On Sure to and Bound to: A Corpora-based Analysis

Namkil Kang
Far East University South Korea

Abstract

The main purpose of this paper is to demonstrate that *sure to* and *bound to* are low similarity synonyms. With respect to the Movie Corpus, it is significant to note that *bound to* may have been preferable to *sure to* except the 2000s and the 2010s. A major point to note is that the frequency of *sure to* and *bound to* reached a peak (688 tokens vs. 442 tokens) in the 2010s. With respect to the BNC, it is worth pointing out that *sure to* and *bound to* show the same property in the spoken genre, whereas they show a different property in the other six genres. It can thus be inferred that in the seven genres of the BNC, *sure to* is 14.28% the same as *bound to*. The BNC clearly shows that *sure to get* (25 tokens) and *bound to say* (42 tokens) are the most preferable ones for the British. The BNC further shows that 35% of forty verbs are the collocations of *sure to* and *bound to*. It is thus reasonable to assume that *sure to* and *bound to* are low similarity synonyms.

Keywords: Movie Corpus, British National Corpus, type, token, sure to, bound to

Introduction

As pointed out by Murphy (2016, 2019), *sure to* and *bound to* are used interchangeably. The main purpose of this article is to compare *sure to* with *bound to* in two corpora (the Movie Corpus and the British National Corpus). First, we consider the diachronic characteristics of *sure to* and *bound to* in the Movie Corpus (1930s-2010s). We consider when two types reached a peak (the highest frequency). On the other hand, we consider when they had the lowest frequency. Second, we investigate the similarity between *sure to* and *bound to* in the BNC. By observing the frequency of *sure to* and *bound to* in seven genres, we can examine how close they are. Third, we investigate the similarity between *sure to* and *bound to* by examining their collocations. We consider how close they are in the British National Corpus. The organization of this article is as follows. In section 2, we argue that

bound to may have been favored over *sure to* except the 2000s and the 2010s. More interestingly, the frequency of *sure to* and *bound to* reach a peak (688 tokens vs. 442 tokens) in the 2010s. In section 3, we further argue that *sure to* and *bound to* show the same pattern in the spoken genre, whereas they show a different pattern in the other six genres. It can thus be inferred that in the seven genres of the BNC, *sure to* is 14.28% the same as *bound to*. In section 4, we contend that *sure to get* and *bound to say* are the most preferred (25 tokens vs. 42 tokens) by the British. Additionally, we show that 35% of forty verbs are the collocations of *sure to* and *bound to*.

Sure to and Bound to in the Movie Corpus

In section 2, we consider the diachronic characteristics of *sure to* and *bound to* in the Movie Corpus. Table 1 shows the frequency of *sure to* and *bound to* from the 1930s to the 2010s:

Table 1 Frequency of sure to and bound to in the Movie Corpus

Type	Sure to	Bound to
1930s	96	117
1940s	151	212
1950s	142	241

1960s	93	234
1970s	121	176
1980s	129	182
1990s	256	258
2000s	505	414
2010s	688	442
ALL	2,181	2,276
US/CA	1,780	1,456
UK/IE	247	629
AU/NZ	28	41
MISC	126	150

As can be seen from Table 1, *bound to* was preferable to *sure to* except the 2000s and the 2010s. The frequency of *bound to* (2,276 tokens) is slightly higher than that of *sure to* (2,181 tokens). The following graph shows the diachronic characteristics of *sure to* and *bound to* from the 1930s to the 2010s:

Figure 1 Frequency of sure to and bound to in the Movie Corpus
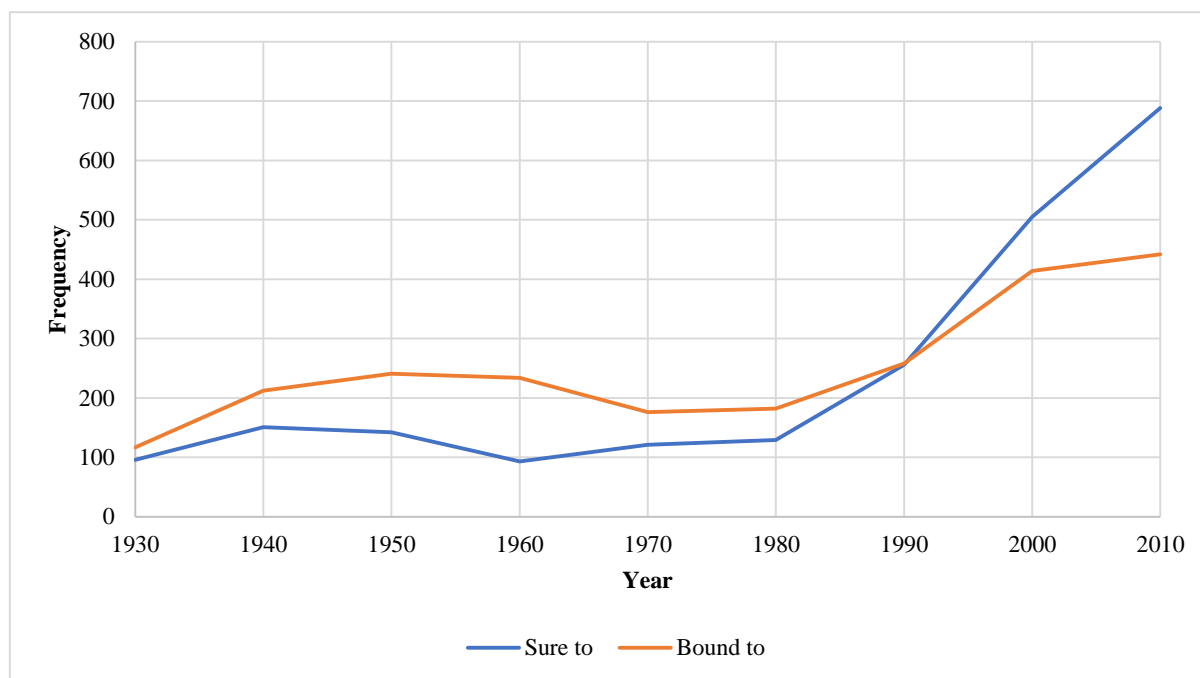
As shown in Table 1, there was a minimal increase (a rise of 55 tokens) in the figure of *sure to* from the 1930s to the 1940s. It should be noted, however, that there was a small drop (a reduction of 58 tokens) in the figure of *sure to* from the 1940s to the 1960s. More interestingly, it can be seen from Table 1 that there was a small rise (an increase of 36 tokens) in the figure of *sure to* from the 1960s to the 1980s. Quite interestingly, as shown in Table 1, there was a marked increase (a rise of 559 tokens) in the figure of

sure to from the 1980s to the 2010s. Most importantly, *sure to* had the highest frequency (688 tokens) in the 2010s, whereas it had the lowest frequency (96 tokens) in the 1930s. From this, it is evident that there was a continuing upward trend except in the 1950s and the 1960s in the frequency of *sure to*.

It is worth pointing out that there was a small increase (a rise of 124 tokens) in the figure of *bound to* from the 1930s to the 1950s. It should be pointed out, however, that there was a slight drop (a

reduction of 65 tokens) in the figure of *bound to* from the 1950s to the 1970s. Again, there was a minimal increase (a rise of 82 tokens) in the figure of *bound to* from the 1970s to the 1990s. Quite interestingly, there was a large rise (an increase of 184 tokens) in the figure of *bound to* from the 1990s to the 2010s. Most importantly, *bound to* was preferable to *sure to* except the 2000s and the 2010s. In the 2000s and the 2010s, *sure to* was the preferable one in the films of six countries. More importantly, the

frequency of *bound to* reached a peak (442 tokens) in the 2010s, which in turn suggests that it was the most preferred one in that period. Finally, it is noteworthy that *bound to* was the most preferred (1,456 tokens) by American and Canadian film writers, followed by British and Irish ones (629 tokens), and Australian and New Zealand ones (41 tokens), in that order. We thus conclude that *bound to* was preferable to *sure to* except the 2000s and the 2010s.

Sure to and Bound to in the BNC

In this section, we consider the similarity between *sure to* and *bound to* by examining their frequency in seven genres. Table 2 shows the genre frequency of *sure to* and *bound to* in the BNC:

Table 2 Genre Frequency of sure to and bound to in the BNC

GENRE	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	NON-ACAD	ACADEMIC	MISC
Sure to	631	31	130	104	117	66	23	160
Bound to	2,253	217	371	154	228	355	532	396

More interestingly, the frequency of *bound to* (2,253 tokens) is three times higher than that of *sure to* (631 tokens). We take this as suggesting that *bound to* may be preferred over *sure to* by the British.

It is interesting to note that *sure to* rank first (160 tokens) in the miscellaneous genre, whereas *bound to* ranks first (532 tokens) in the academic genre. Quite interestingly, *sure to* and *bound to* show no similarity in rank-one. It should be pointed out that the frequency of *bound to* (396 tokens) is two times higher than that of *sure to* (160 tokens) in the miscellaneous genre. We take this fact as implying that *bound to* may be favored over *sure to* in the mixed genre. It is worth pointing out that the frequency of *bound to* (532 tokens) is twenty-three times higher than that of *sure to* (23 tokens) in the academic genre. This in turn suggests that the British prefer using *bound to* rather than using *sure to* in the academic genre.

It is worthwhile noting that *sure to* ranks second (130 tokens) in the fiction genre, whereas *bound to* ranks second (396 tokens) in the miscellaneous genre. Again, *sure to* and *bound to* show no similarity in rank-two. It must be noted that the frequency of *bound to* (371 tokens) in the fiction genre is two times higher than that of *sure to* (130 tokens). This in turn implies that *bound to* is preferred over *sure to* by British writers.

It is noteworthy that *sure to* ranks third (117 tokens) in the newspaper genre, whereas *bound to* ranks third (371 tokens) in the fiction genre. Again, there is no similarity between *sure to* and *bound to* in rank three. With respect to the newspaper genre, it is worth observing that the frequency of *bound to* (228 tokens) is somewhat higher than that of *sure to* (117 tokens). This in turn indicates that British journalists prefer to use *bound to* rather than use *sure to*.

It is worthwhile pointing out that *sure to* ranks fourth (104 tokens) in the magazine genre, whereas *bound to* ranks fourth (355 tokens) in the non-academic genre. Again, there is no similarity between

sure to and *bound to* in rank-four. With respect to the magazine genre, it is worthwhile noting that the frequency of *bound to* (154 tokens) is much higher than that of *sure to* (104 tokens). This in turn suggests that *bound to* is preferred over *sure to* by British journalists. It is interesting to note that the frequency of *bound to* (355 tokens) in the non-academic genre is five times higher than that of *sure to* (66 tokens). It can thus be inferred that *bound to* is preferable to *sure to* in the non-academic genre.

It is worthwhile mentioning that *sure to* ranks fifth (66 tokens) in the non-academic genre, whereas *bound to* ranks fifth (228 tokens) in the newspaper genre. Again, there is no similarity between *sure to* and *bound to* in rank five.

It is interesting to note that *sure to* and *bound to* rank sixth (31 tokens vs. 217 tokens) in the spoken genre. Quite interestingly, there is a close similarity between *sure to* and *bound to* in rank-six. It should be noted, however, that the frequency of *bound to* (217 tokens) in the spoken genre is seven times higher than that of *sure to* (31 tokens). This in turn implies that the British prefer using *bound to* rather than using *sure to* in daily conversation.

Finally, it is interesting to point out that *sure to* ranks seventh (23 tokens) in the academic genre, whereas *bound to* ranks seventh (154 tokens) in the magazine genre. More interestingly, there is no similarity between *sure to* and *bound to* in rank-seven. To sum up, *sure to* and *bound to* show the same property in the spoken genre, but they show a different property in the other genres (six genres). Thus, in the seven genres of the BNC, *sure to* is 14.28% the same as *bound to*. The following graphs show the use of *sure to* and *bound to* in seven genres:

Figure 2 Use of sure to in seven genres
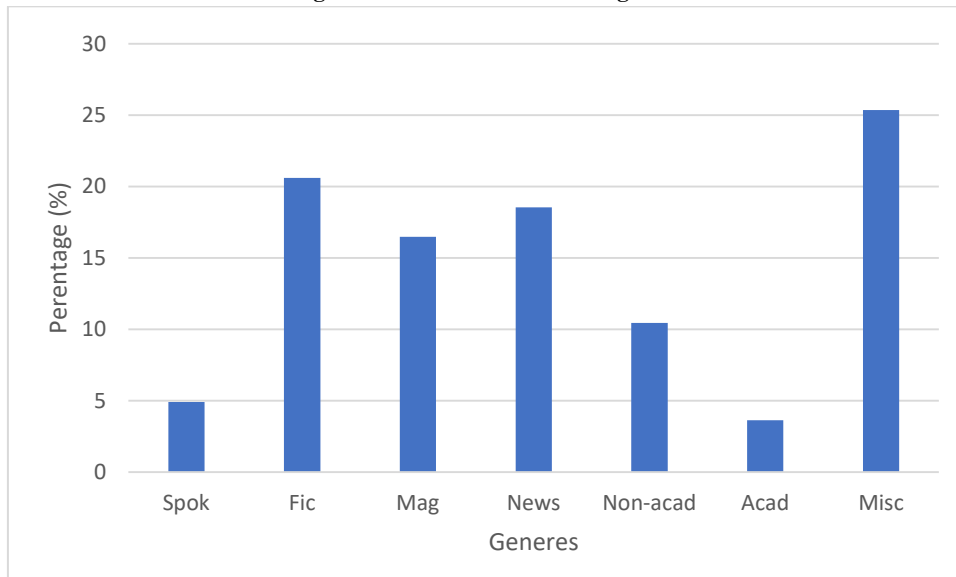
As indicated in Figure 2, the miscellaneous genre is the most influenced by *sure to*, followed by the fiction genre, the newspaper genre, the magazine genre, the non-academic genre, the spoken genre, and the academic genre, in that order.

Figure 3 Use of bound to in seven genres
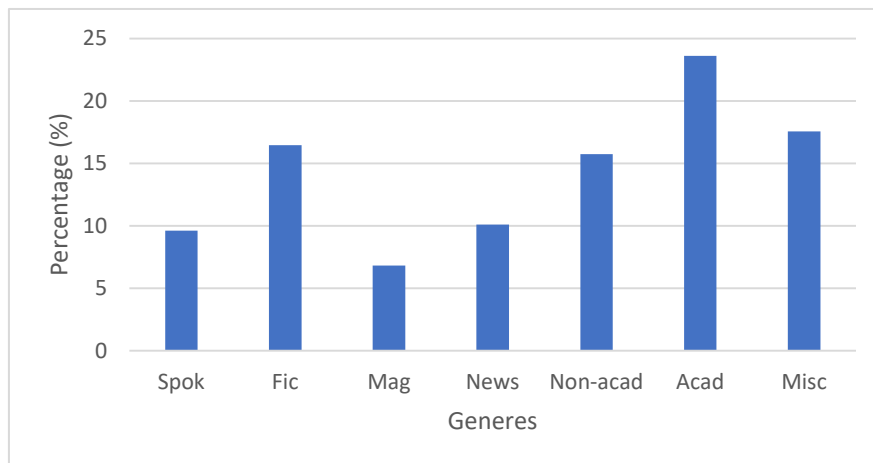
As exemplified in Figure 3, the academic genre is the most influenced by *bound to*, followed by the miscellaneous genre, the fiction genre, the non-academic genre, the newspaper genre, the spoken genre, and the magazine genre, in descending order.

The Collocations of sure to and bound to in the BNC

In this section, we consider the collocations of *sure to* and *bound to* in the BNC. We compare the collocation of *sure to* and that of *bound to* in the top 27. The following table shows the collocation of *sure to* in the BNC:

Table 3 Collocation of sure to in the BNC

Number	Collocation	Frequency
1	sure to get	25
2	sure to find	15

3	sure to take	15
4	sure to go	14
5	sure to make	13
6	sure to come	12
7	sure to give	12
8	sure to use	11
9	sure to keep	7
10	sure to check	7
11	sure to remember	7
12	sure to win	7
13	sure to ask	6
14	sure to include	6
15	sure to look	6
16	sure to meet	6
17	sure to put	6
18	sure to know	5
19	sure to leave	5
20	sure to bring	4
21	sure to allow	4
22	sure to fail	4
23	sure to follow	4
24	sure to provoke	4
25	sure to remain	4
26	sure to say	4
27	sure to see	4

As alluded to in Table 3, *sure to get* is the most frequently used one (25 tokens) in the UK. This in turn implies that *sure to get* is the most preferable one (25 tokens) for the British. As indicated in Table 3, *sure to get* is the most preferred (25 tokens) by the British, followed by *sure to find* (*sure to take*), *sure to go*, *sure to make*, *sure to come* (*sure to give*), and *sure to use*, in that order. Quite interestingly, *sure to go* ranks fourth (14 tokens) in the BNC, whereas *sure to come* ranks sixth (12 tokens). It is worth observing that the expression *sure to remember* ranks ninth (7 tokens) in the BNC. Finally, the expression *sure to win* also ranks ninth (7 tokens) in the BNC. We thus conclude that *sure to get* is the most preferred one (25 tokens) in the UK.

The following table shows the collocation of *bound to* in the top 27:

Table 4 Collocation of bound to in the BNC

Number	Collocation	Frequency
1	bound to say	42
2	bound to come	37
3	bound to get	36
4	bound to happen	35
5	bound to make	30
6	bound to take	28
7	bound to fail	26

8	bound to find	25
9	bound to affect	23
10	bound to go	21
11	bound to pay	19
12	bound to give	17
13	bound to lead	15
14	bound to suffer	15
15	bound to become	14
16	bound to follow	14
17	bound to provide	13
18	bound to ask	12
19	bound to cause	12
20	bound to increase	12
21	bound to see	12
22	bound to win	12
23	bound to change	11
24	bound to accept	11
25	bound to fail	11
26	bound to feel	11
27	bound to remain	11

As can be seen from Table 4, *bound to say* is the most widely used one (42 tokens) in the UK. This in turn suggests that *bound to say* is the most preferable one (42 tokens) for the British. As illustrated in Table 4, *bound to say* is the most preferred (42 tokens) by the British, followed by *bound to come*, *bound to get*, *bound to happen*, *bound to make*, and *bound to take*, in descending order. It is important to note that *bound to say* and *sure to get* are the most preferred ones (42 tokens vs. 25 tokens) in the UK. It is interesting to point out that *bound to come* ranks second (37 tokens) in the BNC, whereas *sure to come* ranks sixth (12 tokens). More interestingly, *bound to go* ranks tenth (21 tokens) in the BNC, whereas *sure to go* ranks fourth (14 tokens). We thus conclude that *bound to say* is the most preferable one (42 tokens) among the British.

Now, attention is paid to the visualization of the collocations of *sure to* and *bound to* in the top 27:

Figure 4 Visualization of the collocations of sure to and bound to

As exemplified in Figure 4, twenty-seven verbs are linked to *sure to* and *bound to*, respectively. Most importantly, fourteen verbs are linked to both *sure to* and *bound to*, which indicates that the fourteen verbs are the collocations of both *sure to* and *bound to*. The collocations include *say*, *come*, *get*, *make*, *take*, *fail*, *go*, *give*, *follow*, *see*, *win*, *ask*, *remain*, and *find*. It can thus inferred that 35% of forty verbs are the collocations of *sure to* and *bound to*. We thus conclude that *sure to* and *bound to* are low similarity synonyms.

Conclusion

To sum up, we have compared *sure to* with *bound to* in two corpora (the Movie Corpus and the British National Corpus). In section 2, we have argued that *bound to* may have been preferable to *sure to* except the 2000s and the 2010s. We have shown that the frequency of *sure to* and *bound to* reached a peak (688 tokens vs. 442 tokens) in the 2010s. In section 3, we have maintained that *sure to* and *bound to* show the same property in the spoken genre, whereas they show a different property in the other six genres. It can thus be inferred that in the seven genres of the BNC, *sure to* is 14.28% the same as *bound to*. In section 4, we have contended that *sure to get* is the most preferred (25 tokens) by the British, followed by *sure to find* (*sure to take*), *sure to go*, and *sure to make*, in that order. We have also contended that *bound to say* is the most preferred (42 tokens) by the British, followed by *bound to come*, *bound to get*, and *bound to happen*. Additionally, we have shown that 35% of forty verbs are the collocations of *sure to* and *bound to*. We thus conclude that *sure to* and *bound to* are low similarity synonyms.

References

1. British National Corpus (BNC). 1, May 2022. Online https://corpus.byu.edu/bnc
2. Murphy, R. (2016). *Grammar in Use*. Cambridge University Press.
3. Murphy, R. (2019). *English Grammar in Use*. Cambridge University Press.
4. TV Corpus (TVC). 1, May 2022. Online https://english-corpora.org /movies/