

On Must be+ing, May be+ing, and Might be+ing

 Namkil Kang¹

¹College of Liberal Arts, Far East University, South Korea

*Corresponding author. Email: somerville@hanmail.net

Abstract

The main purpose of this article is to provide a detailed analysis of *must be+ing*, *may be+ing*, and *might be+ing* within the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), the Corpus of Historical American English (COHA), and the Time Magazine Corpus (TMC). The COCA clearly shows that the type *may be+ing* is the most preferred by Americans. The BNC clearly shows, on the other hand, that *may be+ing* is the type most preferred by British people, followed by *must be+ing*, and *might be+ing*. A major point to note is that *must be+ing* is the most commonly used in the fiction genre in the UK, whereas *might be+ing* is the most widely used in America. A further point to note is that *may be+ing* is the most widely used of these three modal verb types in the spoken genre in America, whereas *might be+ing* is the most frequently used in the UK. This in turn suggests that Americans are fond of using statements with moderate possibility when they talk, whereas British people are fond of using statements with low possibility. With respect to the magazine genre, *may be+ing* is the type most preferred by British and American writers. It is significant to note, however, that British journalists prefer statements with moderate possibility or necessity to statements with low possibility. However, American journalists prefer statements with moderate or low possibility to statements with necessity. With respect to the academic genre, it is worth noting that *may be+ing* is the type most preferred by both British people and Americans. The reason why *may be+ing* is the most widely used in America and the UK may be that *may be+ing*, with its moderate possibility, is the most suitable for conveying knowledge, academic information, and actual facts. Additionally, the COHA clearly shows that *might be+ing* was the type most preferred by Americans from 1810 to 2000, followed by *may be+ing*, and *must be+ing*. Finally, the TMC clearly indicates that *may be+ing* was the type most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*. That *may be+ing* was the type most preferred by educated Americans suggests that they were fond of using statements with moderate possibility.

Keywords: COCA, BNC, COHA, TMC, Must be+ing, May be+ing, Might be+ing

1. Introduction

It is clear that the three modal verb types *must be+ing*, *may be+ing*, and *might be+ing* move from strong conjecture to weak conjecture (*must be+ing* > *may be+ing* > *might be+ing*). The main purpose of this article is to provide a detailed frequency analysis of the three modal verb types within the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), the Corpus of Historical American English (COHA), and the Time Magazine Corpus (TMC). In section 3.1, we show that *may be+ing* is the one of these three modal verb types most commonly used in America. More specifically, *may be+ing* is most preferred by Americans, followed by *might be+ing*, and *must be+ing*, in that order. In section 3.2, we argue that *may be+ing* is the type most preferred by British people, followed by *must be+ing*, and *might be+ing*. Also, we argue that *must be+ing* is the most commonly used of these three modal verb types in the fiction genre in the UK, whereas *might be+ing* is the most widely used in America. We further argue that *may be+ing* is the most widely used in the spoken genre in America, whereas *might be+ing* is the most frequently used in the UK. This in turn suggests that Americans are fond of using statements with moderate possibility when they talk, whereas British people are fond of using statements with low possibility. We contend, on the other hand, that *may be+ing* is the type most preferred by British and American writers in the magazine genre. We maintain, however, that British journalists widely use statements with moderate possibility or necessity, whereas American journalists widely use statements with moderate or low possibility. In addition, we show that *may be+ing* is most preferred by British people and Americans in the

academic genre. The reason why *may be+ing* is the most widely used of these three modal verb types in America and the UK may be that *may be+ing*, with its moderate possibility, is the most suitable for conveying knowledge, academic information, and actual facts. In section 3.3, we contend that *might be+ing* was the type most preferred by Americans from 1810 to 2000, followed by *may be+ing*, and *must be+ing*. The key difference between the frequency in the COCA and in the COHA is that *might be+ing* was the modal verb type most commonly used in America from 1810 to 1989, whereas *may be+ing* was the one most frequently used from 1990 to 2019. In section 3.4, we maintain that *may be+ing* was most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*. That *may be+ing* was the type most preferred by educated Americans suggests that they were fond of using statements with moderate possibility from the 1920s to the 2000s. It should be noted, however, that *must be+ing* has the lowest frequency and the lowest proportion of the three types. From this, it can be inferred that educated Americans tended to avoid statements with necessity.

2. Methodology

The main goal of this article is to provide an in-depth frequency analysis of the three modal verb types (*must be+ing*, *may be+ing*, *might be+ing*) within the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), the Corpus of Historical American English (COHA), and the Time Magazine Corpus (TMC). The following modal verb types move from strong conjecture to weak conjecture, as illustrated in (1):

(1) *must be+ing* > *may be+ing* > *might be+ing*

In this article, we aim to answer the following main questions: Which of these three modal verb types is the most commonly used in America and the UK? Which is most preferred by speakers or writers in America and the UK? Which is the most widely used in the five genres of the COCA? Which is the most frequently used in the seven genres of the BNC? What is the main difference between frequency in the COCA and in the BNC? What does the frequency in the COHA and TMC indicate?

3. Results and Discussion

3.1. The Corpus of Contemporary American English and the Frequency of the Three Modal Verb Types

In the following, we provide an in-depth analysis of the use and frequency of *must be+ing*, *may be+ing*, and *might be+ing* in the COCA (1990-2019):

Table 1. Frequency of the Three Types in the COCA

Type	Spoken	Fiction	Magazine	Newspaper	Academic	Total
<i>must be+ing</i>	503	1443	408	334	133	2,821
<i>may be+ing</i>	3247	433	2613	2193	1749	10,235
<i>might be+ing</i>	1866	1941	1276	1184	471	6,738

An important question to be asked is “which type of modal verb is most frequently used in America?” Table 1 clearly shows that *may be+ing* is the most widely used of these three modal verb types in America. As alluded to in Table 1, the preferable type for Americans is *may be+ing* since it has the highest frequency (10,235 tokens) and the highest proportion (51.74%) among the three modal verb types. Then, why is *may*

be+ing the one most commonly used in America? We wish to argue that *must be+ing* carries strong conjecture, so it conveys necessity, while *might be+ing* carries weak conjecture, so it conveys low possibility. We wish to argue, on the other hand, that *may be+ing* carries moderate conjecture, so it conveys moderate possibility. That may be why Americans prefer *may be+ing* to *must be+ing* and *might be+ing*. In addition, the reason why *must be+ing* is used less frequently than *may be+ing* is that *must be+ing* is semantically marked. It is evident that a type that is semantically marked is used less frequently than a type that is not semantically marked. Thus, Americans may prefer *may be+ing* with its moderate possibility to *must be+ing* with its necessity. Another reason why *must be+ing* is used less frequently than *may be+ing* is that speakers or writers can use *must* only when they have evidence of some kind (Palmer 1987: 122). Let us observe the following sentences:

(2) a. John may be in his office.

b. John must be in his office.

(Palmer 1987: 97)

As pointed out by Palmer (1987: 97), *may* can be interpreted in terms of some kind of possibility, whereas *must* expresses some kind of necessity. (2b) makes a judgment that John is bound to be in his office. Palmer (1987: 122) points out that “*must* essentially makes a conclusive judgment, usually from evidence of some kind” (Palmer 1987: 122). Thus, it is clear that the modal verb type *must be+ing* cannot be used without any evidence. This is why Americans prefer using *may be+ing*, with its moderate possibility, to *must be+ing*, with its necessity.

Now, a question to be asked is “which of the three modal verb types is most commonly used in the spoken genre?” Table 1 clearly indicates that of the three types, *may be+ing* is the one most widely used in the spoken genre. Why is *may be+ing* the one most frequently used in the spoken genre? It is more common for people to talk with moderate possibility than with necessity. Table 1 clearly shows that Americans do not prefer *must be+ing*, since it carries some kind of necessity. A further point to note is that Americans prefer *might be+ing* to *must be+ing*. This may imply that Americans like speaking to people not with necessity but with low possibility. As pointed out by Palmer (1987: 119), “*might* and *could* are used as tentative forms of epistemic *may* in all its possible environments, to express a lower degree of possibility” (Palmer 1987: 119).

Now an important question to be asked is “which of the three modal verb types is the one most widely used in the fiction genre?” Table 1 indicates that *might be+ing* is the one most frequently used in the fiction genre. More specifically, *might be+ing* has the highest frequency (1941 tokens) and the highest proportion (50.85%) among the three types. This suggests that Americans prefer using weak conjecture to using strong conjecture or moderate conjecture in the fiction genre. Then, why do Americans prefer *might be+ing* to *must be+ing* and *may be+ing* in the fiction genre? The reason why *might be+ing* is the type most preferred by Americans in the fiction genre may be that to use *might be +ing* with its low possibility in fiction can excite readers’ curiosity. In addition, writers can develop stories by giving clues about the hidden themes. Thus, *must be+ing*, with its necessity, can be used less frequently in the fiction genre than *might be+ing*, with its low possibility.

Now, let us consider the magazine genre and the newspaper genre. An immediate question to be asked is “which modal verb type in these genres is the type most preferred by Americans?” Table 1 clearly shows that *may be+ing* is the one most widely used in these genres. More specifically, of the three types *may be+ing* has the highest frequency (2613 tokens) and the highest proportion (60.80%) in the magazine genre,

and it has the highest frequency (2193 tokens) and the highest proportion (59.09%) in the newspaper genre. So, why does *may be+ing* have the highest frequency in these two genres? Note that *may be+ing*, with its moderate possibility, is neither a strong statement nor a weak statement. Thus, the reason why *may be+ing* obtains the highest frequency in these genres is that *may be+ing*, with its moderate possibility, may be suitable for conveying factual information.

Let us turn our attention to the academic genre. Which type of modal verb is the one most commonly used in America? Table 1 indicates that of the three types, *may be+ing* is the type most preferred by Americans in the academic genre. *May be+ing* has the highest frequency (1749 tokens) in the academic genre, while *must be+ing* has the lowest frequency (133 tokens). Why does this happen? We wish to argue that *may be+ing* is the most suitable among the three types for conveying factual information, since it carries moderate possibility. Exactly the same thing can be said about the magazine genre and the newspaper genre. *May be+ing* is the most preferred among the three types by American writers since it maintains a neutral position. Just as in the magazine genre and the newspaper genre, *may be+ing* is the most preferred among the three types by Americans in the academic genre since it is neutral and thus is suitable for conveying knowledge, academic information, and actual facts.

Consider the ranking of the type *must be+ing* in the COCA. As indicated in Table 1, *must be+ing* has the highest frequency (1443 tokens) and the highest proportion (51.15%) in the fiction genre. On the other hand, it has the lowest frequency (133 tokens) and the lowest proportion (4.71%) in the academic genre. This indicates that *must be+ing* is the modal verb type most preferred by American writers in the fiction genre, whereas it is not preferred in the academic genre. Why is *must be+ing* the type most preferred by American writers in the fiction genre? We wish to argue that American writers use *must be+ing*, with its necessity, in order to give their fiction a dynamic impact. Also, it must be noted that writers often give clues about the hidden themes by using statements with low possibility. Thus, it is reasonable to conclude that both *might be+ing* with its low possibility and *must be+ing* with its necessity are widely used by Americans in the fiction genre.

Now, let us consider *must be+ing* in the academic genre. As indicated in Table 1, *must be+ing* has the lowest frequency (133 tokens) and the lowest proportion (4.71%) in the academic genre. Why does *must be+ing* have its lowest frequency in the academic genre? This may be due to the fact that the progressive combined with *must* cannot be widely used in the academic genre. The progressive essentially indicates dynamic situations (Quick & Greenbaum 1993, Saeed 2009). In addition, “the progressive indicates a temporary state, whereas the simple present indicates a permanent state” (Kang 2018). The academic genre can be characterized as a knowledge database, and knowledge in the database has a permanent property, but the progressive indicates a temporary state, which leads to the infrequency of *must be+ing* in the academic genre. Furthermore, *must* can be used only when speakers or writers have evidence of some kind. Koster (1997: 244) argues that a form A is more specialized than B if A can fulfill fewer functions than B. As such, *must be+ing* is a specialized form since it can be used only when speakers have evidence of some kind. Simply put, *must be+ing* fulfills fewer functions than *may be+ing* or *might be+ing* in the academic genre. Thus, it is reasonable to conclude that for these reasons, *must be+ing* has its lowest frequency in the academic genre.

Now let us observe the frequency of *may be+ing* in the COCA. A major point to note is that *may be+ing* has its highest frequency (3247 tokens) and highest proportion (31.72%) in the spoken genre. Why does *may be+ing* have the highest frequency in the spoken genre? This may be due to the fact that *may be+ing* carries moderate possibility. *Must be+ing*, with its necessity, is semantically marked to be used in daily

conversation, so *may be+ing* or *might be+ing* may be preferred over *must be+ing* in the spoken genre. It should be noted that *may be+ing* is the type most preferred by Americans, followed by *might be+ing*, and *must be+ing*, in that order.

Now let us turn our attention to the status of *may be+ing* in the fiction genre in the COCA. Table 1 clearly indicates that *may be+ing* has the lowest frequency (433 tokens) and the lowest proportion (4.23%) in the fiction genre. Why does *may be+ing* have its lowest frequency in the fiction genre? As observed earlier, *might be+ing* is most preferred in the fiction genre by Americans, followed by *must be+ing*, and *may be+ing*, in that order. The reason why *may be+ing* has its lowest frequency in the fiction genre may be that *may be+ing* carries moderate possibility. It is evident that writers encourage readers to read a book with excitement. In order to excite readers, writers can use inclusive strategies for readers by using *might be+ing* with its low possibility or *must be+ing* with its necessity. However, half possibility (*may be+ing*) would neither astonish nor excite readers, since *may be+ing* carries moderate possibility.

Now, let us consider *might be+ing* in the fiction genre in the COCA. Table 1 clearly shows that *might be+ing* has its highest frequency (1941 tokens) and highest proportion (28.80%) in the fiction genre. An important question to be asked is “why does *might be+ing* obtain its highest frequency in the fiction genre?” We wish to argue that fiction can become more emotional, intense, and vivid if the progressive is used. Cowan (2010: 363) argues that “the progressive expresses much more emotion than the simple present” (Cowan 2010: 363).

(3) a. This operation is really costing a lot of money.

b. This operation costs a lot of money.

(Cowan 2010: 363)

In addition, as observed earlier, by using statements with low possibility, writers can excite readers' curiosity.

It is worth noting that *might be+ing* has its lowest frequency (471 tokens) and lowest proportion (6.99%) in the academic genre. This may be due to the fact that *might be+ing* carries low possibility. Academic fields require statements with moderate possibility or necessity since they deal with knowledge, scientific, technological, and academic information, and actual facts. Thus, it is reasonable to conclude that for this reason, *might be+ing* has its lowest frequency in the academic genre.

3.2. The Frequency of the Three Types of Modal Verbs in the British National Corpus

In the following, our discussion will center around frequency in the BNC and in the COCA. We will consider the BNC and COCA in order to examine the national variation of the three modal verb types. Table 2 shows the genre frequency of the three types of modal verbs:

Table 2. Genre Frequency of the Three Types in the British National Corpus

Type	Spoken	Fiction	Magazine	Newspaper	Non-academic	Academic	Misc	Total
must be+ing	145	366	48	84	62	46	82	833
may be+ing	86	89	124	137	256	174	330	1196
might be+ing	180	279	32	68	81	68	109	817

As indicated in Table 2, *may be+ing* is the most frequently used of the three types of modal verbs in the UK, followed by *must be+ing*, and *might be+ing*, in that order. More specifically, *may be+ing* has the highest frequency (1196 tokens) and the highest proportion (42.02%), whereas *might be+ing* has the lowest frequency (817 tokens) and the lowest proportion (28.70%). On the other hand, *must be+ing* ranks second among the three modal verb types in the BNC. This suggests that *may be+ing* is the type most preferred by British people, followed by *must be+ing*, and *might be+ing*. It is worth noting that British people's preference regarding the three types of modal verbs is different from Americans'. *May be+ing* is the type most preferred by Americans and British people, but British people's second preference is different from Americans'. Americans prefer using *might be+ing* to using *must be+ing*, whereas British people prefer using *must be+ing* to using *might be+ing*. This in turn suggests that Americans prefer using weak statements with low possibility to using strong statements with necessity, whereas British people prefer using strong statements with necessity to using weak statements with low possibility. Thus, it is reasonable to assume that British people tend to make conclusive judgments from evidence of some kind, whereas Americans do this less frequently.

Now, let us consider the three types of modal verb in the spoken genre, comparing the frequency in the BNC and in the COCA. It is noteworthy that *might be+ing* is the most widely used modal verb type in the spoken genre in the UK, whereas *may be+ing* is the most frequently used in America. What does this mean? We take this as indicating that British people are fond of using statements with low possibility when they talk, whereas Americans are fond of using statements with moderate possibility. It must be noted that *might be+ing* is the most commonly used modal verb type in the spoken genre in the UK, followed by *must be+ing*, and *may be+ing*, whereas in America it is *may be+ing*, followed by *might be+ing*, and *must be+ing*. It is worth emphasizing that *may be+ing* is used less frequently than *must be+ing* in the UK, whereas *must be+ing* is used less frequently than *might be+ing* in America. This suggests that British people do not prefer statements with moderate possibility, whereas Americans do not prefer those with necessity. We thus conclude that *might be+ing* with its low possibility is the type most preferred by British people in the spoken genre, whereas *may be+ing* with its moderate possibility is the type most preferred by Americans.

Now let us examine the usage of the three types in the fiction genre, in the BNC. It is interesting to note that *must be+ing* is the most widely used of these modal verb types in the fiction genre. More specifically, *must be+ing* is the type most preferred in the fiction genre by British people, followed by *might be+ing*, and *may be+ing*. On the other hand, among Americans *might be+ing* is the type most preferred, followed by *must be+ing*, and *may be+ing*. This suggests that British writers are fond of using statements with necessity in their fiction, whereas American writers are fond of using statements with low possibility in their fiction. This may imply that British writers are fond of using a strong and decisive style of writing, whereas American writers are fond of using an allusive style of writing.

Looking at the use of the three types in the magazine genre in the BNC, it seems significant that *may be+ing* is the type most frequently used in the UK. More specifically, *may be+ing* is the type most preferred by British people, followed by *must be+ing*, and *might be+ing*. This indicates that British writers' preferences toward the three types in the magazine genre is slightly different from American writers'. *May be+ing* is the type most preferred by British and American writers. However, British writers prefer using *must be+ing* to using *might be+ing* in the magazine genre, whereas American writers prefer using *might be+ing* to using *must be+ing*. The reason why *may be+ing* is the type most preferred by British and American writers may be that statements with moderate possibility are the most suitable for articles, ads, or information. That *must be+ing* is preferred over *might be+ing* in the UK indicates that in the magazine genre, statements with

necessity are more preferred than statements with low possibility. On the other hand, that *might be+ing* is preferred over *must be+ing* in America indicates that statements with low possibility are more preferred than statements with necessity in the magazine genre.

Now let us examine the frequency of the modal verb types in the newspaper genre in the BNC. It is worth pointing out that *may be+ing* is the type most preferred by British journalists, followed by *must be+ing*, and then *might be+ing*. This pattern is the same as the pattern of the magazine genre in the BNC. This indicates that British journalists are fond of using statements with moderate possibility in magazines and newspapers. Exactly the same can be said of the COCA. *May be+ing* is the type of modal verb most widely used in these genres. As alluded to in Table 2, however, British journalists are fond of using statements with necessity rather than statements with low possibility. On the other hand, American journalists like using statements with low possibility rather than statements with necessity in the newspaper genre. This suggests that British journalists prefer statements with moderate possibility or statements with necessity to statements with low possibility in the newspaper genre, whereas American journalists prefer statements with moderate possibility or statements with low possibility to statements with necessity. We thus conclude that *may be+ing*, with its moderate possibility, is the type most preferred by British and American journalists.

Now let us consider the three types of modal verbs in the academic genre. Table 2 clearly shows that *may be+ing* has the highest frequency (174 tokens) and the highest proportion (60.41%) in the academic genre. As illustrated in Table 1 and Table 2, *may be+ing* is the type most preferred by British people and Americans in the academic genre. The reason why *may be+ing* is the type most preferred by British people and Americans may be that *may be+ing* is neutral among the three types, and thus is the most suitable for conveying factual information. As observed earlier, the academic genre requires statements with moderate possibility or those with necessity since it includes knowledge, scientific, technological, and academic information, and actual facts. Thus, it is reasonable to assume that for this reason, *may be+ing* is the type most preferred by both British people and Americans.

Now let us turn our attention to the seven genres of *must be+ing* in the BNC. It is worth noting that among the seven genres, *must be+ing* is the type most preferred by British writers in the fiction genre. *Must be+ing* has the highest frequency (366 tokens) in the fiction genre, whereas it has the lowest frequency (46 tokens) in the academic genre. That statements with necessity are the most preferred by British writers in the fiction genre may imply that British writers like using a strong and decisive style of writing. On the other hand, the reason why *must be+ing* has the lowest frequency in the academic genre may be that the progressive combined with *must* indicates a temporary state. Clearly, academic knowledge indicates a permanent state, which leads to the infrequency of *must be+ing* in the academic genre.

Now let us consider the use of *might be+ing* in the fiction genre and the magazine genre, in the BNC. *Might be+ing* has the highest frequency (279 tokens) and the highest proportion (34.14%) in the fiction genre, and it has the lowest frequency (32 tokens) and the lowest proportion (3.91%) in the magazine genre. The reason why *might be+ing* is the type most preferred by British writers in the fiction genre may be that as mentioned in the COCA, by using statements with low possibility, writers can excite readers' curiosity. In addition, writers can give clues about the hidden themes of a story by using statements with low possibility. The reason why *might be+ing* has the lowest frequency in the magazine genre may be that it carries the lowest possibility among the three modal verb types. The COCA and BNC clearly show that *may be+ing* is the type most preferred by British people and Americans in the magazine genre. This may imply that statements with moderate possibility are the most suitable for articles, information, and ads. We thus

conclude that *might be+ing* is the type most preferred by British people and Americans in the fiction genre of the BNC and COCA.

3.3. The Frequency of the Three Types in the COHA

In the following, we examine the frequency of the three types in the COHA (1810-2000). Table 3 shows the use and frequency of *must be+ing*, *may be+ing*, and *might be+ing* for one hundred and ninety years (1810-2000):

Table 3. Frequency of the Three Types in the COHA

Type	must be+ing	may be+ing	might be+ing
1810	7	5	0
1820	15	26	14
1830	44	44	31
1840	53	53	45
1850	58	59	51
1860	65	57	67
1870	74	63	86
1880	111	73	89
1890	127	66	89
1900	153	106	108
1910	264	148	152
1920	194	133	174
1930	206	139	196
1940	214	163	232
1950	193	198	243
1960	213	245	242
1970	209	291	270
1980	223	362	264
1990	256	307	340
2000	189	364	386
Total	2868	2902	3079

An important question to be asked is “which modal verb type was the most frequently used in America from 1810 to 2000?” The COHA clearly indicates that *might be+ing* was the type most preferred by Americans for one hundred and ninety years, followed by *may be+ing*, and *must be+ing*, in that order. On the other hand, the COCA (1990-2019) shows that *may be+ing* was the most commonly used modal verb type in America, followed by *might be+ing*, and *must be+ing*. This indicates that the frequency shown in the COCA is different from that of the COHA. The main difference between the result of the COCA and that of the COHA is that *might be+ing* was the most widely used modal verb type in America from 1810 to 1989, whereas *may be+ing* was the most commonly used modal verb type from 1990 to 2019. We take this as

indicating that statements with low possibility were the most preferred by Americans from 1810 to 1989, whereas statements with moderate possibility were the most preferred by Americans from 1990 to 2019. Note that the COCA and COHA have one thing in common. As indicated in Table 3, *must be+ing* has the lowest frequency among the three types of modal verbs in both COHA (1810-2000) and COCA (1990-2019), which implies that Americans did not prefer statements with necessity from 1810 to 2019.

A major point to note is that as can be seen in Table 3, the frequency of *must be+ing* continued to increase to 257 tokens over the 100 years from 1810 to 1910. A further point to note is that the figure of *must be+ing* reached a peak in the 1910s (264 tokens). In the 1910s, *must be+ing* was the most frequently used type of modal verb in America, followed by *might be+ing*, and *may be+ing*, in that order. This in turn suggests that in the 1910s, Americans preferred using statements with necessity to using statements with low possibility or statements with moderate possibility. After this period, there were slight fluctuations in the frequency of *must be+ing* from 1920 to 2000. It should be noted that in the 2000s, the number of tokens of *must be+ing* is 189, but it ranks third among the three types of modal verbs.

As can be seen from Table 3, the frequency of *may be+ing* continued to increase from 1810 to 1910. After this period, there was a decline in the frequency of *may be+ing* from 1910 to 1920. After the 1920s, there was a dramatic increase in the frequency of *may be+ing*, and the frequency of *may be+ing* reached a peak in the 2000s (364 tokens). In this period, *may be+ing* is the most widely used one in America, which we take as indicating that *may be+ing* was the type most preferred by Americans.

Now let us examine the frequency of *might be+ing* in the COHA. A point to note is that the frequency of *might be+ing* continued to increase to 270 tokens from 1810 to 1970. After this period, there was a decline in its frequency. A further point to note is that the frequency of *might be+ing* reached a peak in the 2000s (386 tokens). In this period, *might be+ing* was the most commonly used of the modal verb types in America, followed by *may be+ing*, and *must be+ing*, which in turn suggests that *might be+ing* was the type most preferred by Americans in the 2000s. Thus, it can be concluded that in the 2000s, statements with low possibility were the most preferred by Americans among the three types.

3.4. The Frequency of the Three Types in the Time Magazine Corpus

In this section, we examine the frequency of the three types in the TMC. The following table shows the use and frequency of the three types of modal verbs from the 1920s to the 2000s.

Table 4. The Frequency of the Three Types in the Time Magazine Corpus

Type	must be+ing	may be+ing	might be+ing
1920s	12	25	22
1930s	22	36	31
1940s	33	90	117
1950s	40	141	113
1960s	41	179	93
1970s	42	276	100
1980s	29	314	85
1990s	39	337	80
2000s	26	263	62

Total	284	1661	703
-------	-----	------	-----

An important question that must be addressed is “which modal verb type is the most frequently used among the three types?” Table 4 clearly illustrates that *may be+ing* is the one most commonly used among the three types. This indicates that *may be+ing* was the type most preferred by educated Americans from the 1920s to the 2000s. Table 4 implies that among the three types, *may be+ing* was the type most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*, in that order. That *may be+ing* was the type most preferred by educated Americans indicates that they were fond of using statements with moderate possibility from the 1920s to the 2000s. It must be noted, however, that *must be+ing* has the lowest frequency (284 tokens) and the lowest proportion (10.72%) among the three types. From this it can be inferred that educated Americans tended to avoid statements with necessity.

Regarding the frequency of *must be+ing* in the TMC, it continued to increase from the 1920s to the 1970s, where it reached a peak (42 tokens). In the 1970s, *may be+ing* was the type most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*. After this period, there were fluctuations in the frequency of *must be+ing* from the 1970s to the 2000s. To sum up, a major characteristic of *must be+ing* is that it was the type most preferred by educated Americans in the 1970s.

Now let us examine the frequency of *may be+ing* in the TMC. The frequency of *may be+ing* continued to increase from the 1920s to the 1990s, where it reached a peak (337 tokens). This implies that *may be+ing* was the type most preferred by educated Americans in the 1990s. After this period, there was a decline in the frequency of *may be+ing* in the 2000s.

Now let us observe the frequency of *might be+ing* in the TMC. The frequency of *might be+ing* continued to increase from the 1920s to the 1940s, after which there were fluctuations. A major point to note is that the frequency of *might be+ing* reached a peak (117 tokens) in the 1940s. This indicates that *might be+ing* was the type most preferred by educated Americans in the 1940s. From this it can be inferred that at this time, educated Americans were fond of using statements with low possibility. After this period, the frequency of *might be+ing* continued to decrease from the 1940s to the 2000s, and then saw a dramatic decline in frequency in the 2000s. Consequently, one of the main characteristics of *might be+ing* is that it was the modal verb type most preferred by educated Americans in the 1940s.

4. A Comparative Analysis of the Four Corpora

In this section, we aim to provide a comparative analysis of the three types of modal verbs considered in the COCA, BNC, COHA, and TMC. As indicated in Table 1, *may be+ing* was the most commonly used modal verb type in America from 1990 to 2019. More specifically, the type *may be+ing* was the most widely used in America, followed by *might be+ing*, and *must be+ing*. This suggests that *may be+ing* is the most favorable type for American speakers or writers. This may be due to the fact that *may be+ing* is interpreted in terms of some kind of possibility and is suitable for conveying something reasonable. On the other hand, the BNC clearly indicates that *may be+ing* is the most frequently used modal verb type in the UK. More specifically, *may be+ing* is the most commonly used modal verb type in the UK, followed by *must be+ing*, and *might be+ing*. This suggests that both Americans and British people are fond of using statements with half possibility rather than statements with low possibility or statements with necessity. It is interesting to point out that Americans are fond of using statements with low possibility rather than statements with necessity, whereas British people are fond of using statements with necessity rather than statements with low possibility. From this it can be inferred that British people tend to make conclusive judgments from

evidence of some kind, but Americans do this less frequently than British people do. On the other hand, the COHA clearly indicates that *might be+ing* was the most commonly used modal verb type in America from 1810 to 2000. More specifically, *might be+ing* was the most widely used modal verb type in America, followed by *may be+ing*, and *must be+ing*. This in turn suggests that the frequency of the COHA is different from that of the COCA. The key difference between the frequency results of the COHA and COCA is that *might be+ing* was the most widely used modal verb type in America from 1810 to 1989, whereas *may be+ing* was the most commonly used modal verb type in America from 1990 to 2019. From this it can be conceived that statements with low possibility were the most preferred by Americans from 1810 to 1989, whereas statements with half possibility were the most preferred by Americans from 1990 to 2019. As observed earlier, the TMC clearly shows that *may be+ing* was the most commonly used modal verb type in America from the 1920s to the 2000s. More specifically, *may be+ing* was the modal verb type most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*. That *may be+ing* was the modal verb type most preferred by educated Americans suggests that educated Americans were fond of using statements with moderate possibility. Note, however, that educated Americans preferred using statements with low possibility to using statements with necessity. It should be highlighted here that the COCA, BNC, and TMC have one thing in common with respect to the frequency of the three types of modal verbs. That is to say, *may be+ing* was the type most preferred by Americans, British people, and educated Americans, which in turn suggests that *may be+ing* is the most favorable type for Americans and British people.

Now let us observe the spoken genre of the three types, comparing the frequency of the COCA and that of the BNC. It is noteworthy that *may be+ing* is the most commonly used modal verb type in America, whereas *might be+ing* is the most widely used modal verb type in the UK. This in turn suggests that Americans are fond of using statements with moderate possibility when they talk, whereas British people are fond of using statements with low possibility.

Now let us turn our attention to the use of the three types of modal verb in the fiction genre. When it comes to the fiction genre, *might be+ing* is the type most preferred modal verb type of the three types by Americans, whereas *must be+ing* is the type most preferred by British people. This in turn indicates that American writers are fond of using an allusive style of writing, whereas British writers are fond of using a strong and decisive style of writing.

Now let us turn our attention to the use of the three types of modal verb in the magazine genre. It is interesting to note that in the magazine genre, *may be+ing* is the most commonly used modal verb type in both America and the UK. This may be due to the fact that statements with half possibility are the most suitable for conveying articles, ads, or information. Exactly the same can be said for the newspaper genre. Just as in the magazine genre, the type *may be+ing* is the most commonly used modal verb type in America and the UK, which may be due to the fact that *may be+ing* carries moderate possibility and that it is neutral among the three types.

Finally, let us turn to the academic genre. Interestingly, *may be+ing* is the most frequently used of the modal verb types in both America and the UK. The reason why *may be+ing* is the most commonly used modal verb type in the academic genre in both America and the UK may be that *may be+ing*, with its moderate possibility, is the most suitable for conveying knowledge, academic information, and actual facts. Knowledge, academic information, and actual facts require statements with moderate possibility, which *may be+ing* carries. It is thus reasonable to conclude that *may be+ing* is the most commonly used modal verb type in America and the UK for this genre since it carries moderate possibility.

5. Conclusion

To sum up, we have provided a detailed frequency analysis of the three types of modal verbs within the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), the Corpus of Historical American English (COHA), and the Time Magazine Corpus (TMC). In section 3.1, we have shown that the most preferable type for Americans among the three modal verb types is *may be+ing*. More specifically, *may be+ing* is the type most preferred by Americans, followed by *might be+ing*, and *must be+ing*, in that order. In section 3.2, we have argued that *may be+ing* is the type most preferred by British people, followed by *must be+ing*, and *might be+ing*. Also, we have argued that *must be+ing* is the most commonly used of these modal verb types in the fiction genre in the UK, whereas *might be+ing* is the one most widely used in America. We have further argued that *may be+ing* the most frequently used one in the spoken genre in America, whereas *might be+ing* is the one most commonly used in the UK. This in turn suggests that Americans are fond of using statements with moderate possibility when they talk, whereas British people are fond of using statements with low possibility. We have contended, on the other hand, that *may be+ing* is the type most preferred by British and American writers in the magazine genre. We have maintained, however, that British journalists prefer statements with moderate possibility or necessity to statements with low possibility, whereas American journalists prefer statements with moderate possibility or low possibility to statements with necessity. Additionally, we have shown that in the academic genre, *may be+ing* is the most preferred of these modal verb types by both British people and Americans, which may reflect the fact that *may be+ing*, with its moderate possibility, is the most suitable for conveying knowledge, academic information, and actual facts. In section 3.3, we have contended that *might be+ing* was the type most preferred by Americans from 1810 to 2000, followed by *may be+ing*, and *must be+ing*. The main difference between the result of the COCA and that of the COHA is that *might be+ing* was the most frequently used of these modal verb types in America from 1810 to 1989, whereas *may be+ing* was the one most widely used from 1990 to 2019. In section 3.4, we have shown that *may be+ing* was the modal verb type most preferred by educated Americans, followed by *might be+ing*, and *must be+ing*. The fact that *may be+ing* was the type most preferred by educated Americans from the 1920s to the 2000s implies that they were fond of using statements with moderate possibility. It must be noted, however, that *must be+ing* has the lowest frequency and the lowest proportion among the three types. From this, it can be inferred that educated Americans tended to avoid statements with necessity.

References

- British National Corpus (BNC). <https://corpus.byu.edu/bnc>.
- Corpus of Contemporary American English (COCA). <https://corpus.byu.edu/coca>.
- Corpus of Historical American English (COHA). <https://corpus.byu.edu/coha>.
- Cowan, R. *The teacher's grammar of English*. Cambridge: Cambridge University Press. 2010.
- Kang, N-K. (2018). On Non-progressive Verbs: A Corpora-Based Analysis. *International Journal of Language and Linguistics* 5(4): 166-177. 2018.
- Kang, N-K. (2020). Three types of Auxiliaries: A Corpus-based Analysis. *International Journal of Interactive Storytelling* 4 (1). 2020.
- Koster, Jan. Anaphora and the Uniformity of Grammar. In H. Bennis, P. Pica, and J. Rooryck, eds., *Atomism and Binding*. 235-254. Foris Publications. 1997.
- Palmer, F. R. *The English Verb*. New York: Longman Publishers. 1987.

Quirk, R. and Greenbaum, S. (1993). *University grammar of English*. Longman.

Saeed, J. *Semantics*. Oxford. Blackwell Publishers. 2009.

Time Magazine Corpus (TMC). <https://corpus.byu.edu>