

# Development of the Subject of BE *going to* in Grammaticalisation from the 1820s to 2010s in Comparison with BE *about to*

Michiko Yaguchi

*Institute of Arts and Science, Kanazawa University, Kanazawa, Japan*

Contact Michiko Yaguchi (yaguti58mitiko@gmail.com)

## **Abstract**

This study clarifies the development of the subject of BE *going to* in relation to its increase in frequency from the 1820s to the 2010s by analysing the Corpus of Historical American English. To effectively demonstrate this, BE *about to* is also analysed. Whilst the use of BE *going to* in the passive and *there* constructions is attested to in the early stage of its grammaticalisation towards becoming a semi-auxiliary, this study suggests that there is a possibility that its use with the *what* subject was also a sign of initial stage of grammaticalisation. It will be also shown that subjects with simple content appeared earlier than content-rich ones. This study also demonstrates that the frequency order of these subjects in current English corresponds to the chronological order of when the earliest token appeared in the corpus.

Keywords: BE *going to*, BE *about to*, grammaticalisation, subject, semi-auxiliary

## 1. Introduction

The development of the semi-auxiliary BE *going to*, a future marker, has attracted tremendous attention in both theory and corpus-based attestation since grammaticalisation established its status in the field of historical linguistics (e.g. Pertejo,<sup>1</sup> Traugott and Dasher,<sup>2</sup> Mair,<sup>3</sup> Bybee,<sup>4</sup> Fischer,<sup>5</sup> Hilpert,<sup>6</sup> Garrett,<sup>7</sup> Traugott,<sup>8</sup> Traugott and Trousdale,<sup>9</sup> Budts and Petré,<sup>10</sup> inter alia). It underwent a significant change in terms of frequency at the turn of the twentieth century. There are several theoretical hypotheses to account for its longitudinal developmental path. For instance, Bybee<sup>11</sup> focuses on semantic change: BE *going to* evolved from the original meaning of *go*, i.e. ‘movement with purpose’, to ‘intention without movement’, then ‘intention with future’ and finally to ‘future without intention’. In another instance, Traugott and

---

<sup>1</sup> Pertejo, “*Be Going To + Infinitive*”.

<sup>2</sup> Traugott and Dasher, *Regularity in Semantic Change*.

<sup>3</sup> Mair, “Corpus Linguistics and Grammaticalization Theory”.

<sup>4</sup> Bybee, “From Usage to Grammar”.

<sup>5</sup> Fischer, *Morphosyntactic Change* and “An Analogical Approach to Grammaticalization”.

<sup>6</sup> Hilpert, *Germanic Future Constructions*.

<sup>7</sup> Garrett, “The Historical Syntax Problem”.

<sup>8</sup> Traugott, “On the Persistence of Ambiguous Linguistic Contexts over Time”.

<sup>9</sup> Traugott and Trousdale, *Constructionalization and Constructional Changes*, 94-148.

<sup>10</sup> Budts and Petré, “Reading the Intentions of *Be Going To*”.

<sup>11</sup> Bybee, “From Usage to Grammar”.

Trousdale<sup>12</sup> propose a hypothesis from syntactic and semantic points of view.

Syntactically, as the verb *go* was fused with the progressive construction BE *V-ing*, the purposive construction  $V_{\text{activity}}$  to  $V_{\text{activity}}$  and the passive construction BE *V-en*, BE *going to* became a fixed chunk in the seventeenth century. Along with the syntactic fusion, since the passive complement came to be allowed, selective restrictions on the VP, whose meaning was originally limited to motion-related activity, weakened, which also triggered further bleaching of the meaning of purpose. Semantically, the phrase first denoted an immediate future after a given point, then acquired a deictic future; afterwards, the purpose connotation disappeared in the eighteenth century. It should be noted that they indicate that the phrase's occurrence in passive constructions was a catalyst for further change that diluted the original sense of *go*. Indeed, Garrett<sup>13</sup> maintains that grammaticalising of BE *going to* toward a semi-auxiliary can be confirmed by its function in passive constructions. Furthermore, Traugott and Trousdale<sup>14</sup> argue that its constructionalisation was established in the eighteenth century when it started to be used in *there* constructions.

---

<sup>12</sup> Traugott and Trousdale, *Constructionalization and Constructional Changes*, 94-148.

<sup>13</sup> Garrett, "The Historical Syntax Problem".

<sup>14</sup> Traugott and Trousdale, *Constructionalization and Constructional Changes*, 117-18.

This study will explore change from the nineteenth century to the beginning of the twenty-first century by concentrating on the subject of BE *going to* against the background that previous studies primarily focus on phenomena related to change in the infinitive part including verbs or sentence types. Since the subject of the verb *go* moves or has purpose or intention as mentioned above, it is reasonable to assume that in the early stage of its development toward a grammaticalised phrase, the animate subject was used as the default before the restrictions of subjects disappeared as the original sense of *go* bleached. It can be expected that employing a large-sized corpus will offer extensive data concerning the development of subject types. Hence, this study will elucidate how the subject expanded and how its development is interrelated with frequency in the process of grammaticalisation. To comprehensively illustrate the development, BE *going to* will be compared with BE *about to*, another future marker, whose subject was originally animate as well.

This study will analyse the Corpus of Historical American English (COHA), which compiles 475 million words of 200 years of data in five subcorpora: fiction and non-fiction books (academic), newspapers, magazines, and TV/movies between the 1820s and 2010s. Also, grammaticalisation is defined as “the process whereby lexical items and constructions come in certain linguistic contexts to serve grammatical

functions, and, once grammaticalised, continue to develop new grammatical functions”, following Hopper and Traugott.<sup>15</sup> This study distinguishes present-day English as English since 1900, contemporary English as English after 1950, and current English as English after 2000.

## **2. Previous Studies**

BE *going to* and BE *about to* share two points regarding their diachronic developments.

First, they took only animate subjects in their origin and then the characteristics of the subject changed greatly in Late Modern English. Second, they demonstrated a sudden increment in frequency at a certain time in their development. By reviewing previous studies, this section overviews the diachronic development of the two semi-auxiliaries.

### ***2.1 Overview of Diachronic Development***

#### ***2.1.1 BE going to***

Several previous studies point out that the sequence of BE + *going to* as a grammaticalised phrase emerged around 1500. According to Pertejo,<sup>16</sup> many scholars

---

<sup>15</sup> Hopper and Traugott, *Grammaticalization*, xv.

<sup>16</sup> Pertejo, “*Be Going To + Infinitive*”.

accept the following example of the year 1482 as in (1a) cited by the *Oxford English Dictionary* (OED) as one of the earliest tokens showing grammaticalising features because it no longer connotes purpose or intention of the subject because of passive construction:

- (1) a. ...thy unhappy sowle by the vycторыse pompys of her enmyes was goying to be broughte into helle for the synne and onleful lustys of her body.

(c. 1482, Monk of Evesham, *Revelation*)

- b. So, for want of a Cord, hee tooke his owne garters off; and as he was going to make a nooze, I watch'd my time and ranne away.

(1611, Tourneur, *The Atheist's Tragedie*)

- c. He is fumbling with his purse-strings, as a school-boy with his point when he is going to be whipped, till the master weary with long stay forgives him.

(1628, Earle, *Microcosmography* §19)

- d. I am afraid there is going to be such a calm among us, that...

(1725, Odingsells, *The Bath Unmask'd*)

Meanwhile, according to Traugott and Trousdale,<sup>17</sup> who examined the context of (1a), the intention of the subject ‘*thy unhappy sowle*’ is still expressed. They show the token of (1b) in 1611 presented by Garret<sup>18</sup> as one of the earliest possible examples of BE *going to* expressing a temporal sense. They add a passive token of (1c) in 1628 expressing a temporal sense, cited by Garret.<sup>19</sup> If we follow their contention, grammaticalisation can be considered to have started in the beginning of the seventeenth century. They also show a *there* construction in the year 1725 as in (1d), as an early example to evince the constructionalisation of the phrase, as mentioned in the introduction. It deserves our attention that there is 100 years in gap between the passive token and the *there* construction token.

It is important to review the diachronic transition of frequency. By investigation of quotation texts in the OED, Mair<sup>20</sup> shows that BE *going to* was used sporadically until the 1700s, and then, the frequency began to increase suddenly around 1900, maintaining a constant augmentation until the present. Mair provides Figure 1, showing the sudden increase in frequency in the 1901–1925 period:

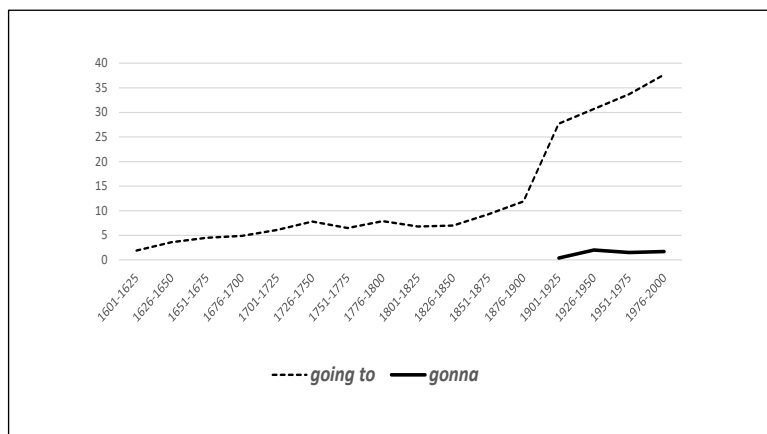
---

<sup>17</sup> Traugott and Trousdale, *Constructionalization and Constructional Changes*, 217-24.

<sup>18</sup> Garrett, “The Historical Syntax Problem”, 69.

<sup>19</sup> Ibid.

<sup>20</sup> Mair, “Corpus Linguistics and Grammaticalization Theory”.



**Figure 1.** *Going to* and *gonna* 1600-2000. (occurrences per 10,000 quotation texts)

(excerpted from Mair<sup>21</sup>)

Several previous studies have analysed various aspects of the development quantitatively. For instance, Hilpert<sup>22</sup> clarifies the transition of verbs employed in the infinitive part through the investigation of Corpus of Late Modern English Texts version 3.0 (CLMET 3.0). His data show that frequently used verb types dramatically changed, from telic, dynamic verbs (e.g. *fight*, *publish*, *answer*, *observe* and *embrace*) in the 1710–1780 period to general, light verbs (e.g. *be*, *do*, *get*, *have*, *try* and *die*) in the 1850–1920 period. Thus, he argues that the data verify the hypothesis that the main meaning and function of BE *going to* shifted from intentional movement to future

---

<sup>21</sup> Mair, “Corpus Linguistics and Grammaticalization Theory”, 129.

<sup>22</sup> Hilpert, *Germanic Future Constructions*, 106-23.



without movement between 1780 and 1850. Furthermore, Budts and Petré<sup>23</sup> investigated the CLMET 3.0, the Penn Parsed Corpus of Modern English and ECCO-TCP and found that BE *going to* showed a gradual but steady semantic and functional change in the six periods (every 35 years) between 1710 and 1920. They found that the ratio of tokens expressing imminent future dropped from 79% to 29%; sentence types other than statements, i.e. directive, conditional, *wh*-question and polarity questions, increased from 2% to 19%; and tokens expressing intention decreased from 90% to 81%. What is relevant to this study is that Budts and Petré<sup>24</sup> also show that the ratio of inanimate subjects increased from 4% to 10%. Thus, the quantitative data of the previous studies clearly illustrate that a great change in semantic function took place between 1710 and 1920. In connection with the data of Figure 1, it is possible to conclude that by the time a sudden increase occurred around 1900, semantic and functional change had proceeded gradually but significantly.

### 2.1.2 BE *about to*

In present-day English, BE *about to* expresses the immediate future. Unlike BE *going to*,

---

<sup>23</sup> Budts and Petré, “Reading the Intentions of *Be Going To*”.

<sup>24</sup> Ibid.

there are only a few previous studies, synchronic or diachronic, that exclusively analyse BE *about to* in the literature. The OED defines the origin of the phrase in Middle English as representing “engaged in or busied with plans or preparations to do something; planning, conspiring, or scheming to do something” (s.v. *about* adv. and prep., A11),<sup>25</sup> which intrinsically called for the use of animate subjects, and suggests that it began to express the future around the sixteenth century (s.v. *about* adv. and prep., A12).

Watanabe<sup>26</sup> also states that Pool, a grammarian, wrote in 1646 that *about to* is “About to, *or* going to, is thee signe of the Participle of the future”, which verifies the OED’s explanation. However, he counterargues this by analysing quotation texts in the OED (15 tokens in the 1751–1800 period and 129 tokens in the 1801–1850 period) and a corpus Watanabe<sup>27</sup> created by compiling data from drama, fiction and newspapers from the eighteenth century to the first half of the nineteenth century (317 occurrences in the 1700–1850 period) to make up for the insufficient number of the OED’s quotation texts. By examining each instance’s meaning in context and considering each novelist’s

---

<sup>25</sup> Watanabe notes that although the OED explains that usage to express the subject’s intention is obsolete, it is still available in present-day English. Also see the discussion in 2.2.

<sup>26</sup> Watanabe, “On the Development of the Immediate Future Use of *Be About To*”, 62.

<sup>27</sup> Ibid.

writing style, Watanabe<sup>28</sup> proposes that whilst from Middle English to the eighteenth century the semi-auxiliary primarily expressed intention of animate subjects, it is not until the turn of the nineteenth century that it began to express the immediate future, freely occurring with inanimate subjects and less agentive or stative verbs, and use in passive constructions, regardless of writers and their style.<sup>29</sup> Indeed, in his OED's data, all subjects were animate until 1800 (before 1800, there were three inanimate tokens, whose subjects were personified) and the earliest token of an inanimate subject was attested to in 1804. He also demonstrates that the first tokens of passive construction and less agentive, light verbs (e.g. *I was about to be angry*) appeared in 1794 and 1779 respectively in the OED. Additionally, his corporal data exhibit that in drama and fiction, some instances expressed the subject's intention and/or willingness until 1800 and other instances expressed both intention and future simultaneously until the early 1800s.<sup>30</sup> Observe the following graph by Watanabe<sup>31</sup> showing the transition of frequency of the OED's quotation texts:

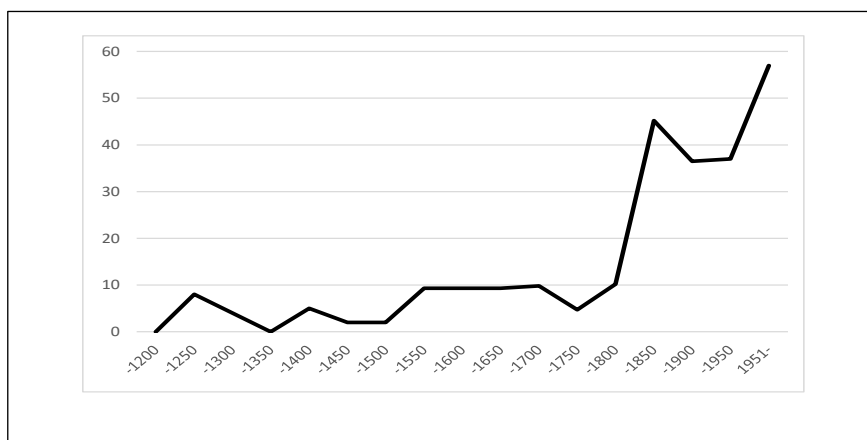
---

<sup>28</sup> Ibid.

<sup>29</sup> Watanabe indicates that the use differed according to novelists and script writers in his corpus until the end of the eighteenth century.

<sup>30</sup> However, Watanabe also reports that the inanimate subject and its use in passive constructions were already observable in newspapers of his corpus in the early 1700s.

<sup>31</sup> Watanabe, "On the Development of the Immediate Future Use of *Be About To*", 66.



**Figure 2.** Frequency of BE *about to* in OED. (occurrences per 100,000 quotation texts)

(excerpted from Watanabe<sup>32</sup>)

It is notable that the frequency showed a sudden increase around 1800. In Watanabe's<sup>33</sup> corpus as well, it suddenly doubled or tripled (depending on the genre) before and after 1800 in an identical way as demonstrated in Figure 2.<sup>34</sup> He makes a reference to the argument by Hopper and Traugott<sup>35</sup> that such a sudden rise in frequency is a phenomenon signifying grammaticalisation. Thus, he concludes that the semantic

---

<sup>32</sup> Ibid.

<sup>33</sup> Watanabe, "On the Development of the Immediate Future Use of *Be About To*".

<sup>34</sup> According to Watanabe's data, from the second half of the eighteenth century to the first half of the nineteenth century, the frequency increased from 0.2 to 0.5 occurrences per 100,000 words in newspapers, from 3.4 to 7.1 occurrences per 100,000 words in drama and from 3.5 to 10.2 occurrences per 100,000 words in fiction.

<sup>35</sup> Hopper and Traugott, *Grammaticalization*, 106, 126-130.

function as immediate future marker was fully entrenched around 1800, not in the sixteenth century as the OED demonstrates.

## ***2.2 Level of Grammaticalisation***

This subsection aims to review how much advanced the two items' grammaticalisations were by examining their separability. According to Hopper and Traugott,<sup>36</sup> grammaticalisation entails “semantic fading, phonological reduction, positional fixing, erasure of word boundaries”. Since semi-auxiliaries, per se, are defined as items that possess inseparability (cf. Quirk et al.<sup>37</sup>), the two semi-auxiliaries under discussion also show inseparable features in principle (cf. Westney<sup>38</sup>). A close evaluation of their grammatical behaviours in contemporary English, however, reveals some differences in the degree of grammaticalisation as unanalysable chunks. Indeed, Collins<sup>39</sup> argues that BE *about to* has a lower degree of grammaticalisation than BE *going to*. Wada<sup>40</sup> demonstrates their differences in separability as follows:

---

<sup>36</sup> Hopper and Traugott, *Grammaticalization*, 127.

<sup>37</sup> Quirk et al., *A Comprehensive Grammar of English Language*, 137.

<sup>38</sup> Westney, *Modals and Periphrastics in English*, 1-37.

<sup>39</sup> Collins, *Modals and Quasi-Modals in English*, 156.

<sup>40</sup> Wada, “*Be Going To* and *Be About To*”, 326-327.

- (2) a. \*and at times the sheer cliffs seemed going to close in.  
 b. and at times the sheer cliffs seemed about to close in.  
 (Brown Corpus)
- (3) a. \*Someone threw a beer bottle at me, and hit a guy who was behind me going to pour a pint over my head.  
 b. Someone threw a beer bottle at me, and hit a guy who was behind me about to pour a pint over my head. (BNC)
- (4) a. \*Maude swooped up the cup and hiked up her top hoop as is going to take off with a racing start.  
 b. Maude swooped up the cup and hiked up her top hoop as is about to take off with a racing start. (Brown Corpus)  
 (2)–(4) from Wada<sup>41</sup>

The difference in function as a complement of ‘*seem*’ in (2), the post-nominal use without *being* in (3) and the omission of a subject in an *as*-clause in (4) indicate that BE *about to* can function separably, in contrast to BE *going to*, which fails to do so. Additionally, the following examples also show that *going to* does not work independently in the same way as *about to*:

- (5) a. \*There is a man going to play the part of an old woman.

---

<sup>41</sup> Ibid.

(Heggie<sup>42</sup>)

b. There is a man about to play the part of an old woman.

(6) a. \*Hannah heard the Wilsons going to take a trip to Egypt.

(Kirsner and Thompson<sup>43</sup>)

b. Hannah heard the Wilsons about to take a trip to Egypt.

Unlike *about to* in (5b), *going to* cannot appear in a small clause of the *there* existential construction, and, as in (6a), it cannot be used as a complement of perception verbs, such as *hear* and *see*. The above examples all suggest that BE *going to* is more advanced in its degree of grammaticalisation. The presence of the shortened form *gonna* as a written form is revelatory evidence that shows fusion, in contrast to the lack of the written form *abouta*, etc.<sup>44</sup>

However, corpus data of current English still include tokens of *going to* in which the lexical function of *go* is in effect. This study manually retrieved *going to* and *about to* tokens showing separability in the 2010s' data of the COHA, as in (7):

---

<sup>42</sup> Heggie, *The Syntax of Copular Structures*, 27.

<sup>43</sup> Kirsner and Thompson, "The Role of Pragmatic Inference in Semantics", 207.

<sup>44</sup> In contemporary vernacular English, *gointa* is also used.

- (7) a. The other driver is halfway out his window, frozen in the motion of going to pull Chris from the burning vehicle.  
 (2010, magazine, COHA)
- b. Jordanian Palestinians see their tax dollars going to support job creation... (2012, non-fiction book, COHA)
- c. and gradually people are hearing about someone going to fight for World title... (2018, magazine, COHA)

In (7a), *going to* appears in a prepositional phrase. In (7b), it is used as a complement of the verb ‘*see*’ and in (7c) it modifies the preceding noun phrase. These grammatical behaviours show *go* functioning as a lexical item. In the same token, *about to* also appears independently from BE verbs in the 2010s’ data of the same corpus, as in (8):

- (8) a. Connie has never asked for anything before and seeing this young woman about to cry is more than Madeline can bear.  
 (2013, fiction, COHA)
- b. He seemed about to say more when he turned quickly, as if about to strike at an interloper... (2012, fiction, COHA)



Indeed, BE *about to* frequently display separable features, as in (2b), (3b), (4b), (5b), (6b) and (8). The present analysis found that out of 18,510 tokens of BE *going to*, 38 tokens (0.2%) show *go* functioning as a lexical verb, whilst 2,856 tokens of BE *about to* contain 222 separable tokens (7.8%). Thus, we can attest to lexical behaviours of *go* in BE *going to* in current English, although the percentage is very low, compared with *about to*. This finding provides evidence that BE *going to* is polysemous in that it expresses meanings ranging from its original lexical sense to the simple future. This naturally confirms the validity of the well-shared view that the motion sense of *go* representing purpose and intention is also in operation in current English (cf. Wada<sup>45</sup>).

It is significant to point out that BE *about to* is also polysemous in that it may denote the animate subject's intention. As Watanabe<sup>46</sup> maintains that some tokens in the early 1800s still ambiguously denote both the intention of the subject and future as mentioned above, we can presume that in current English, BE *about to* occasionally expresses the original sense. Observe the following tokens:

- (9) a. 'We're seriously about to jump out of this plane right now' ...'

---

<sup>45</sup> Wada, "Be Going To and Be About To", 191-245.

<sup>46</sup> Watanabe, "On the Development of the Immediate Future Use of Be About To".

(2019, magazine, COHA)

- b. Viv said she had been about to eat one but had instead made herself paint it for an hour while her stomach twinged. (2018, fiction, COHA)
- c. Since I did not consider my marriage a temporary situation, I was not about to settle for a temporary solution. (2012, fiction, COHA)

The example of (9a) indicates that the subject ‘*we*’ has the intention to ‘*jump out*’ due to the presence of ‘*seriously*’, and in (9b), ‘*Viv*’ had the intention to ‘*eat*’, because ‘*instead*’ connotes changing her intention. In these tokens, the subject’s intention is second to the sense of future. On the other hand, Swan<sup>47</sup> argues that in negative sentences, as in (9c), the subject’s willingness is expressed as the primary sense.<sup>48</sup> Thus, the subject-oriented original sense is still available in current English.

### 3. Transitions in frequency

Here, we grasp how the frequency of BE *going to*, BE *gonna* and BE *about to* have

---

<sup>47</sup> Swan, *Practical English Usage*, Section 31: 353.

<sup>48</sup> The OED explains that this usage is observable in North America (s.v. *about* adv. and prep., A11b).

changed since the 1820s in the COHA's data. Note that *BE gonna* is added in the present analysis. Before conducting the analysis, it is important to consider the impact of the subcorpus of TV/movies because it contains speech data basically whilst the other four subcorpora (i.e. fiction and non-fiction books, magazines and newspapers) compile only written data and because its data are available only after the 1930s in the COHA. Since the speech data is a different representation from written texts, it is essential to heed the data of the TV/movies subcorpus. Table 1 shows the frequencies of the three items' tokens in the 2010s in the five subcorpora.

**Table 1.** Frequency of BE *going to*, BE *gonna* and BE *about to* in the 2010s' COHA data.<sup>49</sup>

	fiction	magazine	news	non-fiction books	TV/movies
BE <i>going to</i>	570.4	360.5	584.7	64.9	846.2
BE <i>gonna</i>	28.0	9.3	2.4	0.9	400.7
BE <i>about to</i>	112.6	40.9	40.7	17.5	127

(occurrences per one million words)

As Table 1 shows, BE *going to*, BE *gonna* and BE *about to* mark a much higher frequency in TV/movies (846.2, 400.7 and 127.0 occurrences per one million words, respectively) than in the other subcorpora. Conversely, they are used the least frequently in non-fiction books. In sum, these future-oriented semi-auxiliaries are used in conversation more abundantly than in formal writing. This characteristic is especially prominent for BE *gonna*. Also, it is notable that BE *going to* is used more often than BE *gonna* in all genres, even in the speech data of TV/movies.

---

<sup>49</sup> The frequencies of BE *going to* and BE *about to* were calculated by manually examining all the tokens. On the other hand, the frequency of *gonna* was calculated by using the frequency yielded by randomly extracting 1000 tokens twice.

The present analysis is further complicated by TV/movies emerging only after the 1930s in the COHA, as referred to above. Table 2 shows the breakdown of word counts in the 2010s' data in the COHA according to the genre:

**Table 2.** Breakdown of word counts in the 2010s' data in COHA.

fiction	magazines	news	non-fiction books	TV/movies
43%	21%	13%	9%	14%

As Table 2 shows, TV/movies account for 14% of the total word count. It should also be mentioned that 10.5% to 17.3% in each decade's data comes from TV/movies since the 1930s in the COHA. Table 3 shows the distribution of tokens of BE *going to*, BE *gonna* and BE *about to* in the 2010s' data of the COHA according to the genre:

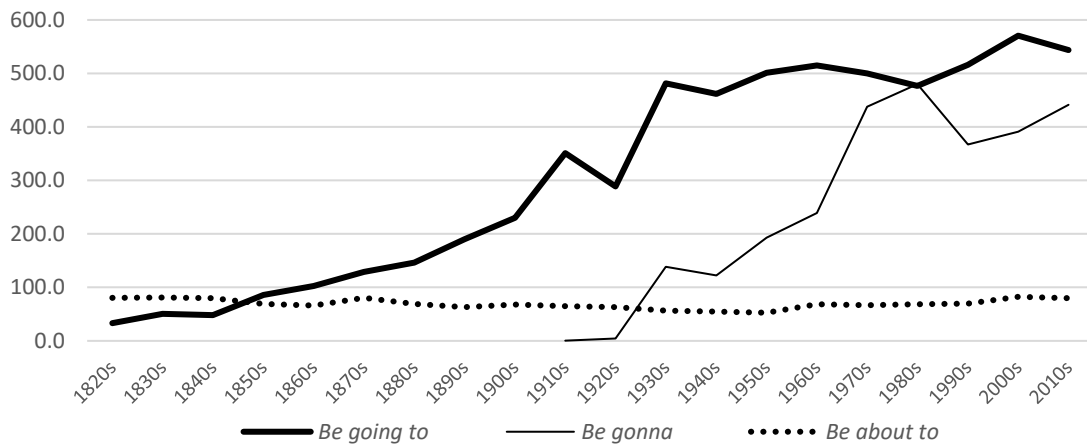
**Table 3.** Distribution of tokens of BE *going to*, BE *gonna* and BE *about to* in the 2010s' data in COHA.

	fiction	magazines	news	non-fiction books	TV/movies
BE <i>going to</i>	46.7%	14.7%	14.4%	1.1%	23.1%
BE <i>gonna</i>	6.4%	2.1%	0.6%	0.2%	90.8%
BE <i>about to</i>	58.9%	10.7%	6.4%	1.9%	22.2%

Table 3 exhibits that BE *gonna* shows an extremely unbalanced distribution: more than 90% of the tokens are used only in TV/movies, although TV/movies account for only 14% of the total word count of the 2010s' data, as seen in Table 2. In addition, 23.1% of BE *going to* tokens and 22.2% of BE *about to* tokens are used in TV/movies. Hence, it is necessary to acknowledge that this kind of speech data are included only from the 1930s. Observe that Figure 3, which shows the historical transition of frequency of the three items (also note Appendix 1 for detailed data):<sup>50</sup>

---

<sup>50</sup> The frequencies of BE *going to*, BE *gonna* and BE *about to* were yielded by applying *going to*, *gonna* and *about to* using a search function of the COHA. As far as BE *going to* and BE *about to* are concerned, by manually examining 200 randomly retrieved tokens each decade manually, the frequency was calculated after excluding the ratio of *going to*



**Figure 3.** Change in frequency in COHA.  
(occurrences per one million words)

Figure 3 demonstrates clear trends. First, the frequency of *BE going to* exhibited a steady increase. It remained on a 50-occurrences-per-million-words level until the 1840s and showed a constantly increasing tendency until the present. A sudden increase can be observed around 1900, reaching more than 500 occurrences per million since the 1930s. Also, after it first appeared in the 1910s in the COHA, *BE gonna* accelerated its frequency every decade, particularly after the 1930s, attaining a 480-occurrences-per-million-words level in the 1980s. Even after the inclusion of data in TV/movies after the 1930s, both showed a constant increase.

---

expressing place tokens and irrelevant tokens of *BE about to*. The frequencies of *gonna* were used directly since there should be no *gonna* tokens irrelevant to *BE gonna*.

Here, it is important to compare the data of the present analysis with those of the OED's quotation texts by Mair,<sup>51</sup> as seen in Figure 1. The transition concerning BE *going to* shows a similar trend to his data: A sudden increase of BE *going to* was seen around 1900 in both Figures 1 and 3. Meanwhile, the frequency of BE *gonna* was extremely high due to the inclusion of the speech texts of TV/movies, compared with the OED's written data in Figure 1: Whereas the frequency in Figure 1 shows no increase after 1950, Figure 3 exhibits a steady increase until the 1980s.

Next, as for BE *about to*, no major change can be observed. Its frequency was around the 80-occurrences-per-million-words level until the 1870s but decreased slightly between the 1880s and 1950s and then its frequency rebounded gradually to around 80-occurrences-per-million-words level in the 2010s. It should be noted that the data seen in Figure 3 correspond with the data from the 1801–1850 period to the 1951–present period in Figure 2, presented by Watanabe.<sup>52</sup> The large-scale corpus discloses that after the sudden increase in the 1801–1850 period, BE *about to* neither increased nor decreased in frequency.

---

<sup>51</sup> Mair, "Corpus Linguistics and Grammaticalization Theory". 129.

<sup>52</sup> Watanabe, "On the Development of the Immediate Future Use of *Be About To*", 66.



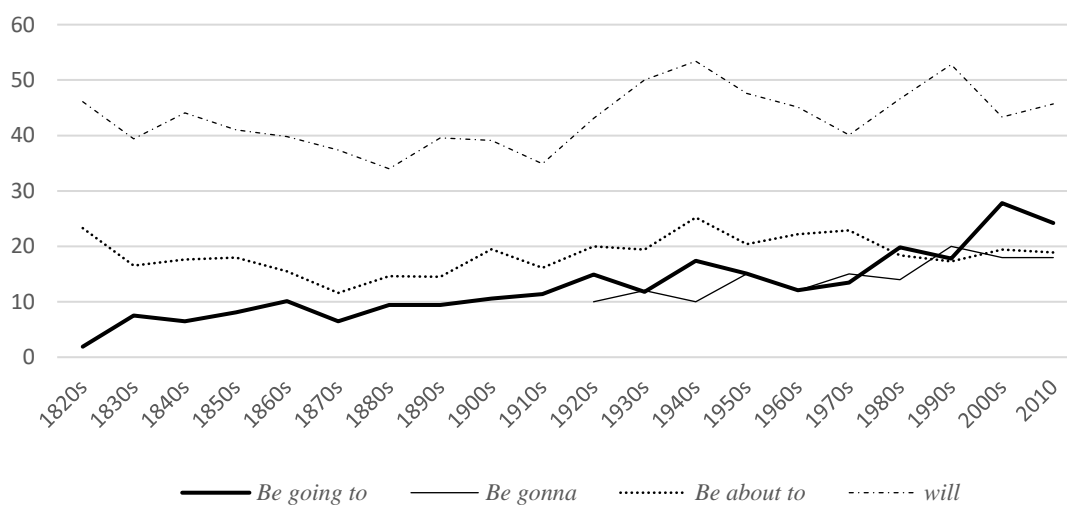
Thus, the present analysis has shown that in the last 200 years, BE *going to* and BE *gonna* showed a constant increase, whereas BE *about to* remained almost unchanged on a certain frequency level. In connection with features of inseparability, as seen in 2.2, the advanced level of grammaticalisation of BE *going to* seems to be sustained by its frequent use.

#### **4. Inanimate subject**

As referred to in the introduction and Section 2, the subject of the two semi-auxiliaries was inherently animate. Through grammaticalisation, the original sense of BE *going to* and BE *about to* diluted to the extent that inanimate subjects came to be freely used.

Along with the loss of analysability of the components and the bleaching of meaning, restrictions for the subject were longitudinally loosened. This section will explore how the subject changed over time. The present analysis adds the data of *will*, which developed the sense of future during the Middle English period, although its subject was inherently animate owing to its original verb's sense of 'to intend' and 'to wish'.

First, the ratio of animate subjects to the total number of subjects will be compared.<sup>53</sup> The following graph shows the transition of the ratio of inanimate subjects to the total use (also note Appendix 2):<sup>54</sup>



**Figure 4.** Ratio of inanimate subjects. (%)

<sup>53</sup> 200 tokens from each 10-year period were randomly retrieved and manually examined.

<sup>54</sup> The data of the 1910s, when the first token appeared, are not shown in Figure 4, because there were only seven tokens (two tokens were descriptions in a dictionary). The ratio of inanimate subjects was 14.3% (1 out of 7).

Figure 4 demonstrates several interesting points. First, BE *going to* showed a gradual increase from 1.9% in the 1820s to 24.2% in the 2010s.<sup>55</sup> This suggests that selective restrictions of the subject were loosened little by little as the ratio of inanimate subjects kept rising to the present, despite a sudden increase in total frequency around 1900, as shown in Figures 1 and 3. The ratio of inanimate subjects of BE *going to* is lower than that of BE *about to* until the 1970s, even with its advanced level of grammaticalisation and abundant use. Judging from there being almost 100 years in the sudden increase between BE *going to* and BE *about to*, which is considered a sign of the establishment as a full-fledged future marker according to Watanabe,<sup>56</sup> the low ratio of BE *going to* in the early 1800s may reflect a late start of grammaticalisation.<sup>57</sup> Or, the sense of *go* may call for animate subjects more strongly than that of *about*. In addition, it is noteworthy that *will*, another future marker, shows a much higher ratio than BE *going to* and BE *about to*.

---

<sup>55</sup> In the study by Budts and Petré, the ratio of inanimate subjects also showed a gradual increase: 4% between 1781 and 1815, 5% between 1816 and 1850, 8% between 1851 and 1885 and 10% in the 1886–1920 period.

<sup>56</sup> Watanabe, “On the Development of the Immediate Future Use of *Be About To*”. His argument is based on Hopper and Traugott (*Grammaticalization*, 106, 126-130).

<sup>57</sup> The previous version of the COHA contained data from the 1810s to 2000s. In its 1810s’ data, the ratio of inanimate subjects of BE *about to* was 19.4%. Conversely, that of BE *going to* was 5.1%.

Second, the informal shortened form BE *gonna* also shows a gradual augmentation at almost the same rate as BE *going to*. The ratio of inanimate subjects was 12.0% in the 1920s and rose to 20.0% in the 2010s. It is on a similar level to that of BE *going to* until the 1990s, whilst in the last two decades, the differences widened slightly. In sum, both forms gradually changed in tandem in terms of the ratio of inanimate subjects. It is reasonable to conjecture that due to the constant increase in frequency, as in Figure 3, and the gradual increase in the use of inanimate subjects, as in Figure 4, BE *going to* and BE *gonna* are undergoing further functional and semantic development at present.

Third, BE *about to* only slightly changed, showing a ratio of approximately 20% throughout the data. In the early 1820s, the ratio (23.3%) was high, but decreased to below the 20% level until the 1920s (as low as 11.6% in the 1870s), and rebounded to approximately the 20% level (18.9%) afterwards. As a whole, however, it is difficult to observe an inclination for the ratio to increase. Considering the unchanged frequency and separability, its grammaticalization seems to be hardly progressing.

## 5. Grammatical subjects

### 5.1 Diachronic development of grammatical subjects

This section will examine what kind of inanimate subjects became available for the two semi-auxiliaries. Inanimate subjects range from physical objects, such as *ship*, *building*, etc., through abstract nouns, such as *justice*, *inequality*, *anger*, etc., to constructions formed by grammatical components, such as gerund-*ing*, *to*-infinitive and *what*, etc.

This study will henceforth call such a subject consisting of grammatical components a grammatical subject. Grammatical subjects—patients in passive construction, existential *there*, *what*, *what*-clause, gerund-*ing*, *to*-infinitive and dummy *it* in the ‘*it* BE Adj *to*-infinitive’ construction—will be examined.<sup>58</sup> Strictly speaking, the patient in passive construction and *there*, however, cannot be categorised under grammatical subjects, but they will be included as grammatical subjects for the sake of convenience. Indeed, it is significant that a subject constituted by grammatical components is more conceptual than physical objects and abstract nouns. The use of a grammatical subject indicates a simple future with a sense of purpose and intention completely bleached.

---

<sup>58</sup> In the present analysis, the *what*-clause includes the *whatever*-clause as well.

Before proceeding with the examination, it is interesting to ascertain how *will* had already developed in terms of the choice of grammatical subjects in the early 1800s.

As expected, it took all grammatical subjects in the early 1820s in the COHA:

- (10) *will*
- a. [patient in passive construction]  
I think, on examination, their constituent parts will be found to have undergone the action of fire, (1820, magazine, COHA)
  - b. [*there* construction]  
There will be no permanent peace or prosperity in the South,... (1820, magazine, COHA)
  - c. [*what*]  
It remains, therefore, to inquire what will be the effect of this event, should it happen, ... (1822, magazine, COHA)
  - d. [*what*-clause]  
...what I am thoroughly convinced will do you good, ... (1827, fiction, COHA)
  - e. [*-ing*]  
Burning this prisoner will only make it worse. (1823, fiction, COHA)
  - f. [*to*-infinitive]  
To become truly her son, will depend upon your own choice,... (1827, fiction, COHA)
  - g. [*it* BE Adj *to*-infinitive]  
..., it will be useful to say a few words on a subject,... (1824, non-fiction, COHA)
  - h. [*whether*-clause]  
Whether this was its real character, or whether it was a mere pretence, will soon become manifest. (1828, magazine, COHA)

Although the word count of the data of the 1820s in the corpus is small,<sup>59</sup> the patient in passive construction (10a), *there* (10b), *what* (10c), *what*-clause (10d), *-ing* (10e), *to*-infinitive (10f), *it* in the ‘*it* BE Adj *to*-infinitive’ (10g) and *whether*-clause (10h) subjects were all retrievable. That the fully grammaticalised auxiliary took all the grammatical subjects in the 1820s can present the hypothesis that the availability of grammatical subjects is an indicator of the degree of grammaticalisation, in that no selective restrictions of subjects mean the dilution of the original sense. Here, the earliest tokens of the grammatical subjects in three items are as follows:

- (11) BE *going to*
- a. [patient in passive construction]<sup>60</sup>  
 ... if she was going to be put into the ground? (1823, fiction, COHA)
  - b. [*there* construction]  
 I almost dread to think of it. I’m afraid there is going to be a worse scrabble next summer to see who shall go (1833, fiction, COHA)
  - c. [*what*]  
 ...there’s something terrible hanging about my spirits, but what is going to happen I can not tell. (1830, fiction, COHA)
  - d. [*what*-clause]  
 ...and whatever you do in this matter is going to be the right thing. (1900, fiction, COHA)

---

<sup>59</sup> The word count of the 1820s and the 1830s is 6,981,389 and 13,711,287, respectively, whereas that of the 2010s is 35,452,806.

<sup>60</sup> The previous version of COHA which contained the data of the 1810s had a passive subject in the 1810s, but not the other grammatical subjects.

e. [-ing]  
Remembering this life is going to help us amazingly,...  
(1869, fiction, COHA)

f. [to-infinitive]  
To do away with this protectionism at home is going to mean changing  
drastically both the entire system of industrial hiring and firing,...  
(1963, magazine, COHA)

g. [it BE Adj to-infinitive]  
I thought it was going to be easy enough to straighten out that little difficulty,  
but it looks different now. (1875, fiction, COHA)

h. [whether]  
None

(12) BE gonna

a. [patient in passive construction]  
Nothing's gonna be done without orders. (1931, TV/movies, COHA)

b. [there construction]  
...I reckon there's gonna be trouble in town to-night. (1922, fiction, COHA)

c. [what]  
What's gonna happen to you, Rose? (1928, fiction, COHA)

d. [what-clause]  
Except what I said was gonna happen. (1967, fiction, COHA)

e. [to-infinitive]  
None

f. [-ing]  
But liking it is gonna be different. (1938, TV/movies, COHA)

g. [it BE Adj to-infinitive]  
It's gonna be hard to face that gang tomorrow if I'm back at that desk.  
(1940, TV/movies, COHA)

h. [whether-clause]  
None



- (13) BE *about to*<sup>61</sup>
- a. [patient in passive construction]
    - ...and until the house was about to be burnt. (1821, non-fiction, COHA)
  - b. [*there* construction]
    - ...there was about to be a shower of good things among some folks... (1834, fiction, COHA)
  - c. [*what*]
    - ...we had little apprehension in our hearts, of what was about to happen. (1823, fiction, COHA)
  - d. [*what*-clause]
    - ...the Apaches knew exactly where their pale-face friends were at that particular moment, which was the precise thing Murray wanted them to feel sure of, considering what he knew was about to be found out. (1882, fiction, COHA)
  - e. [*-ing*]
    - None
  - f. [*to*-infinitive]
    - To protect their independence Was about to come up. (2017, TV/movies, COHA)
  - g. [*it* BE Adj *to*-infinitive]
    - None
  - h. [*whether*-clause]
    - None

To effectively display the chronological order, the graphical presentations of the year of each of the earliest tokens for the three items will be shown in the following three figures:

---

<sup>61</sup> In the previous version of the COHA, the passive subject had an earlier token in 1815, and the *what* subject had in 1811.

	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
patient in passive	█																			
there	█	█																		
what	█	█																		
what-clause																				
-ing																				
to-infinitive																				
it Be Adj to																				
whether																				

**Figure 5.** The year of the first token of BE *going to* in COHA.

	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
patient in passive	█																			
there	█	█																		
what	█																			
what-clause																				
-ing																				
to-infinitive																				
it Be Adj to																				
whether																				

**Figure 6.** The year of the first token of BE *gonna* in COHA.

	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
patient in passive																				
there																				
what																				
what-clause																				
-ing																				
to-infinitive																				
it Be Adj to																				
whether																				

**Figure 7.** The year of the first token of BE *about to* in COHA.

The above data provide five important points. First, the years of the earliest tokens of the patient, *there* and *what* subjects were seen in the early 1800s for BE *going to* and BE *about to*. The cases of the patient and *there* subjects are predictable, as previous studies refer to (cf. Watanabe,<sup>62</sup> Garrett,<sup>63</sup> Traugott and Trousdale<sup>64</sup>). This study would like to

<sup>62</sup> Watanabe, “On the Development of the Immediate Future Use of *Be About To*”.

<sup>63</sup> Garrett, “The Historical Syntax Problem”.

<sup>64</sup> Traugott and Trousdale, *Constructionalization and Constructional Change*, 94-148.

add the *what* subject to the list of early signs of grammaticalisation. Indeed, it seems that *what* BE *going/about to happen*, as in (11c), (12c) and (13c), is based on the sequence of *what* + *happen*, which several English learners' dictionaries cite as a common collocation in contemporary English. In the following subsection of 5.2, this study will propose that the *what* subject emerged immediately after the *there* subject in the English language.

Second, after the *what* subject appeared in the early 1800s, the *what*-clause subject appeared: for BE *going to* in 1900 and for BE *about to* in 1882. As seen in the examples of (10d), (11d), (12d) and (13d), the *what*-clause subject contains more content in terms of tense, agent and activity of the agent, compared with the *what* subject. This may suggest that the content-rich subject is more difficult to co-occur with the future-oriented semi-auxiliaries. This issue will be further addressed below.

Third, the years of the first tokens of the gerund-*ing* subject and *it* in the '*it* BE Adj *to*-infinitive' construction uncover an interesting aspect. (This study would like to make a point based on the presupposition that the *it* in the '*it* BE Adj *to*-infinitive' construction refers to the *to*-infinitive.) BE *going to* started taking them almost at the same time in the 1860s and the 1870s, later than the patient, *there* and *what* subjects but earlier than the *what*-clause subject. What these two subjects denote is more content-

rich than the patient, *there* and *what* subjects, owing to the presence of activity and agent (optional) but less content-rich than the *what*-clause subject, owing to lack of tense. Thus, as far as BE *going to* is concerned, the content richness seems to determine the order in which the grammatical subject emerges earlier. On the other hand, BE *about to* has not co-occurred with the *-ing* subject and *it* in the ‘*it* BE Adj *to*-infinitive’.

Fourth, some grammatical subjects appeared late. The *to*-infinitive appeared late in the 1960s for BE *going to* and in the 2010s for BE *about to*; in addition, a token of the *whether*-clause subject was not observed for both semi-auxiliaries. According to four native English speakers (two American English speakers and two British English speakers), the *to*-infinitive subject is not fully acceptable as a subject for BE *going to* and BE *about to*. Duffley<sup>65</sup> argues that since the *to*-infinitive subject tends to imply that the event expressed is unexpected and undesirable whilst that of the *-ing* subject expresses neutrality, the use of the former connotes the speaker’s subjective attitude, which seems to be incompatible with the non-subjective future sense that BE *going to* and BE *about to* express. Therefore, further analysis about the *to*-infinitive subject will not be conducted in this study. (As seen in (11f) and (13f), several tokens of the *to*-infinitive subject were

---

<sup>65</sup> Duffley, *The English Infinitive*, 126-32.

retrieved, but they are considered to be aberrant tokens.) Meanwhile, the *whether*-clause subject expresses more content-rich features than the *what*-clause subject because of its representation of the interrogative sense. According to the native speakers mentioned above, the *whether*-clause subject is acceptable when co-occurring with BE *going to*. It appears that the word count of the COHA is insufficient to contain such a token.

Fifth, BE *gonna* has taken over the characteristics of BE *going to*, whereas the earliest token of each grammatical subject appeared after the 1920s. The tendency of BE *going to* in which the patient, *there* and *what* subjects appeared earlier than the *-ing* subject and the ‘*it* BE Adj *that*-clause/*to*-infinitive’ construction can be observed in Figure 6 as well. As expected, the *what*-clause subject emerged later than the other grammatical subjects, in the same way as BE *going to*.

After considering the five points thus far, it is justifiable to conclude that BE *going to* took grammatical subjects with simple content earlier than content-rich ones in the process of the dilution of the sense of *go*. Furthermore, as referred to in Section 4, it is possible to infer that BE *going to* is undertaking further grammaticalisation due to a continuous increase in the ratio of inanimate subjects as well as frequency. As a result, the development of grammatical subjects, such as the *where*-clause, *how*-clause and *that*-clause subjects, can be witnessed in larger corpora. By contrast, in the case of BE

*about to*, it fails to show the same ‘content-based’ developmental path. Since the emergence of a *what*-clause subject in 1882, there have been no tokens for an *-ing* subject and an ‘*it* BE Adj *that*-clause/*to*-infinitive’ construction. The data in the present subsection also support our argument that grammaticalisation of BE *about to* seems to have been halted in present-day English, as discussed in Section 4. Further research is required to clarify the detailed development in current English.

Finally, as a caveat, it should be mentioned that there is a possibility that the chronological order of the earliest token of BE *going to* and BE *about to*, as seen in Figures 5 and 7, does not reflect the overall scenario, because there may be earlier tokens which were sifted out when the corpus was compiled, and thereby could lack in accuracy. The next subsection will demonstrate that the chronological order the present analysis has presented in 5.1 has yet to have significance.

## ***5.2 Data in the 2010s***

This subsection will consider the current data. The number of tokens for each grammatical subject appearing in the latest data of the 2010s in the COHA is shown in Table 4 (the occurrences in TV/movies subcorpus are in parentheses to show the

tendency of how often grammatical subjects appear in speech):<sup>66</sup>

**Table 4.** Raw occurrences of grammatical subjects in the 2010s' data of COHA.

	patient <sup>67</sup>	<i>there</i>	<i>what</i>	<i>what</i> - clause	<i>-ing</i>	<i>to</i> - infinitive	<i>it</i> Be Adj <i>to</i>	total tokens
BE <i>going to</i>	270 (50)	161 (26)	132 (21)	16	59 (12)	2 (1)	28 (8)	18,510 (4,283)
BE <i>gonna</i>	94 (83)	72 (64)	55 (45)	1 (1)	24 (22)	0	3 (3)	16,644 (14,266)
BE <i>about to</i>	46 (1)	2 (2)	30 (4)	1 (1)	0	1 (1)	0	2,856 (640)

(TV/movie subcorpus)

Table 4 displays a surprising tendency for BE *going to*. The order of the frequency in

Table 4 is almost the same as that of the year of the earliest token in Figure 5: 1. patient,

2. *there*, 3. *what*, 4. *-ing*, 5. *it* BE Adj *to*, and 6. *what*-clause.<sup>68</sup> Equally important, BE

*gonna* shows the same frequency order as BE *going to*: 1. patient, 2. *there*, 3. *what*, 4. -

---

<sup>66</sup> The data in Table 4 are retrieved by manually examining all tokens of *going to*, *gonna* and *about to*.

<sup>67</sup> Tokens expressing emotions such as *be surprised/thrilled* are excluded.

<sup>68</sup> To be precise, according to the first tokens in (11), the *what* subject of (11c) appeared three years earlier than the *there* subject of (11b), but Table 4 shows that there are more *there* subject tokens (161 occurrences) than *what* subject tokens (132 occurrences). The present investigation found a token in 1833 of (11b) as the earliest token of a *there* subject, whilst Budts' investigation ("The Grammaticalization of *Be Going To* in Late Modern English") into CLMET 3.0 culled a token in 1857 as the earliest token.

*ing*, 5. *it* BE Adj *to*, and 6. *what*-clause. From the viewpoint of historical development, even after considering the possible presence of earlier and aberrant tokens as mentioned in 5.1, the chronological order is considered to reflect the degree of semantic compatibility between the grammatical subject and the sense of BE *going to* in current English. A grammatical subject which appeared earlier is more compatible with the semi-auxiliary, and thereby is used more often than one which appeared later. Conversely, the more compatible, the more is used, and the earlier the subject appears in a corpus, a trend which continued until the 2010s.

Here, we are in a position to validate our hypothesis that the *what* subject also appeared in an early stage of grammaticalisation by analysing the relationship between frequency and the year when each grammatical subject first appeared. First of all, it is necessary to consider the cases of the passive and *there* subjects. As mentioned in 2.1.1, Traugott and Trousdale<sup>69</sup> maintain that a plausible first token of the passive subject appeared in 1628, followed by an early token of the *there* subject in 1725. The gap of almost 100 years seems to be reflected by the difference in raw occurrences, i.e., 270 tokens and 161 tokens respectively, as in Table 4. Interestingly enough, the difference

---

<sup>69</sup> Traugott and Trousdale, *Constructionalization and Constructional Changes*, 221.



in raw occurrences between the *there* and *what* subjects (161 and 132 tokens respectively) is smaller than that between the passive and *there* subjects, which empirically indicates that the *what* subject appeared slightly later than the *there* subject. Indeed, the following token in 1741 was retrieved from Eighteenth Century Collections Online (ECCO):<sup>70</sup>

(14) The Capucbin, who was no Stranger to Madame d’Eftival’s Vagaries, trembled, as did all the Company, at such a formidable Question, as if what was going to happen, had been foreseen.

(1741, Charles de Fieux, *The Fortunate Country Maid*)

Thus, the earliest token of the *what* subject in ECCO appeared only 16 years apart from the early token of the *there* subject in 1725, which certainly proves that the *what* subject also appeared in the early stage of grammaticalization of BE *going to*.

In the case of BE *about to*, the frequency order also corresponds with the chronological order: 1. patient, 2. *what*, 3. *there* and 4. *what*-clause. Its incompatibility

---

<sup>70</sup> Traugott and Trousdale also cite a token of the year 1741 as an early example of a *there* construction (2013: 222).

with the *-ing* subject and the *it* BE Adj *to* construction, which was revealed in the analysis of 5.1, is evinced in Table 4 as well. Also, it is worth noting that the raw occurrences of the *there* subject are very low (only two tokens) in Table 4, although it emerged as early as in 1834. Such a weak collocation in current English between BE *about to* and the *there* subject, compared with the *what* subject mirrors the use of the time in the early 1800s: The *there* subject emerged later than the *what* subject, as in Figure 7, unlike BE *going to*. In fact, we can reach a conclusion that the *what* subject is an indicator of grammaticalisation in the early stage, as in the case of BE *going to*. To summarize the analyses of the three items, the findings show the parallel features of synchronic and diachronic occurrences.

Regarding additional findings, grammatical subjects show a tendency to appear slightly more frequently in writing than in speech. The ratios of raw occurrences in TV/movies to the total occurrences (16.2% or 108/668 for BE *going to*, 87.6% or 218/249 for BE *gonna*, 11.3% or 9/80 for BE *about to*) are lower than those of TV/movies in word count (as Table 3 shows, 23.1% for BE *going to*, 90.8%, for BE *gonna* and 22.2% for BE *about to*).

## 6. Conclusion

This study has clarified the development of BE *going to* and BE *about to*, both of which originally called for animate subjects, by examining the data between the 1820s and 2010s in the COHA. The development of BE *gonna* after the 1910s is also investigated. BE *going to* showed a sudden increase in frequency around 1900, maintaining a steady increase afterwards until the present, while exhibiting a gradual increase in the ratio of inanimate subjects up to the 2010s. BE *gonna* showed similar tendencies. On the other hand, BE *about to* made no major change in frequency and the ratio to inanimate subjects. Indeed, it has been shown that the level of inseparability and the expansion of the grammatical subjects of the three items seem to reflect their differences in the degree of grammaticalisation. These facts indicate that grammaticalization of BE *going to* is still proceeding at present, whereas grammaticalisation of BE *about to* is halted in contemporary English.

Furthermore, it is also found that the chronological order of the year of the earliest token of each grammatical subject corresponds with its frequency order in current English. BE *going to* started to take the patient, *there* and *what* subjects first, then the *-ing* subject and *it* in ‘*it* BE Adj *to*-infinitive’, and finally the *what*-clause subject before showing a sudden increase in frequency. The frequency of each

grammatical subject presents the same order in current English. As far as BE *going to* is concerned, content-rich grammatical subjects are less likely to co-occur than those with simple content, diachronically as well as synchronically. BE *about to*, on the other hand, took the patient, *what* and *there* subjects, then the *what*-clause subject, but stopped short of the *-ing* subject and an *it* in ‘*it* BE Adj *to*-infinitive’. Thus, the semantic compatibility between the semi-auxiliary and the grammatical subject seems to determine the order of when the earliest token of each grammatical subject appears and how frequently it is used in current English.

**Funding** This study is financially supported by Japan Society for the Promotion of Science under Grant Number 22K00618.

**Disclosure statement.** No potential conflict of interest was reported by the author.

## References

- Budts, Sara. “The Grammaticalization of *Be Going To* in Late Modern English: A Corpus-based Analysis.” <[https://www.academia.edu/21024328/The\\_grammaticalization\\_of\\_be\\_going\\_to\\_in\\_Late\\_Modern\\_English\\_a\\_corpus\\_based\\_analysis](https://www.academia.edu/21024328/The_grammaticalization_of_be_going_to_in_Late_Modern_English_a_corpus_based_analysis)> 2013. Accessed on 31 January, 2021.
- Budts, Sara and Peter Petré. “Reading the Intentions of *Be Going To*: On the Subjectification of Future Markers.” *Folia Linguistica Historica* 37 (2016): 1-32.

- Bybee, Joan. "Cognitive Processes in Grammaticalization." In *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure, Volume II*, edited by Michael Tomasello. Mahwah: Erlbaum, 2003.
- \_\_\_\_\_. "Mechanisms of Change in Grammaticization: The Role of Frequency." In *The Handbook of Historical Linguistics*, edited by Brian D. Joseph, and Richard D. Janda. Blackwell: Oxford, 2003.
- \_\_\_\_\_. "From Usage to Grammar: The Mind's Response to Repetition." *Language* 82, no. 4 (2006): 711-33.
- \_\_\_\_\_. *Language Change*. Oxford: Oxford English Press, 2015.
- Bybee, Joan, and William Pagliuca. "The Evolution of Future Meaning." In *Papers from the 7th International Conference on Historical Linguistics*, edited by Anna Giacalone Ramat, Onofrio Carruba, and Giuliano Bernini. Amsterdam/Philadelphia: John Benjamins, 1987.
- Collins, Peter. *Modals and Quasi-Modals in English*. Amsterdam/New York: Brill Rodop, 2009.
- Corpus of Historical American English. <<https://www.english-corpora.org/coha/>> Accessed from 20 August, 2019 to 30 September, 2022.
- Duffley, Patrick J. *The English Infinitive*. London: Routledge, 1992.
- \_\_\_\_\_. "The Gerund and the *To*-infinitive as Subjects." *Journal of English Linguistics* 31 (2003): 324-52.
- Eighteenth Century Collections Online. <<https://www.gale.com/jp/c/eighteenth-century-collections-online>> Accessed on 25 December, 2022 to 7, January, 2023.
- Fischer, Olga. *Morphosyntactic Change: Functional and Formal Perspectives*. Oxford: Oxford University Press, 2007.
- \_\_\_\_\_. "An Analogical Approach to Grammaticalization." In *Grammaticalization: Current Views and Issues*, edited by Ekaterini Stathi, Elke Gehweiler, and Ekkehard König. Amsterdam: John Benjamins, 2010.
- Garrett, Andrew. "The Historical Syntax Problem: Reanalysis and Directionality." In *Grammatical Change: Origins, Nature, Outcomes*, edited by Dianne Jonas, John Whitman, and Andrew Garrett. Oxford: Oxford University Press, 2012.
- Heggie, Lorie, A. *The Syntax of Copular Structures*. Ph.D. Dissertation. University of Southern California, 1988.
- Hilpert, Martin. *Germanic Future Constructions: A Usage-based Approach to Language Change*. Amsterdam & Philadelphia: John Benjamins, 2008.
- Hopper, Paul, and Elizabeth C. Traugott. *Grammaticalization, Second edition*. New York: Cambridge University Press, 2003.
- Kirsner, Robert S., and Sandra A. Thompson. 1976. "The Role of Pragmatic Inference in Semantics: A Study of Sensory Verb Complements in English." *Glossa* 10 no. 2 (1976): 200-40.

- Mair, Christian. "Corpus Linguistics and Grammaticalization Theory: Statistics, Frequencies, and Beyond." In *Corpus Approaches to Grammaticalization in English*, edited by Hans Lindquist, and Christian Mair. Amsterdam/Philadelphia: John Benjamins, 2004.
- Oxford English Dictionary*. <<https://www.oed.com/>> Accessed in December, 2022.
- Palmer, Frank R. *Modality and English Modals*. Harlow: Longman, 1990.
- Pertejo, Paloma Nuñez. 1999. "Be Going To + Infinitive: Origin and Development. Some Relevant Cases from the Helsinki Corpus." *Studia Neophilologica* 71 no. 2 (1999): 135-42.
- Poole, Joshua. *The English Accidence*. Early English Books Online, 1646.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. *A Comprehensive Grammar of English Language*. Harlow: Longman, 1985.
- Swan, Michael. *Practical English Usage, Fourth Edition*. Oxford: Oxford University Press, 2017.
- Traugott, Elizabeth, C. "On the Persistence of Ambiguous Linguistic Contexts over Time: Implications for Corpus Research on Micro-changes." In *Corpus Linguistics and Variation in English: Theory and Description*, edited by Mugnus Huber, and Joybrato Mukherjee. Amsterdam: Rodopi, 2012.
- Traugott, Elizabeth, C., and Richard Dasher. *Regularity in Semantic Change*. Oxford: Oxford University Press, 2002.
- Traugott, Elizabeth, C., and Graeme Trousdale. *Constructionalization and Constructional Changes*. Cambridge: Cambridge University Press, 2013.
- Wada, Naoaki. "Be Going To and Be About To: Just Because Doc Brown was Going to Take Us Back to the Future Does Not Mean that He was About to Do So." *English Linguistics* 17 no. 2 (2000): 386-416.
- \_\_\_\_\_. *The Grammar of Future Expressions in English*. Kaitakusya: Tokyo, 2019.
- Watanabe, Takuto. "On the Development of the Immediate Future Use of Be About To in the History of English with Special Reference to Late Modern English." *English Linguistics* 28 no. 1 (2011): 56-90.
- Yaguchi, Michiko. "The Process of Grammaticalization of Be Bound To as a Semi-auxiliary in American English." *English Corpus Studies* 27 (2020): 53-72.
- Westney, Paul. *Modals and Periphrastics in English*. Tübingen: Niemeyer, 1995.

**Appendix 1. Change in frequency in COHA. (occurrences/million words)**

	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010s
<i>Be going to</i>	33.1	50.4	48.0	86.0	102.7	129.0	146.4	190.2	230.3	351.2	288.7	481.7	461.9	501.4	514.9	500.1	476.6	516.0	570.6	543.5
<i>Be gonna</i>										0.4	4.6	138.7	122.4	193.2	239.1	437.8	480.0	367.0	391.2	441.3
<i>Be about to</i>	80.6	81.1	80.1	69.0	66.0	80.4	69.2	63.2	68.1	65.1	63.0	56.4	54.8	52.6	68.6	66.7	68.4	69.8	82.4	79.7

**Appendix 2. Ratio of inanimate subjects. (%)**

	1820s	1830s	1840s	1850s	1860s	1870s	1880s	1890s	1900s	1910s	1920s	1930s	1940s	1950s	1960s	1970s	1980s	1990s	2000s	2010
<i>Be going</i>	1.9	7.5	6.5	8.1	10.1	6.5	9.4	9.4	10.6	11.4	14.9	11.8	17.4	15.1	12.1	13.45	19.8	17.8	27.8	24.2
<i>Be gonna</i>											10	12	10	15	12	15	14	20	18	18.0
<i>Be about</i>	23.3	16.5	17.6	18	15.5	11.6	14.6	14.5	19.5	16.1	20	19.4	25.2	20.4	22.2	22.9	18.4	17.3	19.4	18.9
<i>will</i>	46.1	39.4	44.1	41	39.8	37.4	34	39.6	39.1	34.9	43.1	50	53.4	47.6	45.1	40.1	46.6	52.8	43.3	45.7

**Table 1.** Frequency of *BE going to*, *BE gonna* and *BE about to* in the 2010s' COHA data.

	fiction	magazine	news	non-fiction books	TV/movies
<i>BE going to</i>	570.4	360.5	584.7	64.9	846.2
<i>BE gonna</i>	28.0	9.3	2.4	0.9	400.7
<i>BE about to</i>	112.6	40.9	40.7	17.5	127

(occurrences per one million words)

**Table 2.** Breakdown of word counts in the 2010s' data in COHA.

fiction	magazines	news	non-fiction books	TV/movies
43%	21%	13%	9%	14%

**Table 3.** Distribution of tokens of BE *going to*, BE *gonna* and BE *about to* in the 2010s' data in COHA.

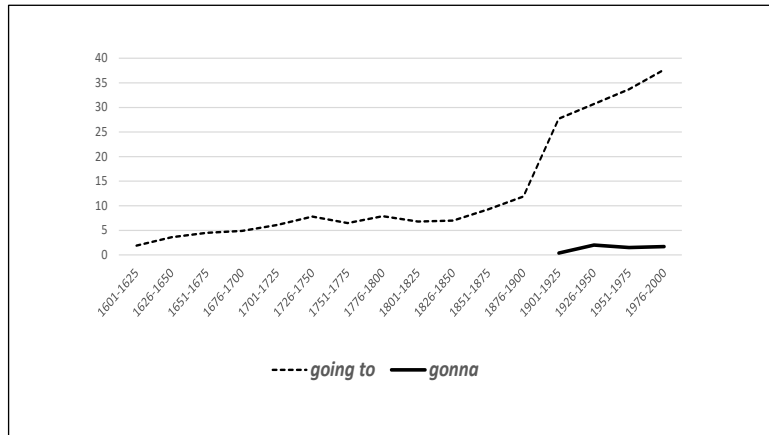
	fiction	magazines	news	non-fiction books	TV/movies
BE <i>going to</i>	46.7%	14.7%	14.4%	1.1%	23.1%
BE <i>gonna</i>	6.4%	2.1%	0.6%	0.2%	90.8%
BE <i>about to</i>	58.9%	10.7%	6.4%	1.9%	22.2%

**Table 4.** Raw occurrences of grammatical subjects in the 2010s' data of COHA.

	patient	<i>there</i>	<i>what</i>	<i>what-</i> <i>clause</i>	<i>-ing</i>	<i>to-</i> <i>infinitive</i>	<i>it Be Adj</i> <i>to</i>	total tokens
BE <i>going to</i>	270 (50)	161 (26)	132 (21)	16	59 (12)	2 (1)	28 (8)	18,510 (4,283)
BE <i>gonna</i>	94 (83)	72 (64)	55 (45)	1 (1)	24 (22)	0	3 (3)	16,644 (14,266)
BE <i>about to</i>	46 (1)	2 (2)	30 (4)	1 (1)	0	1 (1)	0	2,856 (640)

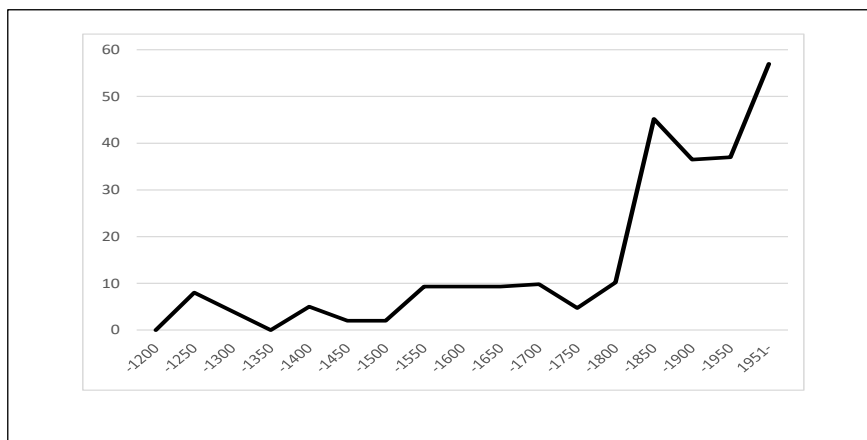


(TV/movie subcorpus)



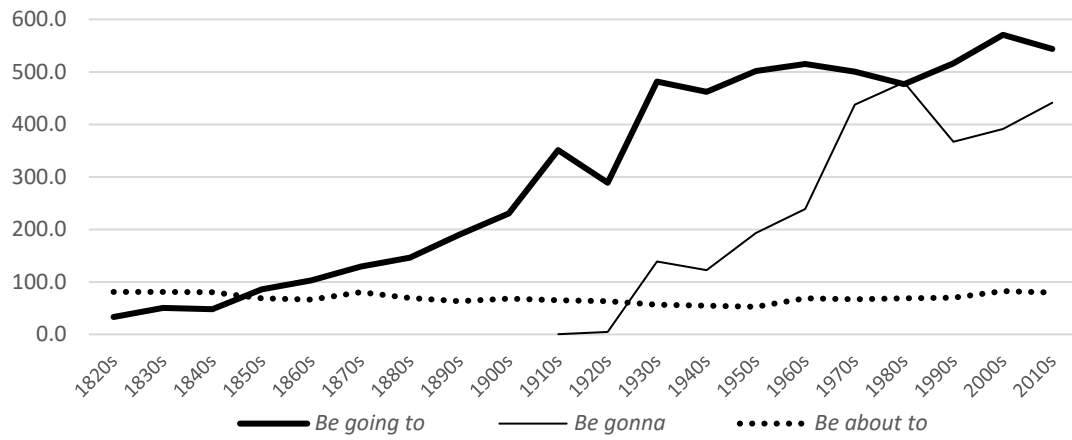
**Figure 1.** *Going to* and *gonna* 1600-2000. (occurrences per 10,000 quotation texts)

(excerpted from Mair)

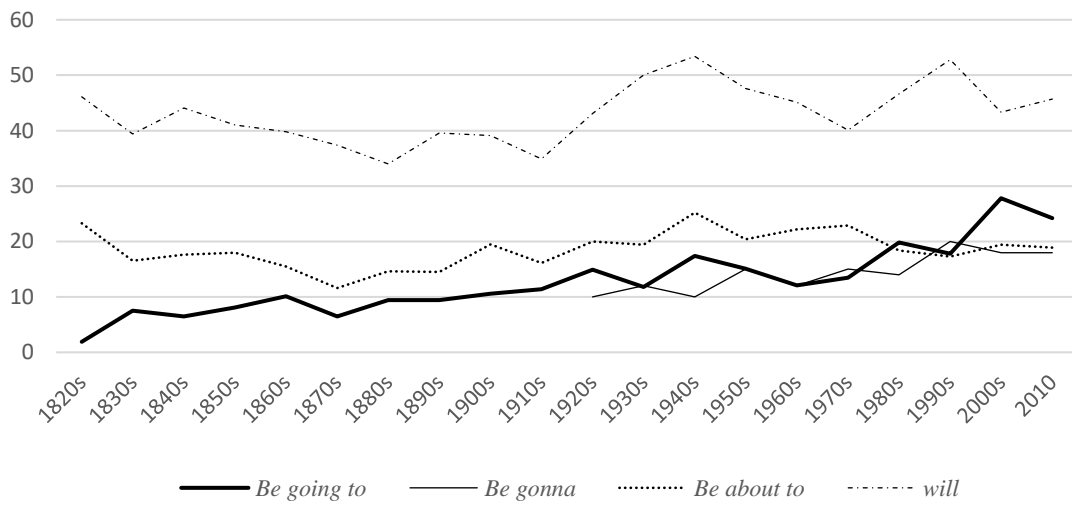


**Figure 2.** Frequency of BE *about to* in OED. (occurrences per 100,000 quotation texts)

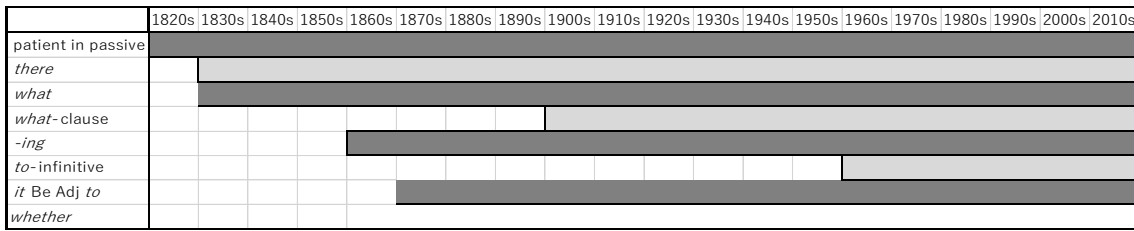
(excerpted from Watanabe)



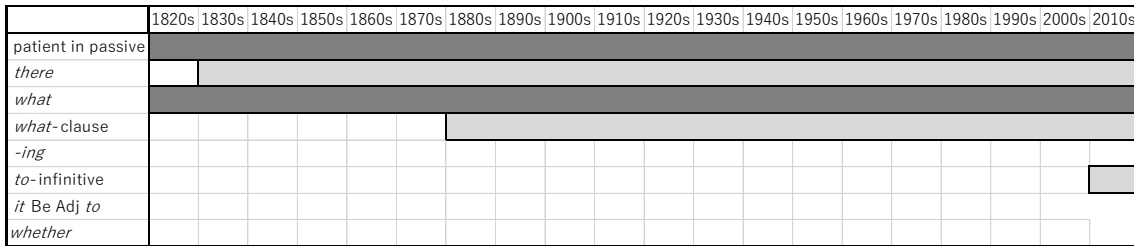
**Figure 3.** Change in frequency in COHA.  
(occurrences per one million words)



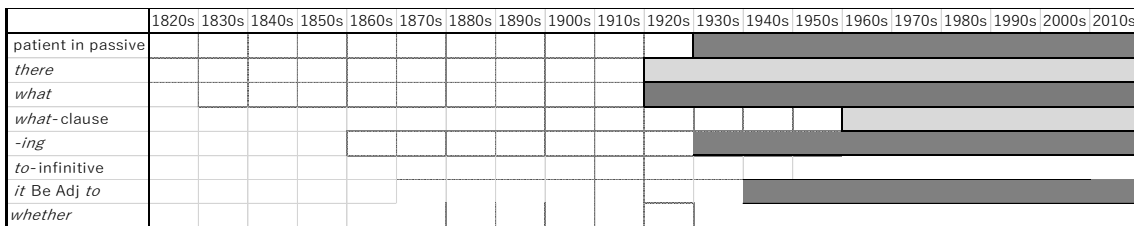
**Figure 4.** Ratio of inanimate subjects. (%)



**Figure 5.** The year of the first token of BE *going to* in COHA.



**Figure 6.** The year of the first token of BE *gonna* in COHA.



**Figure 7.** The year of the first token of BE *about to* in COHA.