

A Comparative Analysis of Die of and Die from in Four Corpora

Namkil Kang
 Far East University South Korea

ABSTRACT

The ultimate goal of this paper is to provide a comparative analysis of *die of* and *die from* in the TV Corpus, the Movie Corpus, the Corpus of Contemporary American English, and the British National Corpus. A major point to note is that *die of* is the preferred one in America, the UK, Canada, Australia, New Zealand, and Ireland. A further point to note is that *die of* and *die from* in the TV Corpus reached a peak in the 2010s. With respect to the Movie Corpus, it is worth noting that *die of* was the preferable one in America, the UK, Canada, Australia, New Zealand, and Ireland. It is interesting to point out that *die of* had the highest frequency in the 2000s, whereas *die from* had the highest frequency in the 2010s. More interestingly, the COCA clearly shows that *die of cancer* is the most widely used one in America, whereas *die from heart* is the most commonly used one. The COCA further shows that 38.88% of thirty six nouns are the collocations of both *die of* and *die from*. Similarly, the BNC shows that *die of starvation* and *die from starvation* are the most frequently used by the British. With respect to the BNC, it should be noted that 19.04% of forty two nouns are the collocations of both *die of* and *die from*. It is thus reasonable to assume that *die of* and *die from* are low similarity synonyms.

Keywords: *TV Corpus, Movie Corpus, COCA, BNC, type, token, die of, die from*

***Corresponding Author**

Namkil Kang
 Far East University South Korea



© Copy Right, IJHSS, 2022. All Rights Reserved

INTRODUCTION

As Murphy [1,2] points out, *die of* and *die from* are used synonymously. The main purpose of this paper is to provide a comparative analysis of *die of* and *die from* in the TV Corpus, the Movie Corpus, the Corpus of Contemporary American English[3], and the British National Corpus[4]. We examine which type is the preferred one in the TV Corpus and the Movie Corpus. We try to provide a diachronic analysis of *die of* and *die from* in the TV Corpus and the Movie Corpus[5]. Also, we consider the collocations of *die of* and *die from* in the COCA and the BNC. We examine which nouns are the preferable ones in America and the UK. Additionally, we examine similarities between *die of* and *die from* in terms of the software NetMiner. The organization of this paper is as follows. In section 2.1, we argue that *die of* is the preferred one in America, the UK, Canada, Australia, New Zealand, and Ireland. It is interesting to note that *die of* and *die from* reached a peak in the 2010s. In section 2.2, we further argue that *die of* was the preferable one in America, the UK, Canada, Australia, New Zealand, and Ireland. We also show that *die of* had the highest frequency in the 2000s, whereas *die from* had the highest frequency in the 2010s. In section 3.1, we maintain that *die of cancer* is the most widely used one in America, whereas *die from heart* is the most commonly used one. We also show that 38.88% of thirty six nouns are the collocations of both *die of* and *die from*. In section 3.2, we contend that *die of starvation* and *die from starvation* are the most frequently used ones in the UK. Finally, we show that 19.04% of forty two nouns are the collocations of both *die of* and *die from* in the BNC.

The Frequency of die of and die from in Two Corpora

The TV Corpus

In what follows, we provide a diachronic analysis of *die of* and *die from* in the TV Corpus[6]. Also, we provide a comparative analysis of *die of* and *die from* in the TV Corpus. Table 1 shows the frequency of *die of* and *die from* from the 1950s to the 2010s:

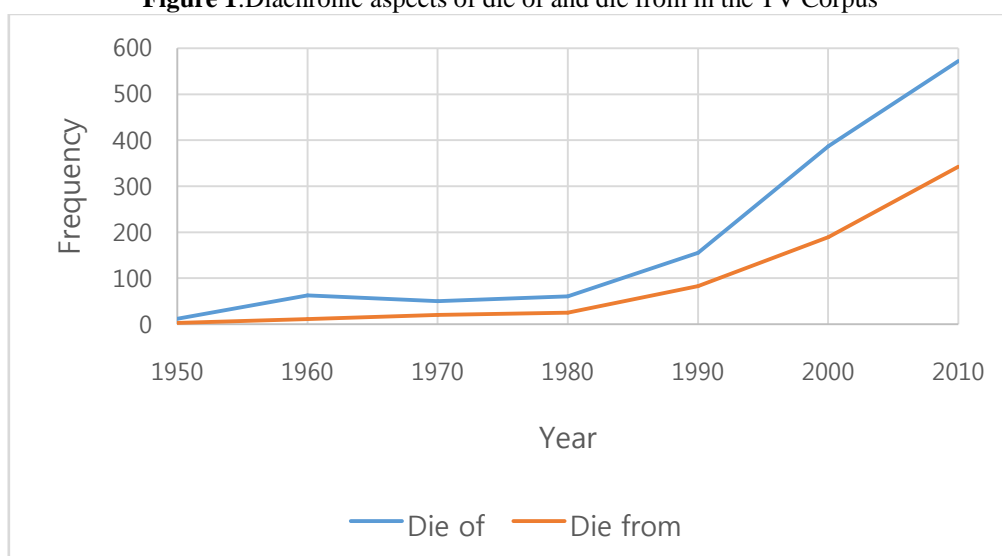
Table 1: Frequency of die of and die from in the TV Corpus

Type	Die of	Die from
1950s	12	3
1960s	63	11
1970s	50	20

1980s	61	25
1990s	155	83
2000s	387	189
2010s	572	342
All	1,300	673
US/CA	895	566
UK/IE	369	95
AU/NZ	23	8
Misc	13	4

It is significant to note that *die of* is the preferred one in America, the UK, Canada, Australia, New Zealand, and Ireland. As illustrated in Table 1, the frequency of *die of* is almost two times higher than that of *die from*. The overall frequency of *die of* is 1,300 tokens, whereas that of *die from* is 673 tokens. This in turn suggests that *die of* is favored over *die from* in America, the UK, Canada, Australia, New Zealand, and Ireland. The following graph shows the diachronic aspects of *die of* and *die from* from the 1950s to the 2010s:

Figure 1: Diachronic aspects of *die of* and *die from* in the TV Corpus



As alluded to in Figure 1, *die of* was always favored over *die from* from the 1950s to the 2010s. More specifically, the frequency of *die of* was always much higher than that of *die from* from the 1950s to the 2010s. This in turn indicates that six countries' (America, the UK, Canada, Australia, New Zealand, and Ireland) celebs preferred using *die of* to using *die from* from the 1950s to the 2010s.

Interestingly, there was a slight rise in the figure of *die of* from the 1950s to the 1960s. However, there was a sudden decline in the figure of *die of* in the 1970s. More specifically, there was a decrease of 13 tokens in the 1970s. More importantly, there was a sharp increase in the figure of *die of* from the 1980s to the 2010s. It is interesting to note that *die of* reached a peak in the 2010s. This in turn indicates that *die of* was the most preferable one for six countries' celebs. It should be noted, however, that *die of* had the lowest frequency in the 1950s, which in turn implies that *die of* was the least preferred one in that period.

More interestingly, there was a gradual increase in the figure of *die from* from the 1950s to the 2010s. It might be worth pointing out that *die from* had the highest frequency in the 2010s. This in turn shows that six countries' celebs preferred *die from* the most in that period. It must be noted, however, that *die from* had the lowest frequency in the 1950s, which indicates that *die from* was the least preferred one in that period. It is interesting to point out that *die of* was favored over *die from* in America and Canada. It is also interesting to note that *die of* was preferable to *die from* in the UK and Ireland. The same can be said of Australia and New Zealand. To be more specific, *die of* was preferred over *die from* by two countries' celebs. We thus conclude that *die of* was preferable to *die from* in six countries from the 1950s to the 2010s.

The Movie Corpus

In the following, we consider the diachronic aspects of *die of* and *die from* in the Movie Corpus. Also, we provide a comparative analysis of *die of* and *die from* in the Movie Corpus. Table 2 shows the diachronic aspects of *die of* and *die from*

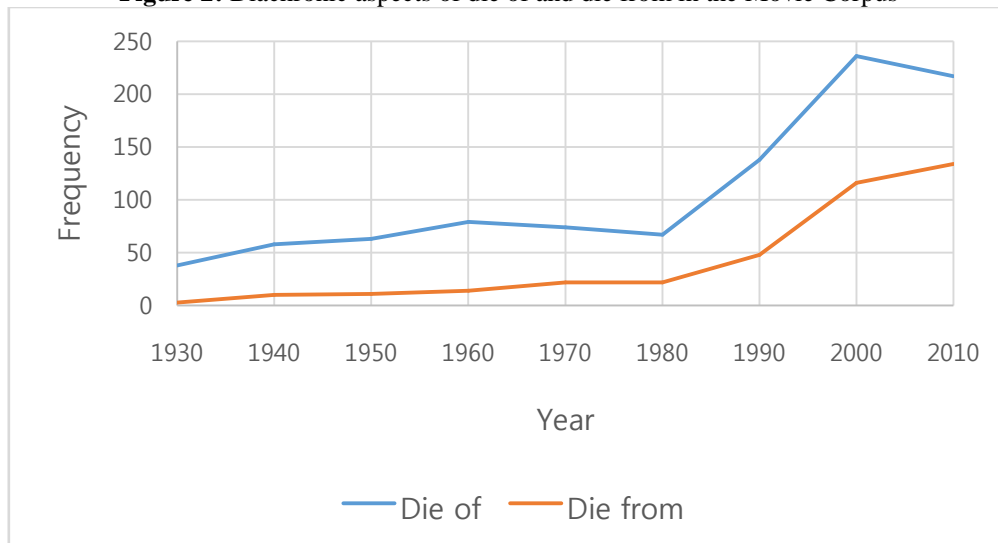
from in the Movie Corpus from the 1930s to the 2010s:

Table 2: Frequency of *die of* and *die from* in the Movie Corpus

Type	Die of	Die from
1930s	38	3
1940s	58	10
1950s	63	11
1960s	79	14
1970s	74	22
1980s	67	22
1990s	138	48
2000s	236	116
2010s	217	134
All	970	380
US/CA	664	285
UK/IE	205	42
AU/NZ	19	7
Misc	82	46

It is worth observing that *die of* was the preferable one in America, the UK, Canada, Australia, New Zealand, and Ireland. The overall frequency of *die of* is 970 tokens, whereas that of *die from* is 380 tokens. This in turn indicates that *die of* was preferred over *die from* by six countries' movie writers. The following graph indicates the diachronic aspects of *die of* and *die from* in the Movie Corpus from the 1930s to the 2010s:

Figure 2: Diachronic aspects of *die of* and *die from* in the Movie Corpus



It might be worth pointing out that there was a gradual rise in the figure of *die of* from the 1930s to the 1960s. There was a rise of 41 tokens from the 1930s to the 1960s. Interestingly, there was a slight decline in the figure of *die of* from the 1970s to the 1980s. More interestingly, there was a sudden rise in the figure of *die of* from the 1990s to the 2000s and after this period, there was a slight decrease in its figure in the 2010s. It is important to note that *die of* had the highest frequency (236 tokens) in the 2000s. This in turn shows that *die of* was the most preferred one in America, the UK, Canada, Australia, New Zealand, and Ireland. It should be noted, however, that *die of* had the lowest frequency in the 1930s. This in turn suggests that *die of* was the most undesired one in the 1930s.

It is significant to note that there was a steady increase in the figure of *die from* from the 1930s to the 2010s. More importantly, *die from* reached a peak in the 2010s, which in turn indicates that movie writers preferred *die from* the most in the 2010s. It must be noted, however, that *die from* had the lowest frequency in the 1930s, which in turn indicates that *die from* was the least preferred one in that period. It is noteworthy that *die of* was always preferred over *die from* by six countries' movie writers from the 1930s to the 2010s. Additionally, *die of* was favored over *die from* in American and Canadian films from the 1930s to the 2010s. Likewise, *die of* was preferable to *die from* in British and Irish films from the 1930s to the 2010s. The same applies to Australian and New Zealand films. *Die of* was favored over *die from* in Australian and New Zealand films from the 1930s to the 2010s. We thus conclude that *die of* was always preferred over

die from by six countries' movie writers from the 1930s to the 2010s.

A Collocation Analysis of *die of* and *die from* in Two Corpora

A Collocation Analysis of *die of* and *die from* in the COCA

In the following, we examine the collocations of *die of* and *die from* in the COCA. We also compare the collocation of *die of* and that of *die from* in terms of the software NetMiner. Table 3 indicates the collocation of *die of* in the top 25:

Table 3: Collocation of *die of* in the COCA

Number	Collocation of <i>die of</i>	Frequency
1	Die of cancer	115
2	Die of starvation	114
3	Die of hunger	95
4	Die of thirst	83
5	Die of AIDS	80
6	Die of heart	70
7	Die of boredom	51
8	Die of breast	48
9	Die of embarrassment	41
10	Die of exposure	39
11	Die of lung	32
12	Die of shame	29
13	Die of dehydration	28
14	Die of hypothermia	24
15	Die of fright	23
16	Die of pneumonia	23
17	Die of heat	18
18	Die of disease	17
19	Die of prostate	17
20	Die of malnutrition	15
21	Die of loneliness	15
22	Die of neglect	14
23	Die of cold	13
24	Die of complications	13
25	Die of asphyxiation	13

It is important to point out that *die of cancer* is the most widely used one in America. This in turn implies that the expression *die of cancer* is the most frequently used by Americans. As illustrated in Table 3, *die of cancer* is the most commonly used one in America, followed by *die of starvation*, *die of hunger*, *die of thirst*, *die of AIDS*, *die of heart*, *die of boredom*, and *die of breast*, in that order. It is interesting to point out that *die of starvation* ranks second in the COCA. More interestingly, the expression *die of boredom* ranks seventh in the COCA. It is worth mentioning that *die of pneumonia* ranks fifteenth in the COCA. We thus conclude that *die of cancer* is the most widely used one in America.

Now attention is paid to the collocation of *die from* in the COCA:

Table 4: Collocation of *die from* in the COCA

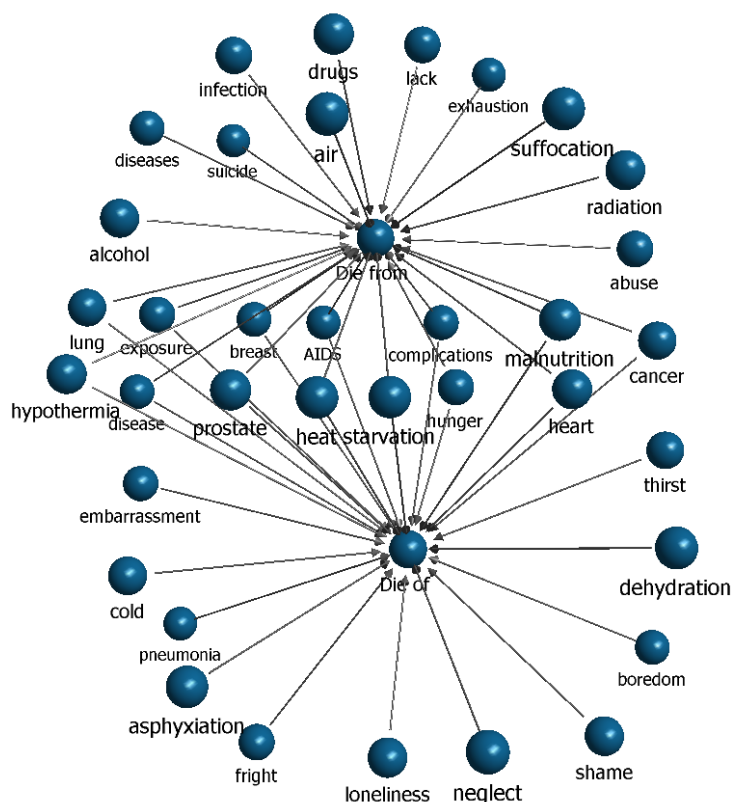
Number	Collocation of <i>die from</i>	Frequency
1	Die from heart	47
2	Die from lack	37
3	Die from breast	34
4	Die from cancer	29
5	Die from AIDS	27
6	Die from starvation	25
7	Die from exposure	20
8	Die from lung	13
9	Die from alcohol	12
10	Die from complications	12
11	Die from disease	12
12	Die from prostate	12

13	Die from suicide	11
14	Die from drugs	10
15	Die from hunger	10
16	Die from hypothermia	9
17	Die from heat	8
18	Die from infection	8
19	Die from malnutrition	8
20	Die from radiation	7
21	Die from air	6
22	Die from abuse	6
23	Die from diseases	6
24	Die from suffocation	6
25	Die from exhaustion	5

As indicated in Table 4, the expression *die from heart* is the most commonly used one in America. This in turn suggests that *die from heart* is the most frequently used by Americans. As alluded to in Table 4, *die from heart* is the most widely used one in America, followed by *die from lack*, *die from breast*, *die from cancer*, *die from AIDS*, *die from starvation*, *die from exposure*, and *die from lung*, in descending order. It is worth noting that *die from breast* ranks third in the COCA, whereas *die of breast* ranks eighth. It is interesting to note, on the other hand, that *die from cancer* ranks fourth in the COCA, whereas *die of cancer* ranks first. Additionally, it should be pointed out that *die from starvation* ranks sixth in the COCA, whereas *die of starvation* ranks second. More interestingly, *die from complications* ranks ninth, whereas *die of complications* ranks twenty third. We thus conclude that *die from heart* ranks first in the COCA, whereas *die of cancer* ranks first.

Finally, let us turn our attention to the visualization of the collocations of *die of* and *die from* in the COCA. This 3-D visualization of the collocations of *die of* and *die from* was performed by the software NetMiner:

Figure 3: Visualization of the collocations of die of and die from in the COCA



As exemplified in Figure 3, particular nouns are linked to *die of* and *die from*, respectively. Most importantly, fourteen nouns are linked to both *die of* and *die from*, which means that they are the collocations of both *die of* and *die from*. The nouns linked to two types are *breast*, *heart*, *cancer*, *AIDS*, *starvation*, *exposure*, *lung*, *complications*, *disease*,

prostate, hunger, hypothermia, heat, and malnutrition. It is noteworthy that 38.88% of thirty six nouns are the collocations of both *die of* and *die from*. This in turn suggests that *die of* and *die from* are low similarity synonyms.

A Collocation Analysis of *die of* and *die from* in the BNC

In what follows, we consider the collocations of *die of* and *die from* in the BNC. We also observe similarities between *die of* and *die from* in the top 25. Table 5 indicates the collocation of *die of* in the BNC:

Table 5: Collocation of *die of* in the BNC

Number	Collocation of <i>die of</i>	Frequency
1	Die of starvation	19
2	Die of AIDS	11
3	Die of hunger	8
4	Die of shame	8
5	Die of exposure	8
6	Die of heart	7
7	Die of shock	6
8	Die of cancer	5
9	Die of pneumonia	4
10	Die of hypothermia	4
11	Die of cholera	4
12	Die of boredom	4
13	Die of lung	3
14	Die of disease	3
15	Die of cold	3
16	Die of embarrassment	3
17	Die of accidents	2
18	Die of breast	2
19	Die of diseases	2
20	Die of dehydration	2
21	Die of fright	2
22	Die of grief	2
23	Die of malnutrition	2
24	Die of heartbreak	2
25	Die of thirst	2

It is significant to note that *die of starvation* is the most frequently used one in the UK. This in turn indicates that the expression *die of starvation* is the most commonly used by the British. As illustrated in Table 5, *die of starvation* is the most widely used by the British, followed by *die of AIDS*, *die of hunger*, *die of shame*, *die of exposure*, *die of heart*, *die of shock*, and *die of cancer*, in that order. It is worth observing that *die of heart* ranks sixth in the COCA and BNC. More interestingly, *die of cancer* ranks first in the COCA, whereas it ranks eighth in the BNC. It should be pointed out, on the other hand, that *die of starvation* ranks second in the COCA, whereas it ranks first in the BNC. We thus conclude that *die of cancer* is the most widely used one in America, whereas *die of starvation* is the most commonly used one in the UK.

Now let us turn to the collocation of *die from* in the BNC:

Table 6: Collocation of *die from* in the BNC

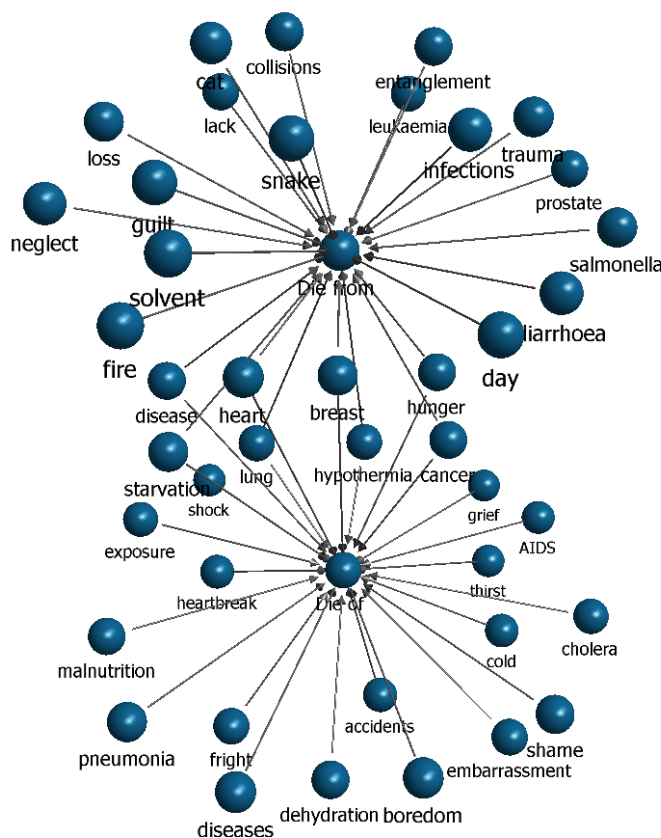
Number	Collocation of <i>die from</i>	Frequency
1	Die from starvation	4
2	Die from heart	3
3	Die from lack	3
4	Die from cancer	3
5	Die from breast	2
6	Die from disease	2
7	Die fromleukaemia	2
8	Die fromhunger	2
9	Die fromlung	2
10	Die fromtrauma	1
11	Die fromloss	1

12	Die fromsolvent	1
13	Die fromsnake	1
14	Die fromsalmonella	1
15	Die fromprostate	1
16	Die fromneglect	1
17	Die frominfections	1
18	Die fromhypothermia	1
19	Die fromguilt	1
20	Die fromfire	1
21	Die fromentanglement	1
22	Die fromdiarrhoea	1
23	Die fromday	1
24	Die fromcollisions	1
25	Die from cat	1

It is important to note that *die from starvation* is the most widely used one in the UK. This in turn suggests that the expression *die from starvation* is the most commonly used by the British. As alluded to in Table 6, *die from starvation* is the most frequently used one in the UK, followed by *die from heart*, *die from lack*, *die from cancer*, *die from breast*, and *die from disease*, in descending order. More interestingly, *die from starvation* and *die of starvation* rank first in the BNC. On the other hand, *die from starvation* and *die of starvation* rank sixth and second in the COCA, respectively. On the other hand, *die from heart* and *die of heart* rank second and sixth in the BNC, whereas *die from heart* and *die of heart* rank first and sixth in the COCA. It is worth noting that *die from cancer* and *die of cancer* rank second and eighth in the BNC, whereas *die from cancer* and *die of cancer* rank fourth and first in the COCA. It is also worth pointing out that *die from disease* and *die of disease* rank fifth and thirteenth in the BNC, whereas *die from disease* and *die of disease* rank ninth and eighteenth in the COCA. We thus conclude that *die from starvation* and *die of starvation* rank first in the BNC.

Finally, let us turn our attention to the visualization of the collocations of *die of* and *die from* in the BNC:

Figure 4: Visualization of the collocations of *die of* and *die from* in the BNC



As illustrated in Figure 4, twenty five nouns are linked to *die of* and twenty five nouns are also linked to *die from*. Most importantly, eight nouns are linked to both *die of* and *die from*, which indicates that they are the collocations of both

die of and *die from*. The nouns linked to two types are *starvation, heart, cancer, breast, disease, hunger, lung, and hypothermia*. This in turn implies that 19.04% of forty two nouns are the collocations of both *die of* and *die from*. From all of this, it is evident that *die of* and *die from* are low similarity synonyms.

CONCLUSION

To sum up, we have provided a comparative analysis of *die of* and *die from* in four corpora. In section 2.1, we have argued that *die of* is the preferred one in America, the UK, Canada, Australia, New Zealand, and Ireland. Interestingly, the types *die of* and *die from* reached a peak in the 2010s. In section 2.2, we have further argued that *die of* was the preferable one in America, the UK, Canada, Australia, New Zealand, and Ireland. We have shown that the type *die of* had the highest frequency in the 2000s, whereas the type *die from* had the highest frequency in the 2010s. In section 3.1, we have maintained that *die of cancer* is the most commonly used one in America, whereas *die from heart* is the most widely used one. Additionally, we have shown that 38.88% of thirty six nouns are the collocations of both *die of* and *die from*. In section 3.2, we have contended that *die of starvation* and *die from starvation* are the most commonly used by the British. Finally, we have shown that 19.04% of forty two nouns are the collocations of both *die of* and *die from*.

REFERENCES

1. Murphy, R. (2016). *Grammar in Use*. Cambridge University Press.
2. Murphy, R. (2019). *English Grammar in Use*. Cambridge University Press
3. Corpus of Contemporary American English (COCA). Online <https://corpus.byu.edu/coca>
4. British National Corpus (BNC). Online <https://corpus.byu.edu/bnc>
5. Movie Corpus (MC). Online <https://english-corpora.org/movies/>
6. TV Corpus (TVC). Online <https://english-corpora.org/tv/>