Split infinitives in English: A corpus-based investigation

Youngjun Jang* · Sunjoo Choi**
(Chung-Ang University)

Jang, Youngjun and Sunjoo Choi. 2014. Split infinitives in English: A corpus-based investigation. Linguistic Research 31(1), 53-68. The focus of this paper is to investigate the corpora of split infinitives. In English, some adverbs occur between the infinitive marker to and the verb, which has been called split infinitives. To split or not to split infinitive has been an issue in English grammar. Modern English, however, witnesses a burst-out of the usage of split infinitives, as has been investigated by numerous works in the literature. In this paper, we have investigated the phenomenon using corpus data such as the Corpus of Contemporary American English (COCA), the Corpus of Historical American English (COHA) and Time Magazine Corpus and we reached a conclusion that split infinitives are being used more often than before. We also tried to analyse various types of split infinitives. It appears that only a limited number of adverbs can split the infinitives and only a number of verbs can be split by these adverbs. There is a big difference between the distribution of adverbs in a finite clause and that of an infinitival clause. (Chung-Ang University)

Keywords: split infinitives, infinitival marker, adverb, COCA, COHA, corpus, intensifiers, clausal structure

1. Introduction

To split or to ‘not’ split? That has been one of the controversial issues in the grammar of English (Crystal, 2003; Mikulova, 2011; Mitrasc, 2009, among others). Split infinitives are the constructions in which an adverb, an adverbal phrase, or a

* Corresponding author ** First author
1 Portions of this paper were presented at Annual Conference of the English Language and Literature Society of Korea, November 7, 2013, at Sookmyung Women's University, at the Fall Joint Conference of the Joongwon Linguistic Society of Korea and the Korean Generative Grammar Circle held at Semyung University, Korea, November 16, 2013, and at the Second Asia Pacific Corpus Linguistics Conference (APCLC 2014) held at Hong Kong Polytechnic University, Hong Kong, March 7-8, 2014. We thank the audience for their questions and comments. All errors are, however, our own. This research was supported by the Chung-Ang University Research Assistant Scholarship in 2013.
negative expression is inserted between the infinitival marker *to* and the verb, as shown in (1) below.

(1) He seemed *to really want* to find a way to serve. (TIME MAGAZINE CORPUS December 22, 1980)

(2) the states he’s won, the red states, is they’re not likely *to all of a sudden turn* blue in November... (Mikulova, 2011:31)

(3) They will give young people yet another reason *to not go* to church. (Mikulove, 2011:28)

In (1), the adverb really is inserted between *to* and the infinitive verb *want*. In (2), the adverbial phrase *all of a sudden* is inserted between *to* and the infinitival verb *turn*. In (3), the negative head not is inserted between *to* and the infinitival verb *go*. Most recently, Pinker (2012) also mentions that splitting infinitives has been one of the important issues in prescriptive English grammar, mentioning examples like *to boldly go where no man has gone before*, which otherwise would be *to go boldly where no man has gone before*.

The earliest record of split infinitives in written English text traces back to the 13th century (Curme, 1931). But the usage suddenly fell down during the 18th century, when the English grammarians including Robert Lowth discouraged the use of split infinitives. The main rationale for this discouragement was that splitting infinitives are not in accordance with English grammar. However, use of split infinitives began to upsurge during the 19th century and nowadays we see abundant use of split infinitives in media and academic writings. It seems still not quite settled whether infinitives should or should not be split in terms of prescriptive grammar, as David Crystal (2003: 193) notes.

In this paper, we investigate the origin and development of the split infinitives in English, using corpus data. In so doing, we could make some observations. First, we found out that the use of split infinitives is drastically increased recently, especially after the year 2000. Second, we found out that only a limited number of adverbs, but not any adverbs, are frequently used in split infinitival constructions. Third, we found out that the most frequently-used adverbs are intensifiers. Lastly, we found out that there is a big difference between the distribution of adverbs in a finite clause and that of an infinitival clause.
2. A historical review

A split infinitive first appeared in the 13th century. The split infinitives of this period, however, are different from those of contemporary American English in that a pronoun, and not an adverb, was used to split the infinitives. The split infinitives in which an adverb is inserted began to be found only from the 14th century. Many literary works of the 14th century such as Chaucer's contain split infinitives. Since the split infinitives appear as early as the 13th century and in many texts of the 14th century, it may not be correct to claim that split infinitives are un-English, as most of the 18th century prescriptive grammarians claimed (for discussion, see Crystal 1985:46). Some of the relevant examples are provided in what follows.

(4) the 14th century
   a. He louied þe lasse auþer to lenge lye or to longe sitte
      'He did not like to either lie or sit long.'
      (Sir Gawayne and the Greene Knight, ll.87-88, quoted from Curme 1931:460)
   b. It is good to not ete flesich and to not drinke wyn
      'It is good to not eat flesh and to not drink wine.'
      (Wycliff, Romans, XIV, 21, Purvey's ed., A.D.1388, quoted from Curme 1931:460)

(5) the 15th century
   Y schall swere to not discouere them
   'I shall pledge myself to not inform on them.'
   (The Folewer to the Donet, E.E.T.S., No.164,1454, quoted from Curme 1931:460)

In (4a), the adverbs lenge and longe each appear between to and the verbs lye and sitte. In (4b), the negative head not appears between to and the verb ete 'eat.' In (5), the negative head not appears between to and the verb discouere 'inform.'

The split infinitives are found in the writing of the 16th and 17th century writers such as Thomas Cromwell, William Tyndale, Samuel Pepys, Sir Philip Sydney, John

---

2 Unfortunately, we have been unable to find relevant examples from the corpora nor from previous works.
Donne, Daniel Defoe, and Dr. Johnson, as Curme (1931) notes. However, the use of split infinitives suddenly died out during the 16th century and non-split infinitives such as (6) below are often found. The exact reason why its use had suddenly dropped is not clear, but many scholars including Curme (1931), Myers (2002), and Visser (1984) note that it was so.

(6) the 16th century

I am ready obediently to conforme myself to his graces' commandments.

(Roper, c. 1557, quoted from Fowler & Burchfield, 1998:738)

If split infinitives had still been used in this period, examples like (6) might have been something like ... to obediently conforme... One can conjecture that the use of split infinitives was discouraged by the traditional English grammarians of these days. As Myers (2002:59) notes, it seems that traditional prescriptive grammarians like Lowth (1762) reasoned “that since it was grammatically impossible to split an infinitive in Latin, the language of learning and prestige in his day, then English shouldn't allow the split infinitive either.” Of course, it might seem rather illogical to quote Bishop Lowth (1762) since he was a 18th century grammarian, as an anonymous reviewer notes. We simply want to point out that Latin might have exerted quite a strong influence on the decline of the English split infinitives. Indeed, Latin does not allow split infinitives only because the Latin infinitives are single words. For example, the Latin infinitive verb facere 'to make' is one word and there is simply no way of splitting this infinitive.

Split infinitives began to reappear in the 18th century and during this period adverbs, instead of pronouns, are inserted between the infinitival marker to and the verb, just like contemporary split infinitives (Visser, 1984:997). They became more common in the nineteenth and twentieth centuries. Curme (1931) found that over sixty renowned authors of literature, science, and political discourse published in those centuries used various forms of the split infinitive. Even though the use of split infinitives has been drastically increased in contemporary English, the issue of whether split infinitives are grammatically acceptable or not does not still seem to be resolved. For example, Eastman (1994:146) argues that “splitting infinitives is not acceptable and we should avoid split infinitives especially in writing.”
3. Pros and cons regarding split infinitives

Those who had traditional education considered splitting infinitives as inappropriate and ungrammatical, since they assumed that to and the infinitive verb are a fixed linguistic unit that cannot be split. Furthermore, traditional grammarians of the 18th century assumed that a language should be controlled and kept in pure form (whatever it may mean).

An extreme case of prohibiting split infinitives can be found in a letter to the editor of the *New England Magazine*, written in 1834 (quoted from Bailey, 2006). The writer of the letter mentions that:

*The practice of separating the prefix of the infinitive mood from the verb, by the intervention of an adverb, is not unfrequent among uneducated persons as “to fully understand it,” instead of “to understand it fully” or “fully to understand it.” This fault is not often found in print, except in newspapers where the editors have not had the advantage of a good education.... the particle to, which comes before the verb in the infinitive mode, must not be separated from it by the intervention of an adverb or any other word or phrase; but the adverb should immediately precede the particle, or immediately follow the verb. (P, letter to the editor, 1834)*

As shown in the letter, the writer of this letter goes as far as to mention that using split infinitives is due to the lack of decent education. Scholars presume that this practice may be related to the influence of the Latin grammar, where the infinitive marker and the infinitive verb are a single word and thus cannot be split, as Peters (2006:775) notes.

On the other hand, split infinitives are so frequently used in contemporary English that it would be nonsense to try to prohibit its use. Crystal (2003) advocates split infinitives by claiming that adverbs can be inserted to modify the infinitival verb, just like an adjective modifies a noun. In other words, the following noun phrase and the split infinitive may have the same structure in a sense.

(7) a. the good man
    b. to really understand
However, what Crystal might have omitted is the fact that in (7b) the adverb can follow the verb so that the resulting form *to understand really* can have the same meaning as (7b), while this kind of positional freedom is not allowed in (7a).

According to Crystal (2003), split infinitives can also be used 'to not destroy' (or 'not to destroy') the rhythm and/or meaning of the infinitives. For example, the split infinitive in examples like *to boldly go where no man has gone before* (*Star Trek*) is needed to keep the rhythm.3 If we change the position of *boldly*, then the rhythm changes. As well-known, adverbs like *boldly* receives stress and moving this adverb to other positions result in different prosody. Huddleston & Pullum (2002: 1183-1184) notes that the only function of the particle *to* is to mark the following infinitival verb as an infinitive and does not constitute any syntactic structure with the following infinitival verb. In other words, this particle is not part of the whole infinitive. Thus, separating the particle *to* from the infinitival verb is not against grammar, according to them. This claim, however, is not free of potential problems. If *to* and the infinitival verb has no relationship, then we also note that the finite tensed auxiliary verb and the main verb have no relationship in whatever sense they would intended to mean. As mentioned in the Introduction, we see a distributional difference between the finite and infinitival constructions with regard to the distribution of adverbs.

One of the advantages of using split infinitives is that they remove the ambiguity of the given infinitives (Calle-Martin & Miranda-Garcia, 2009:361). For example, the following sentences have different meanings according to the position of the adverb.

(8) a. You *really* have to watch him.
   b. You have to *really* watch him.

In (8a), the adverb modifies the modal auxiliary *have to*, while in (8b) the adverb *really* modifies the infinitival verb *watch*, thereby producing different meanings. This kind of meaning difference has been one of the most frequently mentioned rationale for using split versus non-split infinitives in the literature. However, the problem with this explanation is that not all cases of split infinitives are of this kind, as will be shown in the following Section of this paper.

---

3 The prosody of *to boldly go* is weak-strong-weak-strong, while that of *to go boldly* is weak-strong-strong-weak.
Split infinitives in English: A corpus-based investigation 59

As an extension of this line of reasoning, we might consider the case of split infinitives with negation such as the following.4

(9) a. not to buy a book
    b. to not buy a book

To clearly mark the scope of negation, we can split the infinitive with negation or put negation in front of the infinitival. It seems that an intriguing fact about negation might be the negative adverb never. Consider the following.

(10) a. I will never see him again.
    b. *I want to never see him again.
    c. cf. I just want to sleep and never wake up.

While (10a) is quite natural, (10b) sounds somehow awkward. If the position of negation and/or any adverb is determined to decide the target of modification, there is no reason why (10b) is not natural. Based on such data, we have to admit that split infinitives are not just a matter of removing ambiguity by placing negation in different positions, as Calle-Martin & Miranda-Garcia, (2009) claims.5

An anonymous reviewer notes that the split infinitive including negation like *to not buy a book is ungrammatical. We do not see any reason why it is ungrammatical. In general negation can split infinitives quite freely. Thus, if this particular case is ungrammatical, the impending factor must not be the splitting itself, but something else, we assume.

Regarding (10c), the negative adverb never is allowed in this case, but notice that it is quite far from the infinitive marker to. This might be due to some distance effect. For example, consider the following:

(i) *My mom and me don't go.
(ii) Me and my mom don't go.

If an offending element is far from the source position, then it tends to be accepted. Thus, (i) is not allowed, because the pronoun me should be in nominative case since it is in the subject position. (ii) is okay, because the pronoun me is far from the subject position and is separated from the case position by my mom. For more discussion, see Jang (2001).

---

4 An anonymous reviewer notes that the split infinitive including negation like to not buy a book is ungrammatical. We do not see any reason why it is ungrammatical. In general negation can split infinitives quite freely. Thus, if this particular case is ungrammatical, the impending factor must not be the splitting itself, but something else, we assume.

5 Regarding (10c), the negative adverb never is allowed in this case, but notice that it is quite far from the infinitive marker to. This might be due to some distance effect. For example, consider the following:

(i) *My mom and me don't go.
(ii) Me and my mom don't go.

If an offending element is far from the source position, then it tends to be accepted. Thus, (i) is not allowed, because the pronoun me should be in nominative case since it is in the subject position. (ii) is okay, because the pronoun me is far from the subject position and is separated from the case position by my mom. For more discussion, see Jang (2001).
4. A corpus–based investigation

CORPORA
For our purpose, we used the Corpus of Contemporary American English (COCA), the Corpus of Historical American English (COHA), and the TIME Magazine Corpus. These three corpora are all American English. COCA carries 45 billion words from 1990 to 2012, COHA carries 40 billion words from 1810 to 2009, and the Time Magazine Corpus carries 10 billion words from 1923 to 2006.

The reason why we have chosen American English is that our initial investigation shows that American English, rather than British English, does allow split infinitives more often. This difference might be due to the fact that American English is more progressive and liberal than its British counterpart. Especially, as will be shown in what follows, spoken English, rather than does allow more split infinitives.6

RETRIEVAL
In this research, we used the following command keys. to [r*] [v?i*] is a command to retrieve any possible token of infinitives containing an adverb between to and the verb. The initial r of [r*] refers to adverbs and the wild key * in [r*] is a command to retrieve any tokens of adverbs. The command [v?i*] is to retrieve any infinitival verbal form.

RESULTS
The use of split infinitives tends to increase in currently English. Based on the corpus data collected from the COHA during 1920 through 2000, the frequency gradually increases, as shown in Figure 1 below.

---

6 A reviewer raises a question regarding why we have investigated adverbs only. Most of the split infinitives allow adverbs between the infinitival marker to and the verb. Other phrases like prepositional phrases are rarely found. That is why we have chosen adverbs for this initial research.
Split infinitives in English: A corpus-based investigation

As shown in Figure 1, use of split infinitives tend to increase. In 1920s there was found no tokens of split infinitives, while in 2000s split infinitives were found to be used 28 times. The number itself might need to be statistically reinterpreted, because the overall size of the corpus gets larger. As the third line of the Figure shows, the frequency ration per million words also increased from 0.00 in 1920s to 4.36 in 2000s. One noticeable fact from the Figure is that split infinitives are more frequently used during the 2000s. If we compare the frequency of split infinitives of the 1920s and 1930s with that of the 2000s, we see a drastic increase. Faculta (2011) notes that use of split infinitives during 1940s and 1980s decreased due to the decreased use of them on the part of literary works.7

The increased use of split infinitives can also be confirmed by the COCA data, which is given in Figure 2 below.

The COCA data also suggest that split infinitives are more often used, as time goes by. The second line shows the raw tokens of split infinitives and the third line the ration of tokens per million words. The token of split infinitives during the 5 year span between 1990 and 1994 is 11,064, while that between 2005 and 2009 is

---

7 A word of caution is in order here. A corpus might be biased in terms of register, style, or resources. Thus when we interpret the corpus data, we have to take into consideration the characteristics of the given corpus. The more corpora, the more reliable results we can get.
Given the data from COHA and COCA, it is clear that the use of split infinitives is increasing and this tendency is quite conspicuous in the 2,000s.

**DISCUSSION**

From the COCA, we retrieved the total 9,689 tokens of split infinitives. Out of this number, 4,797 tokens are from spoken data, and 4,892 tokens are from written data. What this means is that split infinitives do not show any meaningful difference between spoken and written style of speech. The total tokens of infinitives are 4,237,555 and the those of split infinitives are 9,689. Thus, the ratio of split infinitives against total infinitives is 0.22%.

From the COHA, we could retrieve 1,760 tokens. From the Time Magazine Corpus, we could retrieve 362 tokens of split infinitives. One noticeable fact about the increase of split infinitives is that it suddenly increased in 2,000. Why is that so, as an anonymous reviewer asks? Indeed, not only corpora but also any digital data such as google search or newspapers reveal that split infinitives began to surge during this period. We simply conjecture that digital revolution made spoken English wide spread-out and with it split infinitives began to spread out. In a moment, we will see that split infinitives are properties of spoken English, rather than of written English.

Based on the COCA data, we have found that a limited number of adverbs typically appear between the marker to and the infinitival verb, given in (11).

(11) *better, just, really, still, always, even, fully, further, simply, actively, effectively, carefully, accurately, fundamentally, ever, also, seriously*

In particular, intensifiers such as *just, really, actually, better, even, further, fully,* etc. are most frequently found. Table 1 shows the frequency of the most frequently used adverbs in split infinitives from COCA and the TIME MAGAZINE Corpus, respectively.
Split infinitives in English: A corpus–based investigation

Table 1. Most frequently used adverbs in split infinitives

<table>
<thead>
<tr>
<th>COCA</th>
<th>TIME MAGAZINE</th>
</tr>
</thead>
<tbody>
<tr>
<td>to + adverb</td>
<td>frequency</td>
</tr>
<tr>
<td>to just</td>
<td>4,049</td>
</tr>
<tr>
<td>to really</td>
<td>3,795</td>
</tr>
<tr>
<td>to actually</td>
<td>2,296</td>
</tr>
<tr>
<td>to better</td>
<td>2,091</td>
</tr>
<tr>
<td>to even</td>
<td>1,805</td>
</tr>
<tr>
<td>to further</td>
<td>1,342</td>
</tr>
<tr>
<td>to fully</td>
<td>1,298</td>
</tr>
</tbody>
</table>

From Table 1 above, we see that intensifying adverbs such as *just, really, actually,* etc. are frequently used in split infinitives, as shown by the COCA that includes spoken data. This general tendency is also found in the written data according to the TIME MAGAZINE Corpus, which shows that the same adverbs are frequently used with a slight difference in their frequency. That is, in the TIME MAGAZINE Corpus, the intensifying adverb *really* is most frequently retrieved, which is followed by *further, just, even, better,* and *actually.* The adverb *just* takes 41% of the total tokens of all adverbs in the split infinitives retrieved from the COCA corpus. It also takes 28% in the TIME MAGAZINE Corpus, with the most frequent adverb *really* taking 31%. Interestingly, the adverb *actually* takes 24% in the COCA, the 3rd most frequent adverb, while it takes only 13% in the TIME MAGAZINE Corpus. What this might suggest is that *actually* is most frequently used in spoken English, rather than written English. Table 2 shows the relative frequency of those adverbs from the COHA corpus.

Table 2. Frequency of adverbs in split infinitives (COHA, 1800~2000)

<table>
<thead>
<tr>
<th></th>
<th>1800</th>
<th>1830</th>
<th>1860</th>
<th>1890</th>
<th>1920</th>
<th>1950</th>
<th>1980</th>
<th>2000</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>to just</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5</td>
<td>14</td>
<td>33</td>
<td>76</td>
<td>166</td>
<td>295</td>
</tr>
<tr>
<td>to even</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>21</td>
<td>21</td>
<td>25</td>
<td>67</td>
<td>124</td>
<td>263</td>
</tr>
<tr>
<td>to really</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>8</td>
<td>11</td>
<td>34</td>
<td>36</td>
<td>115</td>
<td>208</td>
</tr>
<tr>
<td>to fully</td>
<td>1</td>
<td>0</td>
<td>8</td>
<td>3</td>
<td>3</td>
<td>4</td>
<td>17</td>
<td>83</td>
<td>191</td>
</tr>
<tr>
<td>to better</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>5</td>
<td>10</td>
<td>28</td>
<td>104</td>
<td>150</td>
</tr>
<tr>
<td>to further</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>9</td>
<td>11</td>
<td>10</td>
<td>27</td>
<td>60</td>
<td>121</td>
</tr>
<tr>
<td>total</td>
<td>1</td>
<td>0</td>
<td>23</td>
<td>48</td>
<td>65</td>
<td>116</td>
<td>251</td>
<td>652</td>
<td></td>
</tr>
</tbody>
</table>
Again, we see in the COHA data that similar intensifiers are frequently used in split infinitives. As shown in the table, the intensifier just is most frequently used, just as the COCA data suggest. If we look at the total number of tokens, we see that split infinitives have increased since the 1950s. The most frequently used split infinitives in spoken and written English from the COCA are shown in the table 3 below:

<table>
<thead>
<tr>
<th>split infinitives</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>to better understand</td>
<td>874</td>
</tr>
<tr>
<td>to just be</td>
<td>361</td>
</tr>
<tr>
<td>to really get</td>
<td>349</td>
</tr>
<tr>
<td>to just go</td>
<td>263</td>
</tr>
<tr>
<td>to just say</td>
<td>252</td>
</tr>
<tr>
<td>to really make</td>
<td>206</td>
</tr>
<tr>
<td>to fully understand</td>
<td>191</td>
</tr>
</tbody>
</table>

As shown in the table, we found that the most frequently used split infinitive in English is to better understand, according to the COCA data. This split infinitive takes almost 10% of all the tokens of split infinitives from the COCA. The table also shows the fact that the infinitival verb understand is readily modified by another intensifier fully, which appears 191 times in the COCA corpus.

The corpus that we have investigated reveals that specific adverbs are readily combined with a specific group of verbs, respectively. As just mentioned, we see that the verb understand is split from its infinitival marker to by adverbs like better, fully, and really. Likewise, other adverbs are found with different verbs, as shown in Table 4 below:
Table 4. Adverbs and most likely collocating verbs

<table>
<thead>
<tr>
<th>to just</th>
<th>say, get, stand, sit, stop, let, keep, go, call, do, give...</th>
</tr>
</thead>
<tbody>
<tr>
<td>to really</td>
<td>get, believe, change, understand, know, make, want...</td>
</tr>
<tr>
<td>to actually</td>
<td>see, do, materialize, go, get, help, have, write, vote...</td>
</tr>
<tr>
<td>to even</td>
<td>think, talk, get, begin, consider, know, make, imagine...</td>
</tr>
<tr>
<td>to further</td>
<td>reduce, complicate, improve, develop, increase, assure...</td>
</tr>
<tr>
<td>to better</td>
<td>understand, protect, manage, serve, compete, anticipate,...</td>
</tr>
<tr>
<td>to fully</td>
<td>understand, fund, appreciate, comprehend, enjoy, engage...</td>
</tr>
</tbody>
</table>

It follows from the table that the adverbs splitting the infinitives are generally the ones that can modify the infinitival verbs in finite clause, too. For example, the adverbs better and fully readily modifies verbs like understand, protect, manage, serve, compete, anticipate, appreciate, comprehend, enjoy, engage, etc in a finite clause. In this respect, these adverbs seem to be of nothing special except that they simply modify infinite form of the verbs. However, we are forced to note that there is a big difference in the distribution of these adverbs between finite and infinitive verb forms. That is, only a limited number of adverbs can split the infinitives and only a number of verbs can be split by these adverbs. In contrast, there is practically no restriction on the distribution of adverbs that modify a finite verb. Given this, we cannot adopt Crystal’s (2003) claim that adverbs can be inserted to modify the infinitival verb, just like an adjective modifies a noun. Rather, we have admit that there is a strong semantic relationship between the splitting adverbs and the split infinitival verbs, no matter what the relationship is. Otherwise, we would find rather a free distribution of adverbs, which is not true.8

SOME IMPLICATIONS
The distribution of the adverbs that we have found through the corpus investigation may pose a potential problem for the clausal structure adopted in the generative grammar. Take a look at the clausal structures below:

---
8 A anonymous reviewer points out that the infinitive marker to is in fact not tense or inflection. We have known of no serious proposal about this claim. We will leave this line of research open for future research.
(12) a. Finite clause

```
IP
  Spec
    I'
      I
        [+tns]
          adverb
  VP
    Spec
      V'
        adverb
```

b. Infinite clause

```
IP
  Spec
    I'
      I
        [-tns]
          adverb
  VP
    Spec
      V'
        adverb
```

In terms of clausal structure, there is no meaningful difference between the tensed clause (12a) and the infinitival clause (12b), as shown in the above structure. That is, in the finite clause (12a), the I is tensed, while that of the infinitival clause (12b) is tense-less. However, the distribution of adverbs in split infinitives is quite different from that of the finite clauses: Only a limited number of adverbs can be inserted between the infinitival marker to and the verb. Why is it the case? The Crystal-style explanation discussed in Section 2 might not be able to explain why it is so. As of now, we do not have a concrete solution for this issue, but we simply speculate that a semantic restriction on collation words may be in operation in the distribution of the pre-modifier adverbs. Or that the relationship between the Inflection and the verb is more tightly connected in the infinitival clause than in finite clause. One potentially-supporting evidence might come from the VP-dislocation test in each of these cases. That is, VP-fronting is more readily allowed in finite clause than in infinitival one. Consider the following.

(13) a. I will study English.
    b. Study English, I will.

(14) a. I want there to be a man in the room.
    b. *Be a man in the room, I want there to.

Compare (13b) and (14b). In terms of clausal structure, there is no significant difference. However, VP-fronting in (13b) is okay, but it is not in (14b). Somehow, we need to restrict fronting an element out of infinitival clause, which may not be
5. Summary and conclusion

In this paper, we first reported on the findings that we made from our investigation of the corpora of split infinitives. We investigated COCA, COHA, and the TIME MAGAZINE Corpus to find out the status of split infinitives. In so doing, we have found out (i) that the use of split infinitives is drastically increased recently, especially after the year 2000, (ii) that only a limited number of adverbs, but not any adverbs, are frequently used in split infinitival constructions, (iii) that the most frequently-used adverbs are intensifiers, (iv) that there is a big difference between the distribution of adverbs in a finite clause and that of an infinitival clause, and finally (v) the distribution of adverbs in split infinitives may require some more elaboration of the clausal structure in the generative grammar framework.

References


---

9 An anonymous reviewer raises a question regarding whether or not splitting infinitives is possible in elliptical sentences. A relevant example, provided by the reviewer is:

i) John tried to kiss Mary, but she tried not to/*to not.

As far as we understand, the above example involves another factor. That is, only focused negative NOT is allowed at the end of a sentence. Thus, we presume that (i) is okay, if the final *not* is focused.


