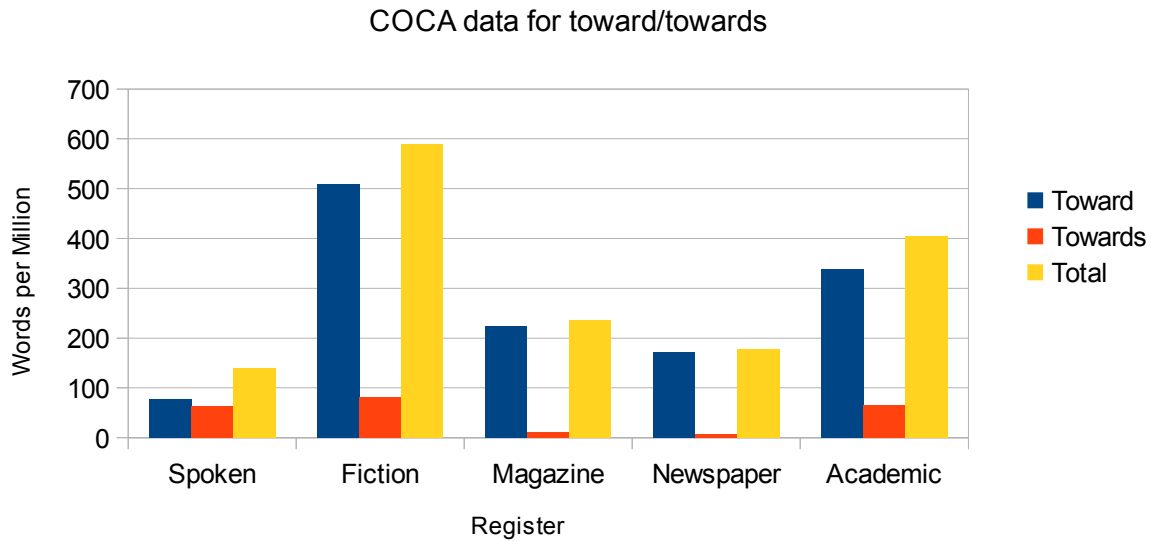


Figure 3.2

COCA: *toward*

SECTION	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
FREQUENCY	6887	43241	20302	14801	28993
PER MILLION	76.47	508.94	224.85	170.77	337.95

COCA: *towards*

SECTION	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC
FREQUENCY	5643	6887	1040	567	5635
PER MILLION	62.65	81.06	11.52	6.54	65.68

As can be clearly seen from figures 3.1 and 3.2, overall uses of *TOWARD* vary greatly between registers, occurring most frequently in fiction, where the narration of a story often necessitates the use of the preposition to describe the action. To a lesser extent, academic prose makes use of the word in describing conceptual progressions and general shifts in data or ideas. The other registers do not use the word as frequently. In general, however, it can easily be seen that American English strongly favors the use of the form *toward*, dropping rather than adding or retaining the /s/, but this tendency is far stronger in written registers than in spoken ones. In the spoken registers, use of *toward* is only 22% more frequent than the use of *towards*, a stark contrast to newspaper writing, where the use of *toward* is

2,511% more frequent. This distinction may result from the differing modes of production in the different registers. Although rarely prescribed as more grammatical, when given time to stop and consider their wording more carefully, Americans seem to prefer to employ the *toward* form almost to exclusion of its counterpart, yet the near-equivalent usage in the spoken register suggests that, although a preference may exist, both forms do, in general, arise quite naturally and intuitively. This is quite unlike the equivalent data obtained from the BNC.

Figure 3.3

BNC data for toward/towards

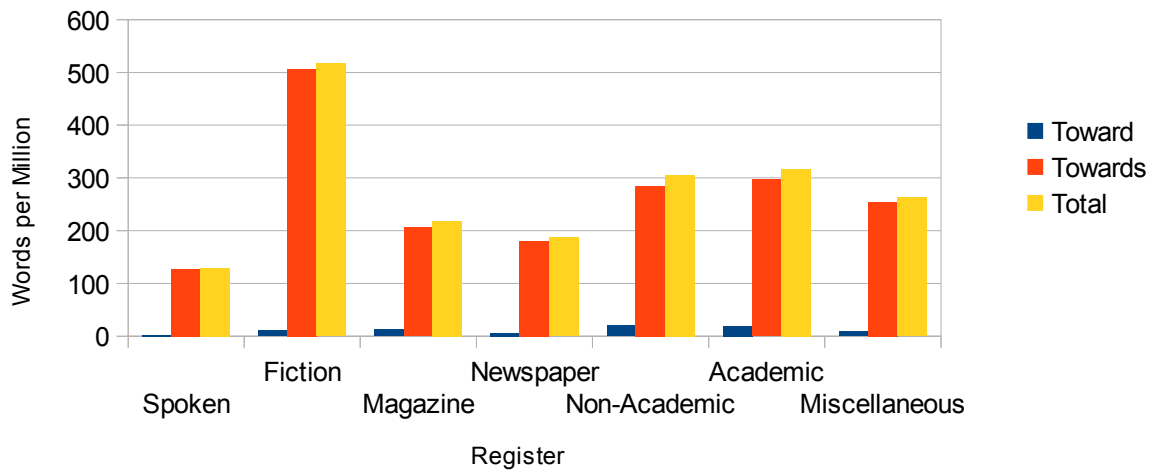


Figure 3.4

**BNC: *toward***

SECTION	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACADEMIC	ACADEMIC	MISC
FREQUENCY	19	174	94	60	336	304	178
PER MILLION	1.91	10.94	12.94	5.73	20.37	19.83	8.54

**BNC: *towards***

SECTION	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACADEMIC	ACADEMIC	MISC
FREQUENCY	1270	8052	1493	1895	4700	4553	5304
PER MILLION	127.46	506.12	205.59	181.06	284.93	296.97	254.57

Figure 3.3 and 3.4 show the equivalent data for British English use obtained from the BNC. As

before, a preference for the use of *TOWARD* in fiction and to a lesser extent in academic and non-academic prose can be seen, but unlike in the data from the COCA, the use of *toward* and *towards* in the spoken register are far from approximate to one another. In fact, where American English favored the use of *toward*, British English favors the use of *towards*, only far more strongly so. While some registers, such as the use in newspaper writing (curiously one of the most imbalanced registers in American English), are almost perfectly inverted (having the same level of contrast but the relative values reversed), for most registers, the ratio of one form to the other marks an even more striking contrast in British English. This is, of course, most noticeable in the spoken register, where American English is only 22% higher in its use of *toward*, while British English favors the use of *towards* by 6,573% (a difference of over 125 words per million, rather than a difference of less than 15), but even registers like fiction must compare a 528% increase from the less favored form to the more favored one in the COCA with a 4,526% increase in the BNC. There are numerous possible explanations for why this might be. Following the given hypothesis for the difference between the spoken and written registers in the COCA, one likely possibility is that the dialect difference stems from a stronger tradition of prescriptivism in British English than exists in American English, even in speech and pronunciation (where British English retains a standardized dialect and American English does not), thus causing speakers of British English to apply more time and concentration for their choice of word forms in speech, creating a similar effect in the spoken register to the one seen in the American written registers. Meanwhile, the stronger prescriptive tradition reflects on the written registers by heightening the differences in frequency between the two word forms as writers are increasingly careful about which form they choose to use in any given situation.

If *toward* may be considered a predominantly American form then and *towards* a predominantly British one, this begs a number of questions. How, for example, did the current distribution come about? Which was the original, and who made the innovation? Is the current mix of two forms in each corpus a case of cross-contamination of dialects? Does it stem from errors in control of the corpora

contents? Or is it somehow the remnant of an older, more complex tradition of usage? To answer some of these questions, we turn to the COHA, in order to examine the historical use of these forms in the United States.

Figure 3.5

COHA data for toward/towards

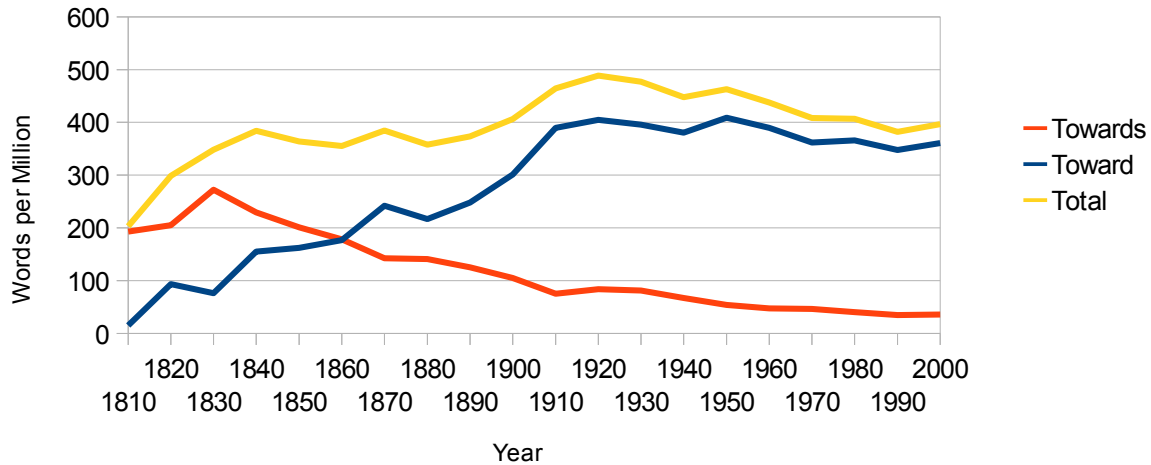


Figure 3.6

**COHA: *toward***

Year	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
Freq	18	646	1048	2487	2675	3019	4495	4406	5116	6658	8843	1034	9731	9269	1003	9347	8622	9265	9708	1066
Per Mil.	15.24	93.26	76.08	154.97	162.4	177.02	242.16	216.8	248.3	301.3	389.5	404.7	395.5	380.6	408.8	389.8	362.0	365.9	347.4	360.7
				7		2	6	7	4		5	8	3	9	8	3	4	7	4	7

**COHA: *towards***

Year	1810	1820	1830	1840	1850	1860	1870	1880	1890	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990	2000
Freq	228	1422	3749	3678	3315	3038	2644	2861	2581	2322	1704	2148	2006	1636	1323	1141	1104	1028	970	1060
Per Mil.	193.0	205.2	272.1	229.1	201.2	178.1	142.4	140.8	125.2	105.0	75.06	83.73	81.54	67.19	53.9	47.59	46.36	40.61	34.72	35.85
	2	8	7	8	5	3	4	3	9	8										

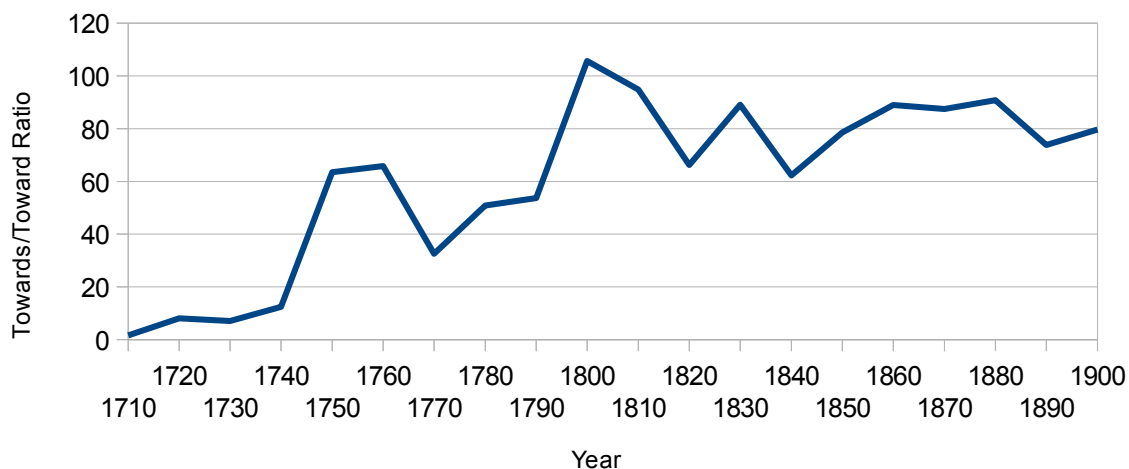
Figures 3.5 and 3.6 show data from the COHA, tracking different forms of TOWARD from 1810 to 2000. After the total uses climb up the chart during the first thirty years, there is, in general, relatively little change in the overall, combined uses of the forms during the time frame represented in the COHA. The use of each individual form, however, undergoes a much more drastic transformation.



Starting in 1810 as a relatively rare form (though not as rare as in the more recent data on British English from the BNC) *toward* grows quickly in frequency over the next century, arriving in the 347-409 words per million range by 1910, where it remains comfortably for the rest of the recorded time line. Meanwhile, *towards* begins as the dominant form and grows in use for twenty years (at least in total word count), up until 1830, at which point it enters a steady decline all the way until 2000, where it achieves a position comparable to the equivalent data found in the COCA. This shows a clear historical transformation in the American context. Initially, as in the later British context, *towards* had been the favored form, but this changed during the 19<sup>th</sup> century, as a gradual shift in usage resulted in an increasing preference for *toward*. Given that the United States had only declared its independence from Great Britain barely thirty four years before the data in the corpus begins, after a period of colonization that had begun at Jamestown, Virginia in 1607 (just over a century and a half prior), it is quite easy to speculate, based on this data, that *towards* was the original form and that the form *toward* was simply an American innovation, originating there and growing in popularity until it began to spill over in small ways back into British English as the United States became a cultural superpower during the 20<sup>th</sup> century. This conclusion, however, is not supported by data from the OBC.

Figure 3.7

OBC data for toward/towards



Some caution should be taken in examining figure 3.7. In particular, it need be noted that the y-axis of the graph indicates the ratio of uses of *towards* to uses of *toward*. For reasons specified in section II, this data must be used here in place of words per million data. Because of this difference, a low mark on this graph does not indicate the number of occurrences of *toward* is approaching zero. Rather, as the line on the graph drops toward<sup>2</sup> one, the number of incidents of *toward* and *towards* in the data approaches equivalency, and as the line climbs higher on the chart, the total instances of the use of *towards* take an increasingly clear majority. Starting in 1710 the ratio was nearly equal, at 1.6, so for every 10 instances of *toward* in the data, there were only 16 instances of *towards*, placing the latter at a majority, but not an overwhelming one. This clearly refutes the prior hypothesis by demonstrating that even as early as 1710, the form *toward* already represented a significant portion of the data in at least one British context, although a shift away from it was already in progress. In fact, by 1740, thirty years later, the instances of *towards* outnumbered instances of *toward* with more than a 12 to 1 ratio, this climbing up to pass a 100 to 1 ratio around 1800, just before the American data begins in the COHA. Being a ratio, rather than a words per million count and using a smaller sample size, the level varies more than it would if only a single word form were tracked, but the data continues within a comparatively set range throughout the 19<sup>th</sup> century. It's worth noting that this ratio is actually higher than the ratio observed in the BNC, indicating an even stronger final preference for *towards* over *toward* than is seen in the more modern data. There is insufficient data to conclude as to whether this is due more to the highly specific choice of dialect/register found in the OBC or whether it is a result of later changes in British English use, perhaps as a result of increased contact with American English. The difference is, however, too large to be accounted for by other data from other dialects being blended in with the BNC data.

## IV. Conclusions

What conclusions, then, can be drawn from this analysis in the end? Several are clear. Firstly,

---

<sup>2</sup> See previous footnote.

the use of *toward* is most strongly favored in the American English context, particularly in written registers, with a weaker, but still noteworthy, favoritism being found in the spoken register. In British English, the *towards* form is preferred, even to a greater extent than the American preference. At some point, in British English, the use of each form was more or less equal, at least in certain registers, but that quickly changed to a very strong favoritism to *towards*. About the time this shift was reaching its conclusion in Britain, English in the United States began a shift in the opposite direction, which transpired over the following century, resulting in the largely opposite preferences seen today.

Further study remains possible, despite this compelling conclusions, with many questions left unanswered. How did the use of TOWARD vary in earlier times? Were there ever distinctions in meaning? What factors influence speaker decisions to occasionally employ their less favored forms still today? Is there a prototypical form of each, or is the distribution haphazard? How do other major English dialects compare? And, of course, the question of the future remains ever open. How will this use continue to evolve? Will changes in meaning develop? As time passes and more data becomes available, followup studies could be pursued, in order to track these ongoing changes. Similarly, studies of similar words could be undertaken. TOWARD is, after all, just one word of a much broader type. The data here is fascinating, but it begs for further investigation.

And so we find ourselves, along with our language, compelled to move ever onward(s).