# A Comparative Analysis of Different to and Different from in Three Corpora

Namkil Kang
*Far East University*
*South Korea*

The main goal of this paper is to compare *different to* and *different from* in the Corpus of Contemporary American English, the British National Corpus, and the TV Corpus. In the COCA, *different from* was always preferred over *different to* by Americans from 1990 to 2019. With respect to *different to*, it had the highest frequency (209 tokens) in the period from 2015 to 2019. When it comes to *different from*, it had the highest frequency (3,422 tokens) in the period from 1990 to 1994. In the BNC, *different to* and *different from* show a different pattern in five genres, whereas they show the same pattern in two genres. That is, *different to* is 28.57% the same as *different from*. With respect to the BNC, it is interesting to note that *different to* is the nearest to *different from* in the misc genre. The TV Corpus clearly shows that *different from life* and *different from earth* are the most preferred ones in the TV programs of America, the UK, Canada, Australia, New Zealand, and Ireland. In addition, the expression *different to others* is the most preferred by six countries' celebs. More importantly, it is worth mentioning that 13.15% of thirty eight nouns are the collocations of both *different from* and *different to*. Interestingly, the COCA shows that the expression *different from men* and *different to me* are the most preferable ones for Americans. Finally, it is worth noting that 6.38% of forty seven nouns are the collocations of both *different from* and *different to*.

**KEYWORDS:** TV Corpus, COCA, BNC, type, token, different from, different to

## I. INTRODUCTION

The main purpose of this paper is to compare *different to* and *different from* in the Corpus of Contemporary American English (COCA), the British National Corpus (BNC), and the TV Corpus (TVC). As Murphy (2016, 2019) points out, *different to* and *different from* are used interchangeably. First, we consider which type is the preferable one for Americans and the British. Second, we provide the diachronic aspects of *different to* and *different from* in the COCA. Third, we observe similarities between *different to* and *different from* in the seven genres of the BNC. Fourth, we examine similarities between *different to* and *different from* in the TV Corpus and the COCA. The organization of this paper is as follows. In section 2, we argue that *different from* was always preferred over *different to* by Americans from 1990 to 2019. With respect to *different to*, it had the highest frequency (209 tokens) in the period from 2015 to 2019. When it comes to *different from*, it had the highest frequency (3,422 tokens) in the period from 1990 to 1994. In section 3, we argue that *different to* and *different from* show a different pattern in five genres, whereas they show the same pattern in two genres. That is, *different to* is 28.57% the same as *different from*. Additionally, it is interesting to note that *different to* is the nearest to *different from* in the misc genre. In section 4, we contend that *different from life* and *different from earth* are the most preferred ones in the TV programs of America, the UK, Canada, Australia, New Zealand, and Ireland. Additionally, we maintain that the expression *different to others* is the most preferred by six countries' celebs. In this section, we also argue that 13.15% of thirty eight nouns are the collocations of both *different from* and *different to*. In section 5, we further argue that the expression *different from men* and *different to me* are the most preferable ones for Americans. We also show in this section that 6.38% of forty seven nouns are the collocations of both *different from* and *different to*.

## II. THE DIACHRONIC ASPECTS OF DIFFERENT TO AND DIFFERENT FROM IN THE COCA

In this section, we provide the diachronic aspects of *different to* and *different from* in the COCA. Table 1 indicates the frequency of *different to* and *different from* from 1990 to 2019:
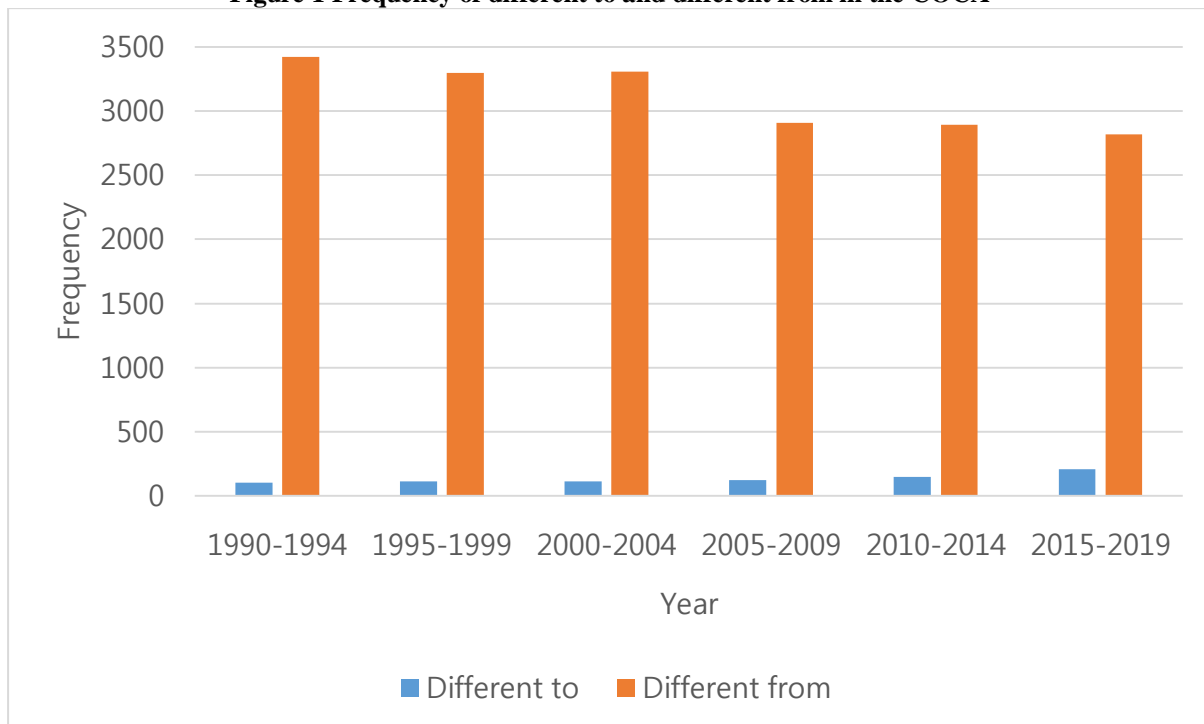
**Table 1 Frequency of different to and different from in the COCA**

| Type | Different to | Different from |
|------|-------------|---------------|
| 1990-1994 | 104 | 3,422 |
| 1995-1999 | 114 | 3,298 |
| 2000-2004 | 111 | 3,309 |
| 2005-2009 | 122 | 2,908 |
| 2010-2014 | 147 | 2,892 |
| 2015-2019 | 209 | 2,816 |
| All | 807 | 18,645 |

It is worth noting that *different from* was the preferable one among Americans from 1990 to 2019. More specifically, the overall frequency of *different to* is 807 tokens, whereas that of *different from* is 18,645 tokens. This in turn suggests that *different from* was preferred over *different to* by Americans. The following graph indicates the frequency of *different to* and *different from* from 1990 to 2019:

**Figure 1 Frequency of different to and different from in the COCA**



It is significant to note that *different from* was always preferable to *different to* in America. More specifically, the frequency of *different from* is twenty three times higher than that of *different to*. From this it can be inferred that *different from* was favored over *different to* for twenty nine years.

It is interesting to point out that there was a slight increase in the frequency of *different to* from 1990 to 1999. There was a rise of 14 tokens from 1990 to 1999. It must be noted, however, that there was a slight decrease (a decrease of 3 tokens) in the frequency of *different to* from 2000 to 2004. It is worth noting that there was a gradual rise (an increase of 87 tokens) in the frequency of *different to*

from 2005 to 2019. It should be pointed out that the frequency of *different to* reached a peak (209 tokens) in the period from 2015 to 2019. This in turn implies that Americans preferred *different to* the most in that period. It should be noted, on the other hand, that *different to* had the lowest frequency (104 tokens) in the period from 1990 to 1994. This in turn suggests that *different to* was the least common type in that period.

It is worthwhile noting that the frequency of *different from* reached a peak (3,422 tokens) in the period from 1990 to 1994. From this it is clear that Americans preferred *different from* the most in that period. After the period, there was a steady decline (a

decrease of 606 tokens from 1994 to 2019) in the frequency of *different from* from 1995 to 2019. It must be pointed out that there was a rise of 11 tokens from 2000 to 2004, compared to the previous period (from 1995 to 1999). It is interesting to note that *different from* had the lowest frequency (2,816 tokens) in the period from 2015 to 2019. This in turn shows that *different from* was the least common type in the period. It is noteworthy that *different from* was always preferred over *different to* by Americans from 1990 to 2019. We thus conclude that Americans preferred using *different from* to using *different to*

from 1990 to 2019.

## III. DIFFERENT TO AND DIFFERENT FROM IN THE BRITISH NATIONAL CORPUS

In the following, we compare *different to* and *different from* in the seven genres of the BNC and examine similarities between them in terms of the Euclidean distance. Table 2 indicates the genre frequency of *different to* and *different from*:

**Table 2 Genre Frequency of different to and different from in the BNC**

| GENRE | ALL | SPOKEN | FICTION | MAGAZINE | NEWSPAPER | NON-ACAD | ACADEMIC | MISC |
|---|---|---|---|---|---|---|---|---|
| **Different to** | 474 | 133 | 44 | 47 | 36 | 54 | 57 | 103 |
| **Different from** | 3,246 | 178 | 379 | 128 | 120 | 689 | 1,036 | 716 |

It is significant to note that *different to* ranks first (133 tokens) in the spoken genre, whereas *different from* ranks first (1,036 tokens) in the academic genre. *Different from* and *different to* show a different pattern with respect to rank-one, thus showing that they are low similarity synonyms. It should be emphasized that *different from* is preferred over *different to* by the British in the spoken genre. This is derived from the fact that the frequency of *different from* is much higher than that of *different to* in the spoken genre. This in turn indicates that the British prefer using *different from* rather than using *different to* in daily conversation. It should be noted, however, that the frequency of *different from* is nineteen times (1,036 tokens vs. 57 tokens) higher than that of *different to* in the academic genre. From this it is clear that the British prefer using *different from* to using *different to* in the academic genre.

It is interesting to point out that *different to* and *different from* rank second (103 tokens vs. 716 tokens) in the misc genre. Interestingly, *different to* and *different from* show the same property in the misc genre, hence the same pattern. It must be pointed out, however, that the frequency of *different from* is six times (103 tokens vs. 716 tokens) higher than that of *different to*. We take this as indicating that *different from* is preferable to *different to* in the misc genre.

It is probably worthwhile pointing out that *different to* ranks third (57 tokens) in the academic genre, whereas *different from* ranks third (689 tokens) in the non-academic genre. Interestingly, *different to* and *different from* show a different pattern in the

academic and non-academic genres, hence low similarity synonyms. It should be noted that *different from* is preferable to *different to* in the non-academic genre.

Noteworthy is that *different to* ranks fourth (54 tokens) in the non-academic genre, whereas *different from* ranks fourth (379 tokens) in the fiction genre. Again, *different to* and *different from* do not show the same pattern with respect to rank–four, thus implying that *different to* and *different from* are low similarity synonyms. As illustrated in Table 2, *different from* is preferred over *different to* by British journalists in the fiction genre. This is derived from the fact that the frequency of *different from* is seven times (54 tokens vs. 379 tokens) higher than that of *different to* in the fiction genre.

It is worth observing that *different to* ranks fifth (47 tokens) in the magazine genre, whereas *different from* ranks fifth (178 tokens) in the spoken genre. Again, *different to* and *different from* show a different pattern with respect to rank-five, hence low similarity synonyms again. It must be pointed out that in the magazine genre, *different from* is preferred over *different to* by British journalists.

It is worth mentioning that *different to* ranks sixth (44 tokens) in the fiction genre, whereas *different from* ranks sixth (128 tokens) in the magazine genre. Again, *different to* and *different from* do not show the same pattern with respect to rank-six. It is noteworthy that *different to* and *different from* rank seventh (36 tokens vs. 120 tokens) in the newspaper genre, hence the same pattern. It should be pointed out, however, that the frequency of
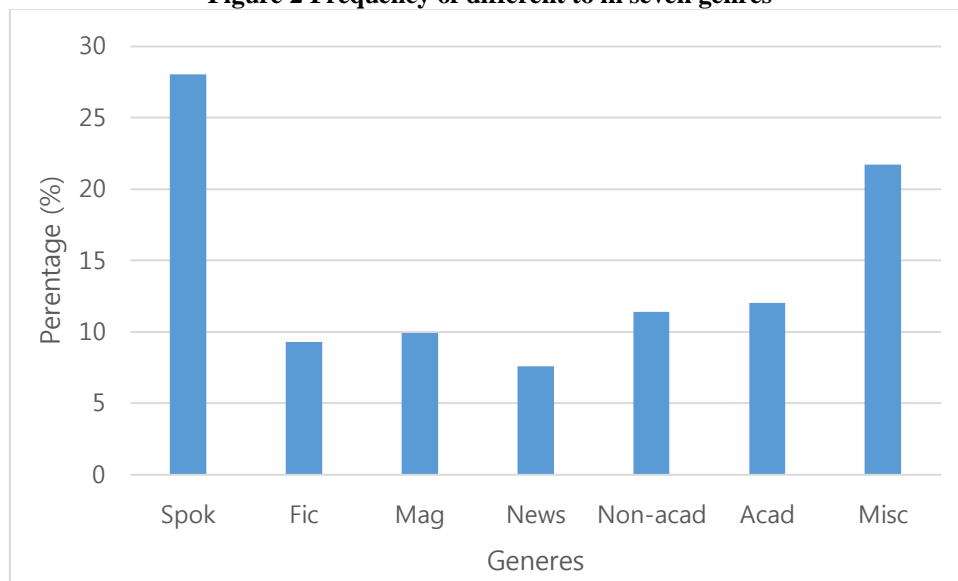
*different from* (120 tokens) is much higher than that of *different to* (36 tokens). It can thus be inferred that *different from* is favored over *different to* in the newspaper genre. To sum up, *different to* and *different from* show a different pattern in the fiction and magazine genres and the spoken, non-academic, and academic genres, whereas they show the same pattern in the newspaper genre and the misc genre. This in turn indicates that *different to* is 28.57% the same as *different from,* hence low similarity synonyms.

Now attention is paid to the frequency of *different to* and *different from* in seven genres:
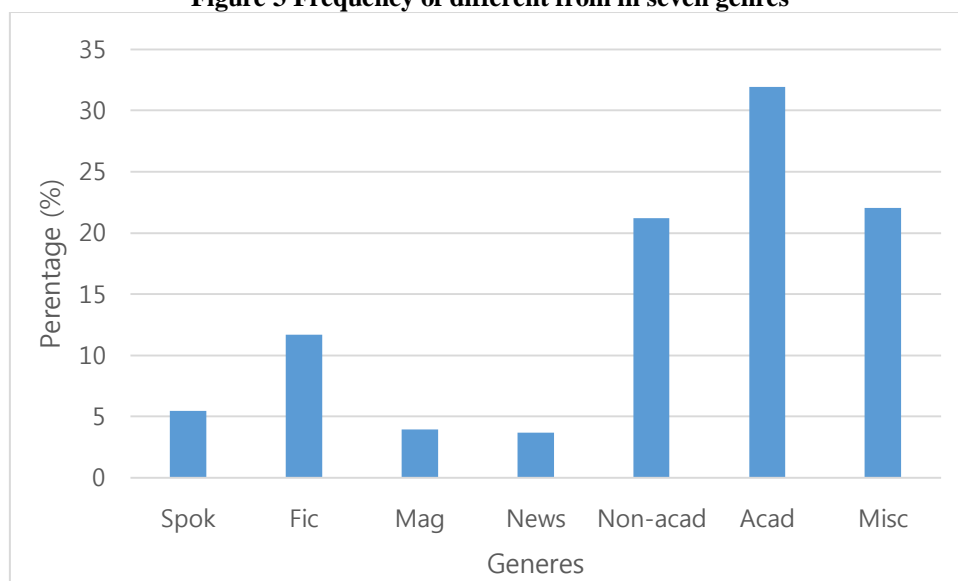
**Figure 2 Frequency of different to in seven genres**



It is interesting to point out that as indicated in Figure 2, the spoken genre is the most influenced by *different to*, followed by the misc genre, the academic genre, the non-academic genre, the magazine genre, the fiction genre, and the newspaper genre, in that order.

**Figure 3 Frequency of different from in seven genres**

It is worth mentioning that as illustrated in Figure 3, the academic genre is the most influenced by *different from*, followed by the misc genre, the non-academic genre, the fiction genre, the spoken genre, the magazine genre, and the newspaper genre, in descending order.

(1) The Euclidean distance:

$$d(p, q) = \sqrt{\sum_{i=1}^{n} (p_i - q_i)^2}$$

Finally, let us turn our attention to the Euclidean distance. In what follows, we consider similarities between *different to* and *different from* in terms of the Euclidean distance. With respect to the Euclidean distance, we adopt the following definition:

**Table 3 Euclidean distance in seven genres**

| GENRE | Spok | Fic | Mag | News | Non-acad | Acad | Misc |
|---|---|---|---|---|---|---|---|
| Percentage of different to | 28.05 | 9.28 | 9.91 | 7.59 | 11.39 | 12.02 | 21.72 |
| Percentage of different from | 5.48 | 11.67 | 3.94 | 3.69 | 21.22 | 31.91 | 22.05 |
| Euclidean distance | 22.57 | 2.47 | 5.97 | 3.9 | 9.83 | 19.89 | 0.33 |

Most importantly, *different to* is the furthest from *different from* in the spoken genre. On the other hand, *different to* is the nearest to *different from* in the misc genre. This in turn indicates that *different to* and *different from* show deep similarities in the misc genre. Interestingly, the Euclidean distance in the fiction genre is 2.47, which is the second lowest. This in turn suggests that *different to* and *different from* show a high degree of similarity in the fiction genre. Conversely, the Euclidean distance in the academic genre is 19.89, which is the second highest. This in turn implies that *different to* and *different from* show a low degree of similarity. We thus conclude that

*different to* is the nearest to *different from* in the misc genre.

## IV. A COLLOCATION ANALYSIS OF DIFFERENT TO AND DIFFERENT FROM IN THE TV CORPUS

In what follows, we compare *different from* and *different to* in the TV Corpus and examine similarities between them. Table 4 indicates the collocation of *different from* in the TV Corpus:

**Table 4 Collocation of different from in the TV Corpus**

| Number | Collocation of different from | Frequency |
|---|---|---|
| 1 | Different from life | 8 |
| 2 | Different from earth | 8 |
| 3 | Different from home | 5 |
| 4 | Different from others | 5 |
| 5 | Different from humans | 4 |
| 6 | Different from animals | 3 |
| 7 | Different from women | 3 |
| 8 | Different from hell | 2 |
| 9 | Different from bein | 2 |
| 10 | Different from men | 2 |
| 11 | Different from mother | 2 |
| 12 | Different from people | 2 |
| 13 | Different from Apple | 2 |
| 14 | Different from stars | 2 |
| 15 | Different from things | 2 |
| 16 | Different from training | 2 |
| 17 | Different from Viking | 2 |

| 18 | Different from appetizers | 1 |
|----|--------------------------|---|
| 19 | Different from porn | 1 |
| 20 | Different from populations | 1 |
| 21 | Different from politics | 1 |
| 22 | Different from police | 1 |
| 23 | Different from pesticides | 1 |
| 24 | Different from person | 1 |
| 25 | Different from parties | 1 |

As indicated in Table 4, *different from life* and *different from earth* are the most frequently used ones in the TV programs of America, the UK, Canada, Australia, New Zealand, and Ireland. This in turn suggests that the expressions *different from life* and *different from earth* are the most widely used by six countries' celebs. As alluded to in Table 4, *different from life* and *different from earth* are the most commonly used by six countries' celebs, followed by *different from home* (*different from others*), *different from humans*, and *different from* animals (*different from women*), in that order. More importantly, *different from life* and *different from earth* rank first (8 tokens) in the TV Corpus. This in turn implies that these expressions are the most preferred by six countries' celebs. Interestingly, *different from humans* is the second most preferred one in the TV Corpus. We thus conclude that *different from life* and *different from earth* are the most frequently used ones in the TV Corpus.

Now attention is paid to the collocation of *different to* in the TV Corpus:

**Table 5 Collocation of different to in the TV Corpus**

| Number | Collocation of different to | Frequency |
|--------|-----------------------------|-----------|
| 1 | Different to others | 3 |
| 2 | Different to women | 2 |
| 3 | Different to earth | 2 |
| 4 | Different to beer | 1 |
| 5 | Different to adults | 1 |
| 6 | Different to vampires | 1 |
| 7 | Different to sunshine | 1 |
| 8 | Different to smash | 1 |
| 9 | Different to radio | 1 |
| 10 | Different to people | 1 |
| 11 | Different to men | 1 |
| 12 | Different to marriage | 1 |
| 13 | Different to management | 1 |
| 14 | Different to lord | 1 |
| 15 | Different to kids | 1 |
| 16 | Different to geology | 1 |
| 17 | Different to gatehouses | 1 |
| 18 | Different to French | 1 |

It is significant to note that *different to others* is the most frequently used one in the TV programs of America, the UK, Canada, Australia, New Zealand, and Ireland. This in turn suggests that the expression *different to others* is the most preferred by six countries' celebs. As alluded to in Table 5, *different to others* is the most frequently used by six countries' celebs, followed by *different to women* (*different to earth*), and *different to beer*, in that order. It is worth mentioning that *different to wo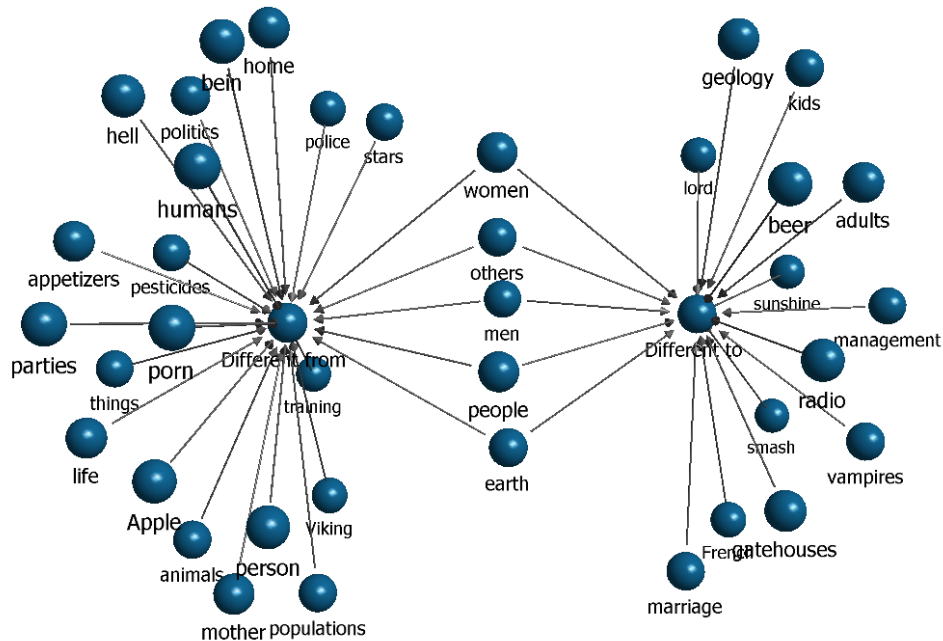men* ranks second in the TV Corpus, whereas *different from women* ranks sixth. It is also worth pointing out that *different to earth* ranks second in the TV Corpus, whereas *different from earth* ranks first. More interestingly, *different to people* ranks fourth in the TV Corpus, whereas *different from people* ranks eighth. We thus conclude that *different to others* is the most preferred by six countries' celebs.

Now let us turn to the visualization of the collocations of *different from* and *different to* in the TV Corpus:

**Figure 4 Visualization of the collocations of different from and different to in the TV Corpus**



As exemplified in Figure 4, twenty nouns are linked to *different from*, whereas five nouns are linked to both *different from* and *different to*. On the other hand, thirteen nouns are linked to *different to*, whereas five nouns are linked to both *different from* and *different to*. Most importantly, five nouns are linked to both *different from* and *different to*, which indicates that these five nouns are the collocations of two types. The nouns linked to the two types are *women*, *earth*, *people*, *others*, and *men*. This in turn indicates that 13.15% of thirty eight nouns are the collocations of both *different from* and *different to*. It can thus be inferred that *different from* and *different* *to* are low similarity synonyms in the TV Corpus.

## V. A COLLOCATION ANALYSIS OF DIFFERENT FROM AND DIFFERENT TO IN THE COCA

In the following, we compare the collocation of *different from* and that of *different to* in the COCA. Also, we examine similarities between *different from* and *different to* in the top 25. Table 6 indicates the collocation of *different from* in the COCA:

**Table 6 Collocation of different from in the COCA**

| Number | Collocation of different from | Frequency |
|---|---|---|
| 1 | Different from men | 72 |
| 2 | Different from people | 40 |
| 3 | Different from earth | 28 |
| 4 | Different from president | 23 |
| 5 | Different from adults | 22 |
| 6 | Different from humans | 21 |
| 7 | Different from life | 20 |
| 8 | Different from women | 19 |
| 9 | Different from home | 15 |
| 10 | Different from mine | 13 |

| 11 | Different from girls | 12 |
|----|---------------------|----|
| 12 | Different from Mr | 12 |
| 13 | Different from State | 12 |
| 14 | Different from control | 11 |
| 15 | Different from English | 11 |
| 16 | Different from person | 10 |
| 17 | Different from place | 10 |
| 18 | Different from adult | 8 |
| 19 | Different from chance | 8 |
| 20 | Different from students | 8 |
| 21 | Different from Americans | 7 |
| 22 | Different from boys | 7 |
| 23 | Different from things | 7 |
| 24 | Different from back | 6 |
| 25 | Different from business | 6 |

It is important to note that *different from men* is the most frequently used one (72 tokens) in America. This in turn implies that the expression *different from men* is the most preferable one for Americans. As illustrated in Table 6, *different from men* is the most preferred by Americans, followed by *different from people*, *different from earth*, *different from president*, *different from adults*, *different from humans*, and *different from life*, in that order. It is worth mentioning that *different from adults* ranks fifth (22 tokens) in the COCA. It must be pointed out, on the other hand, that *different from life* ranks seventh (20 tokens) in the COCA. More interestingly, the everyday expression *different from mine* ranks tenth (13 tokens) in the COCA. We thus conclude that *different from men* is the most widely used one (72 tokens) in America.

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

**Table 7 Collocation of different to in the COCA**

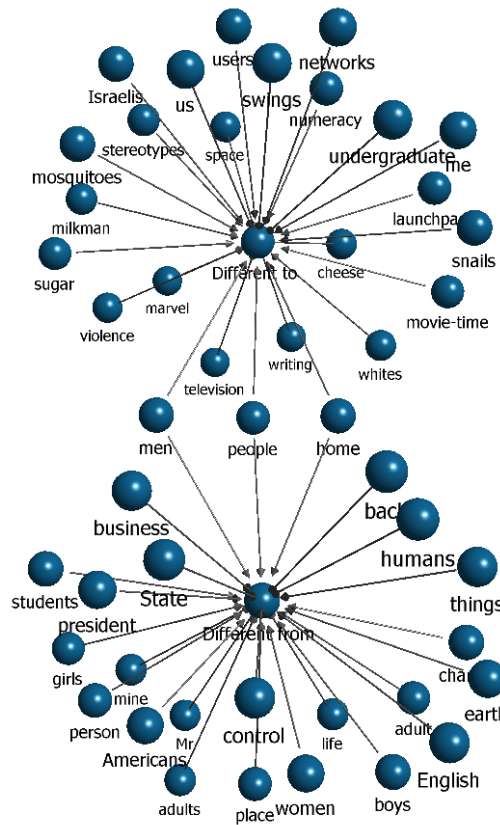| Number | Collocation of different to | Frequency |
|--------|----------------------------|-----------|
| 1 | Different to me | 9 |
| 2 | Different to people | 6 |
| 3 | Different to cheese | 2 |
| 4 | Different to home | 2 |
| 5 | Different to numeracy | 2 |
| 6 | Different to networks | 1 |
| 7 | Different to movie-time | 1 |
| 8 | Different to mosquitoes | 1 |
| 9 | Different to milkman | 1 |
| 10 | Different to men | 1 |
| 11 | Different to marvel | 1 |
| 12 | Different to launchpads | 1 |
| 13 | Different to Israelis | 1 |
| 14 | Different to writing | 1 |
| 15 | Different to whites | 1 |
| 16 | Different to violence | 1 |
| 17 | Different to users | 1 |
| 18 | Different to us | 1 |
| 19 | Different to undergraduate | 1 |
| 20 | Different to television | 1 |
| 21 | Different to swings | 1 |
| 22 | Different to sugar | 1 |
| 23 | Different to stereotypes | 1 |
| 24 | Different to space | 1 |

| 25 | Different to snails | 1 |

It is significant to note that *different to me* is the most commonly used one (9 tokens) in America. As alluded to in Table 7, the expression *different to me* is the most preferred by Americans, followed by *different to people*, *different to cheese* (*different to home*, *different to numeracy*), and *different to networks*, in descending order. It is worth observing that *different to people* (6 tokens) and *different from people* (40 tokens) rank second in the COCA. From this, it is evident that Americans prefer using *different from people* to using *different to people*. It is worthwhile noting that *different to home* ranks third in the COCA, whereas *different from home* ranks ninth. We thus conclude that *different to me* and *different from men* are the most preferred by Americans.

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

**Figure 5 Visualization of the collocations of different from and different to in the COCA**



As shown in Figure 5, twenty two nouns are linked to *different from*, whereas three nouns are linked to both *different from* and *different to*. On the other hand, twenty two nouns are linked to *different to*, whereas three nouns are linked to both *different from* and *different to*. This indicates that three nouns are the collocations of both *different from* and *different to*. These three nouns are *people*, *home*, and *men*. This in turn suggests that 6.38% of forty seven nouns are the collocations of both *different from* and *different to*. It can thus be inferred that *different from* and *different to* are low similarity synonyms.

## VI.    CONCLUSION

To sum up, we have compared *different to* and *different from* in the COCA, BNC, and MC. In section 2, we have argued that *different from* was always preferred over *different to* by Americans from 1990 to 2019. We have further argued that the frequency of *different to* reached a peak (209 tokens) in the period from 2015 to 2019, whereas the frequency of *different from* reached a peak (3,422 tokens) in the period from 1990 to 1994. In section 3, we have shown that *different to* and *different from*

show a different pattern in five genres, whereas they show the same pattern in two genres. That is, *different to* is 28.57% the same as *different from*. We have also shown that *different to* is the nearest to *different from* in the misc genre. In section 4, we have contended that *different from life* and *different from earth* are the most frequently used ones in the TV programs of America, the UK, Canada, Australia, New Zealand, and Ireland. We have also contended that the expression *different to others* is the most preferred by six countries' celebs. Additionally, we have shown that 13.15% of thirty eight nouns are the collocations of both *different from* and *different to*. In section 5, we have argued that the expression *different from men* and *different to me* are the most preferable ones for Americans. Finally, we have shown that 6.38% of forty seven nouns are the collocations of both *different from* and *different to*.

## REFERENCES

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