KOREAN JOURNAL OF ENGLISH LANGUAGE AND LINGUISTICS

All-cleft constructions in English: A corpus-based approach* **

Youn-Gyu Park · Jong-Bok Kim (Kyung Hee University)

Youn-Gyu Park (1st author)
M.A. student, Dept. of English Language and Literature, Kyung Hee Univ.
E-mail: youngyu0406@khu.ac.kr

Jong-Bok Kim
(corresponding author)
Professor, Dept. of English Language and Literature, Kyung Hee Univ.
E-mail: jongbok@khu.ac.kr

ABSTRACT

Park, Youn-Gyu and Jong-Bok Kim. 2023.All-cleft constructions in English: A corpus-based approach. *Korean Journal of English Language and Linguistics* 23, 571-586.

The English *all*-cleft construction, consisting of an *all*-cleft clause with a gap, a copula and an element licensed in the post-copula position, displays a filler-gap dependency with connectivity effect. In addition, previous literature asserts, unlike the pseudo-clef, the all-cleft construction cannot be predicational, casting several analytical and empirical questions. Key research questions include if the construction is derived from derivational processes or base-generated, what grammatical properties distinguish the construction from the related cleft constructions, and so forth. To answer some of these research questions, we have performed a comprehensive corpus investigation. Based on our corpus data, we suggest a construction-based approach to the English *all*-cleft that can account for its syntactic, semantic, and pragmatic properties, while capturing its shared properties with related cleft constructions like the pseudo-cleft.

KEYWORDS

all-cleft, connectivity effect, corpus, construction-based, co-varying collexeme analysis

1. Introduction

English *all*-cleft constructions, one of the members of the pseudo-cleft construction family, consist of a pre-copula *all*-cleft following the word *all*, the copula *be*, and a post-copula element (Kay 2013, Tellings 2020).

(1) All I wanted was the right to develop my ideas. (Bonelli 1992)

In (1), the *all*-cleft *all I wanted* serves as a subject of the copula *was*, and its complement NP *the right to develop my ideas* functions as a post-copula element. An *all*-cleft takes a syntactic gap. For instance, in (1), we can find a gap in the direct object position of the verb *wanted*. The *all*-cleft constructions can induce both specificational and predicational readings, as given below respectively, as the other cleft constructions are (cf. Higgins 1979, Declerck 1983, Homer 2019).

(2) a. All I ate for dinner was a salad. (Tellings 2020)
     b. All that John ate for lunch was tasty. (Homer 2019)

The sentence in (2a) induces the meaning that there is something that *I ate for dinner*, and it is the post-copula element *a salad*. Since *a salad* specifies the variable in the *all*-cleft, such readings are specificational. In this case, the post-copula element is called a focus-bearing 'pivot.' Meanwhile, in (2b), the sentence does not imply such meaning but adds something more about it; it implies that there is a set of things that *John ate for lunch*, and they are all *tasty*.[1] Unlike the pivot in (2a), that in (2b) does not specify any variable but predicates all the entities in the set of variables. In this regard, it is predicational. One thing to note is that *all*-clefts in both readings clearly differ from typical relatives.

(3) a. The video [which [you recommended]] was really terrific. (Kim and Michaelis 2020)
     b. All [that [you recommended]] was that video.
     c. All [that [you recommended]] was really terrific.

In (3a), the relative clause implies two propositions; that *you recommended* something, which is *the video*, and *the video was really terrific*. In contrast, the *all*-cleft in (3b-c) implicates that *you recommended something*, but such *something* does not correspond to *all*.

Then, the question that follows is how such syntactically complex sentences can be constructed. In order to look into linguistic properties that the construction displays and account for syntactic aspects of them, the current study conducts a comprehensive corpus investigation with the actual uses of the construction with authentic corpus data from the COCA (the Corpus of Contemporary American English, Davies 2008). In doing so, the paper tries to see whether there are any peculiar, idiosyncratic instances of them as well as to verify whether the derivation-based approaches can cover all the constructs of the constructions.

In the following section, the study briefly sketches basic syntactic patterns of the construction and its constituents and describes the semantic and pragmatic properties of the construction. In section 3, we briefly sketch the previous literature. We focus on how the derivation-based approaches tried to analyze the construction, and in what aspect they might fail to account for the construction. In section 4, we conduct a corpus investigation to look into the actual usage of the construction in real-life. First, we describe the methodology and introduce variables. Then, we report the distributional properties of the construction. We take this result to the discussion in section 5.

---

[1] Unlike Homer (2019), Tellings (2020) mentions that the construction cannot be predicational.

The section mainly deals with the major variable of the construction, where the embedded verb lexeme *do* licenses a gap and a nonfinite verb appears as a post-copula element. We address the information-structure properties of the construction as well. The section 6 summarizes and finishes the paper.

2. Basic properties

2.1. Syntactic properties

The first argument of the construction is the pre-copula *all*-clefts. It contains the relative pronoun *that*, which can possibly be dropped out.

(4) a. [All (that) one has to do] is to start training earlier. (Kay 2013)
     b. [All *(that) surprised me] was that Mary was there. (Tellings 2020)

In (4), the sentences display different distributions of the relativizer *that*: the former can drop it, but the latter cannot. This optionality of the relative pronoun has to do with the position of the gap inside an *all*-cleft. In other words, the embedded clause shows the anti-*that*-trace effect, so the relativizer *that* cannot be dropped out if the gap is in its subject position (Kay 2013, Tellings 2020).

The next component of the construction is the copula *be*. As a member of the cleft construction family, *all*-cleft constructions need a copula *be* to be a syntactic head of their main clauses. In pseudo-cleft constructions, the readings of a pseudo-cleft sentence are known to depend on the types of the copula *be* (Delahunty 1984, Lambrecht 2001).

(5) What John is is silly. (Higgins 1979)
     a. John is $\left\{ \begin{array}{c} \text{this} \\ \text{the following} \end{array} \right\}$: silly
     b. John is an x. $\left\{ \begin{array}{c} \text{(Being) an } x \\ x - \text{hood} \end{array} \right\}$ is silly.

As given in (5), the sentence *What John is is silly* can be read in at least two different ways. One is the specificational reading, construing into sentences like there is a set indicating *John* and the set itself is *silly*. The other one is the predicational reading, where *silly* is one of the properties *John* has. In *all*-cleft constructions as well, the copula *be* is licensed as the head of its main clause as well (Kay 2013, Homer 2019).

(6) a. All you cooked is this soup.
     b. All that John ate for lunch was tasty. (Homer 2019)

The example in (6a) seems to be specificational whereas a sentence in (6b) is predicational. Thus, the pivot *this soup* in the former example would be the only member of *you cooked* whereas in the latter, the post-copula expression *tasty* describes a property of all the members of the set *John ate for dinner*.[2] In this regard, we can assume that the readings of *all*-cleft constructions depend on the types of copula.

---

[2] The reading of the example in (6a) has to do with exhaustivity of the specificational *all*-clefts. Further details will be given in §2.2.

One thing to notice is that the specificational *all*-clefts display the filler-gap dependency between the gap inside the *all*-cleft and the post-copula pivot. However, it is not the case that the filler-gap dependency always holds; the filler-gap dependency often weakens or even disappears in some cases (Kay 2013, Tellings 2020).

(7) a. [All you cooked _____ $_i$] is *this soup$_i$* (= (6a))
     b. [All they could do _____] was *embrace each other*. (Tellings 2020)
     c. [All that John ate _____ for lunch] was *tasty*. (Homer 2019)

As shown, in the example in (7a), which is specificational, the pivot *this soup* can fill the gap inside the *all*-cleft. However, the post-copula element in (7b-c) cannot fill the pre-copula gap. For instance, the VP$_{[bse]}$ *embrace each other* cannot fill the gap in the complement position of the verb *do* (i.e., *They could *do embrace each other*; see section 5.1 for further reasoning). The filler-gap discrepancy can also be observed in the sentence in (7b) in a similar manner.

2.2 Semantic and pragmatic properties

Semantically, the universal quantification reading of the word *all* seems to fade when the construction displays exhaustive reading (Homer 2019, Kim and Michaelis 2020).

(8) a. All John ate for lunch was a banana.
     b. All John ate for lunch was a banana. # He *also* ate a strawberry. (Homer 2019)

Here, although the preceding sentence starts with *all*, the sentence is not construed with universal quantification. Instead, the construction evokes exhaustivity thus the sentence is infelicitous once the second sentence is licensed. Thus, *all*-clefts get infelicitous once another possible entity is licensed as a member of the set to which the post-copula element belongs, as in (8b). On the other hand, one of the unique properties of the construction is that it induces scalar readings (Fauconnier 1975, Kay 2013, Tellings 2020).

(9) a. All I can eat is *half a pizza*. (Kay 2013)   [quantity-scale]
     b. All he is is *a simple employee*. (Tellings 2020)   [fame-scale]

The examples in (9) presuppose different types of scalar models. In (9a), the sentence induces that there is a quantity-based scale in terms of the amount of pizza. On the other hand, in (9b), the post-copula element *a simple employee* implies that there is a scale of fame. Here, we can capture another idiosyncratic reading of the construction: the smallness reading. In (9) again, we can observe that the sentences commonly imply that the post-copula elements are placed at a lower point. In other words, the smallness reading implies that the post-copula element for specificational *all*-cleft constructions "represent(s) a lower point in a presupposed scalar model" (Kay 2013), and implies that the pivot "*isn't much*" (Tellings 2020). However, it is not the case that all the specificational *all*-cleft constructions deliver the smallness reading on any occasion (Kay 2013).

(10) All that one has to do is to start training earlier.
     (Implication: # To start training earlier *isn't much*)

Notice also, that all these idiosyncratic semantic properties are only denoted under the combination of *all*-cleft and the specificational copula *be*, as Tellings (2020) notes (i.e., '*He deserves all he has got*.' does not evoke the exhaustive reading).

3. Previous analyses

The derivation-based approaches assume that the pseudo-cleft constructions, including the *all*-cleft construction, are generated in a different form in the deep structure, and it gets its surface structure after a series of derivational operations (Ross 1972, Tellings 2020).

Tellings (2020) asserts that the *all*-cleft constructions are a phonological realization of a relativization of the exclusive particle *only*. The *all*-cleft constructions have first generated a sentence with *only*, and finally realized as it is, after a series of derivational operations.

(11) a. Ed only eats a salad.   [non-clefted counterpart]
    b. [$_{\text{FOCP}}$ [a salad]$_i$ FOC^0 [Ed only eats t_i]]   [movement to Spec, FocP]
    c. [$_{\text{FORCEP}}$ all$_m$ [$_{\text{TOPP}}$ [(that/what$_i$) Ed only$_m$ eats t_i]$_j$ TOP^0 [$_{\text{FOCP}}$ [a salad]$_i$ FOC^0 t_j]]]   [*all*-relativization]
    d. [$_{\text{FORCEP}}$ all$_m$ [TOPP [(that/ what$_i$) Ed only$_m$ eats t_i]$_j$ is [$_{\text{FOCP}}$ [a salad]$_i$ FOC^0 t_j]]].   [copula insertion]

This analysis can account for the exhaustiveness of the *all*-cleft constructions. Furthermore, as noted by previous studies including Kay (2013), the construction displays the connectivity effects between post-copula and pre-copula elements (Kay 2013, Tellings 2020):

(12) a. All the president$_i$ wanted was [to succeed *himself$_i$*].   [Binding condition A]
    b. All she$_i$ said to me was [that I should call *Mary$_{*i/j}$*].   [Binding condition C]
    c. All they$_i$ did was [embrace *each other$_i$*].   [Reciprocals]
    d. All I said (that) I had*n't* done, was [read *any* books about syntax].   [NPI]

The analysis suggested by Tellings (2020) can account for such connectivity effects, which is one of the constructional properties, as well as it can explain why the anaphoric expressions and NPIs are allowed to be licensed since it assumes that the pivot is first generated in the binding domain or the domain of negators. For the derivational approaches, this syntactic effect can be a piece of evidence that the pivots displaying the effect are moved from the binding domain of its antecedent or the negation in the pre-copula position to the post-copula position. However, it is not the case this analysis can cover all the constructs of the constructions; The derivational approach certainly can cover neither the cases where the matrix pivot is a verb phrase whose syntactic head is in its nonfinite form and the embedded verb is *do*, nor the predicational *all*-cleft constructions:

(13) a. All the samples have to *do* is *contain protein molecules*, and we are sunk. (Ross 1972: 99)
    b. All that we suspect is *under the sea*. (Kim In progress: 236)

The sentence in (13a) would have been derived from a sentence **the samples only have to do contain protein molecules*. The analysis cannot be applied here since the source sentence is ruled out (as mentioned, *do* here cannot be used in its emphatic usage). As for the sentence in (13b), the source sentence would have been '*We only suspect under the water*', which is far from what its *all*-cleft counterpart implicates.

Another analysis of pseudo-cleft construction is the deletion-based approach, first sketched by Ross (1972). Under this view, a pseudo-cleft sentence undergoes a deletion operation eliding a subject and a verb of a main predicate (see among, Ross 1972, Higgins 1979, den Dikken 2006, Kim and Michaelis 2020).

(14) a. [[What John read _____] [was [~~John read~~ novels]]]
    b. What John read was *novels*. (den Dikken 2006)

However, we also fail to predict instances of the predicational *all*-cleft construction, as in the derivation-based analysis:

(15) a. [[All I ate _____ for dinner] [was [*~~I ate~~ healthy ~~for dinner~~]]]
    b. All I ate for dinner was *healthy*.

To verify whether the derivation-based approaches can cover various types of constructs of the construction, we conduct a corpus investigation in the next section. In what follows, we will describe the methodologies and results of the corpus investigation.

4. A Corpus investigation

4.1 Methodology

To propose a plausible, robust analysis of both usage of the *all*-cleft constructions, we conduct a corpus investigation with authentic usages of the COCA (the Corpus of Contemporary American English, Davies 2008).[3] We adopt the following search strings and gathered 500 randomized tokens from the search result from the search strings given below (abbr., PUNC: punctuation; PRON: pronoun):

(16) a. PUNC all PRON
    b. PUNC all that

The search strings above target such tokens as the following, respectively:

(17) a. [...] , however **. All it** means is the trial judge who [...]
    b. [...] operations manual **. All that** 's left now is to sign [...]

Then, we manually filtered out irrelevant examples from the initial dataset of 1,000 tokens, such as:

(18) a. **All compares are** done using the effective address, [...] . (COCA 1994 ACAD)
    b. **All that** other stuff is secondary. (COCA 1999 NEWS)

After the filtering process, 597 among 1,000 tokens of the *all*-cleft construction are left in the final dataset. The result is briefly illustrated in Table 1.

Table 1. Search Strings and Token Distribution in Dataset (Raw Freq.)

Type	Initial dataset	Filtered token	Final dataset
PUNC all PRON	500	53	447
PUNC all that	500	350	150
Totals	1000	403	597

---

[3] As a member of the BYU Corpora family, the COCA contains one billion+ words from eight different registers: ACAD (academic), FIC (fiction), MAG (magazine), SPOK (spoken), TV/M (TV and movie), BLOG (blog), and WEB (webpage). The tokens in the corpus are parsed in terms of their part of speeches so that users can search for instances of constructions with search strings, a string consisting of part of speech information, lexemes, words, or phrases.

The 597 tokens of authentic corpus data then are analyzed with the following variables to see the syntactic patterns and semantic/pragmatic properties regarding the *all*-cleft constructions:

(19) a. Syntactic categories of post-copula elements: NP, VP, AP...
     b. Filler-gap dependency between *all*-clefts and post-copula elements: Violation / Non-violation
     c. Reading: Specificational / Predicational

The first variable 'syntactic categories of post-copula element' is to look into the distribution of post-copula element and check the relation between them and the function of the whole construction (whether it is specificational or predicational). Its levels would be NP, VP, AP, PP, CP, S, and so forth. Second, the filler-gap dependency is to see whether the previous literature can cover all the cases of the *all*-cleft constructions. The last one, 'two readings of *all*-clefts' is to look into how often specificational and predicational *all*-cleft constructions are used in real-life. In addition, we perform a co-varying collexeme analysis to see the overall constructional properties of the *all*-clefts.

4.2 Data distribution

We first looked into the lexical categories of the post-copula elements of the data. Our dataset indicates that the post-copula elements can be realized in various syntactic forms:

(20) All I need is [$_{NP}$ an opening, a moment of weakness]. (COCA 2008 FIC)
(21) All that we see is [$_{AP}$ *noble and good and pure*]. (COCA 2010 MOV)
(22) a. All that did was [$_{VP[bse]}$ make everyone angrier]. (COCA 1997 SPOK)
    b. All that's the latest is [$_{VP[en]}$ covered 24 hours a day on CNN]. (COCA 1991 SPOK)
    c. All I wanted *was* [$_{VP[to]}$ to be with Bethany]. (COCA 1996 MAG)
(23) All I want to know is [$_{wh\text{-}CL}$ how he found out]. (COCA 2003 TV)
(24) All I can remember is [$_S$ my parents had nothing]. (COCA 2010 NEWS)
(25) All that remains is [$_{CP}$ for you to open the gates]. (COCA 2019 MOV)

The most frequent type of post-copula element was NP type (214 tokens, 35.85%), followed by VP$_{[bse]}$ type (170 tokens, 28.48%), CP$_{[that]}$ type (65 tokens, 10.89%), and so forth. The distribution is shown in Figure 1.
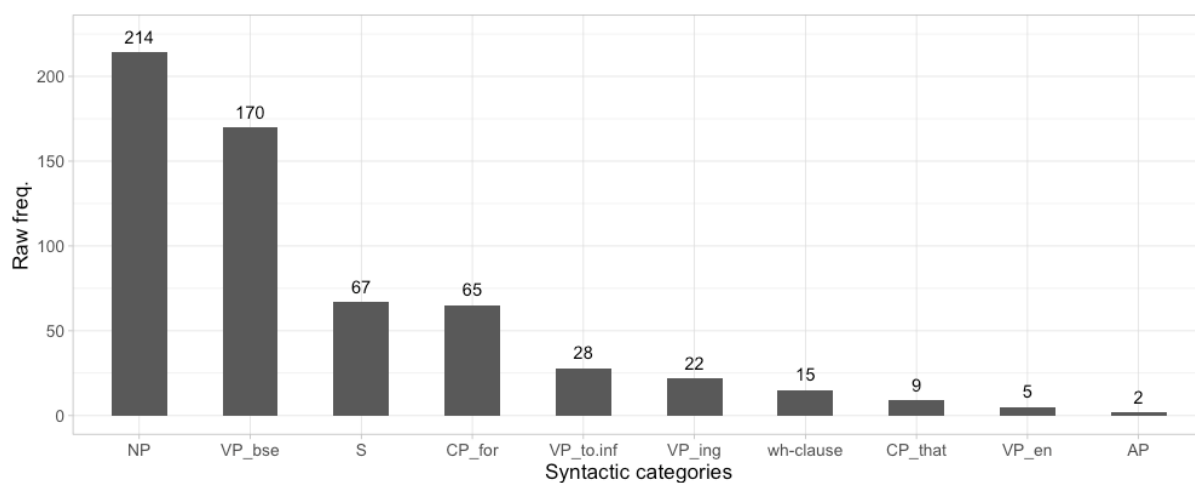

Figure 1. The Distribution of Pre-copula Elements in terms of Their Syntactic Categories (Raw Freq.)

577

Second, we observed how the aspect of the filler-gap dependency between *all*-clefts and post-copula elements. In order to test the filler-gap dependency, we reordered entities inside the sentences in the underlying form with the typical order that Tellings (2020) suggested (see (11) above for more detail). The dataset indicates that overall, the filler-gap dependency between the gap inside *all*-cleft and the post-copula element is preserved:

(26) All you need to add is [NP *a little milk*]. (COCA 2013 MAG)
     = You need to add *a little milk*.

Nevertheless, it also yields examples displaying the filler-gap discrepancy. Here, notice that the verb lexeme *do* cannot serve as the emphatic *do* since it can be licensed in its infinite form, as in (27b-d):

(27) a. All that we see is [AP *noble and good and pure*]. (COCA 2010 MOV)
        = *We see *noble and good and pure*.
     b. All he's going is [VP[ing] *being a cheerleader for them*]. (COCA 1997 NEWS)
        = *He's doing *being a cheerleader for them*.
     c. All that is said and done in public is [VP[en] *observed and recorded*]. (COCA 2005 FIC)
        = *Observed and recorded* is said and done in public.
     d. If I stuck the plug in the socket, all I had to do was [VP[to] *to throw the lever*]. (COCA 1997 FIC)
        = *I had to do *to throw the lever*.

In general, the filler-gap violation cases are often found with VP[bse] post-copula elements (170 tokens, 28.48%), as given in Figure 2.
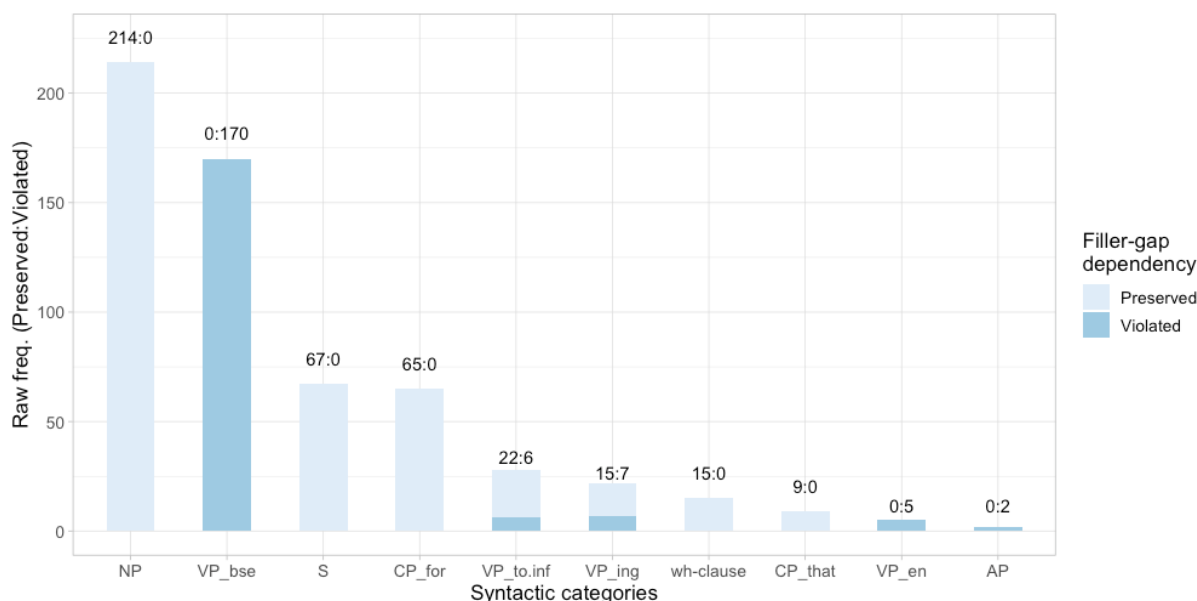

Figure 2. Syntactic Realization of Post-copula Pivots and Their Filler-gap Dependency (Raw Freq.)

One thing that we could observe here is the fact that such VP[bse] post-copula elements are licensed only by the verb lexeme *do*. Here we could assume that there might be some interaction between gap-licensing verbs and a post-copula element at the collocation-level. In order to statistically prove this, we performed a co-varying collexeme analysis (Gries 2022). The analysis is "to quantify how much words in one slot of a construction are

                                                                                            578

attracted to or repelled by words in a second slot of the same construction" (Gries 2019). That is, as for the *all*-cleft constructions, the first slot would be the verb lexeme licensing a gap and the second slot would be the syntactic form of the post-copula element. Before we conduct the analysis, we checked all the verb lexemes in the dataset licensing a syntactic gap in *all*-clefts. The verb lexemes are listed in Table 2 below.

Table 2. Distribution of Gap Licensing Verbs and Their Frequency (Raw Freq.)

No.	V_lex	Freq.	No.	V_lex	Freq.	No.	V_lex	Freq.	No.	V_lex	Freq.
1	*do*	188	2	*know*	51	3	*say*	43	4	*need*	39
5	*want*	23	6	*matter*	22	6	*remind*	22	8	*leave*	18
9	*see*	15	10	*take*	13	11	*hear*	11	11	*have*	11
13	*care.about*	10	14	*be*	8	14	*get*	8	16	*require*	7
17	*tell*	6	17	*be.missing*	6	17	*happen*	6	20	*remember*	5
21	*ask*	4	21	*mean*	4	21	*think*	4	21	*find*	4
21	*think.about*	4	26	*think.of*	3	26	*talk.about*	3	26	*ask.for*	3
29	*hope.for*	2	29	*change*	2	29	*accomplish*	2	29	*bring*	2
29	*be.important*	2	29	*come.out*	2	29	*be.necessary*	2	36	*learn*	1
36	*talk*	1	36	*agree.on*	1	36	*assume*	1	36	*expect*	1
36	*deserve*	1	36	*act.on*	1	36	*become*	1	36	*seem*	1
36	*seek*	1	36	*address*	1	36	*read*	1	36	*write*	1
36	*claim*	1	36	*present*	1	36	*carry.out*	1	36	*give*	1
36	*offer*	1	36	*be.sure.of*	1	36	*go.by*	1	36	*concern*	1
36	*hold.on*	1	36	*love*	1	36	*Imagine*	1	36	*suggest*	1
36	*be.good.for*	1	36	*distinguish*	1	36	*legitimate*	1	36	*coach*	1
36	*achieve*	1	36	*greet*	1	36	*go*	1	36	*come*	1
36	*glitter*	1	36	*be.hold*	1	36	*prove*	1	36	*separate*	1
36	*keep*	1	36	*strand*	1	36	*teach*	1	36	*endure*	1
36	*list*	1									

The analysis was conducted on the R (ver. 4.2.2, R Core Team 2020) environment on the RStudio (ver. 2023.3.0.386), using the package 'Coll.analysis 4.0 (Gries 2022)'. To get a clear result, we got rid of the verb lexemes with frequencies lower than 3. The lexemes excluded from the analysis are given below with their frequencies. Finally, the result indicates that the verb lexeme *do* displays the most strong attraction with the post-copula element whose form is the VP$_{[bse]}$ (LLR = 554.81).[4] The Coll.str. is followed by other combinations such as *say*-S (28 tokens, LLR = 76.90), *need*-NP (36 tokens, 65.55), and so forth.

(28) a. All you gotta *do* is *say yes*. (*do*-VP$_{[bse]}$, COCA 2012 TV)

b. All I can *say* is *thank you for appreciating all my hard work*. (*say*-S, COCA 2008 FIC)

c. All you *needed* was *a good memory*, that's all. (*need*-NP, COCA 1993 FIC)

Meantime, the same verb lexeme *do* shows the tendency of repulsion with other types of post-copula elements such as CP, S, or NP. The detailed result is given in Table 3.

---

[4] Log-Likelihood Ratio (LLR) quantifies whether the distribution of each 'lexeme - post-copula element' pair in a dataset is statistically significant by comparing the raw frequencies of such pairs within the given dataset. In general, an LLR value larger than 3.84 is to be statistically significant ($p < 0.05$) and an LLR value of 15.13 correspond to the *p*-value of 0.0001 (Gries 2019; Brezina 2022).

Table 3. The results of the Co-varying Collexeme Analysis on the Dataset
(Collocational Strength, Coll.str. = Logged Likelihood Ratio, LLR)

W_1 (V_lxm)	W_2 (post.cop.)	W_1*W_2.freq.	W_1.freq.	W_2.freq.	Relation	LLR
do	V_bse	170	188	170	attraction	554.80893
say	S	28	43	66	attraction	76.90187
need	NP	36	39	179	attraction	65.55040
know	CP_that	24	51	57	attraction	51.83528
know	S	21	51	66	attraction	31.42489
care.about	V_ing	6	10	18	attraction	29.76056
see	NP	14	15	181	attraction	24.81946
want	VP_to-inf	8	23	23	attraction	24.67694
have	NP	11	11	181	attraction	24.54558
remain	NP	17	22	181	attraction	18.53018
take	NP	11	13	181	attraction	14.93384
			. . .			
say	NP	0	22	170	repulsion	-14.87831
matter	V_bse	0	22	170	repulsion	-17.02170
remain	V_bse	0	22	170	repulsion	-17.02170
want	V_bse	0	23	170	repulsion	-17.81630
need	V_bse	0	39	170	repulsion	-30.79283
say	V_bse	0	43	170	repulsion	-34.11698
know	NP	1	51	181	repulsion	-35.40693
know	V_bse	0	51	170	repulsion	-40.86588
do	CP_that	0	188	57	repulsion	-52.18782
do	S	0	188	66	repulsion	-61.14694
do	NP	4	188	181	repulsion	-161.55429

To sum up, the result implies that it is not a mere coincidence for the verb *do* to show the highest frequency; the collocation of the lexeme *do* is a major variation of the *all*-cleft construction, indeed. Further, the result proves that it is not a mere coincidence for the verb lexeme *do* to license a VP$_{[bse]}$ post-copula element.

Lastly, the dataset indicates that the construction is usually construed in its specificational reading. On the other hand, examples of the predicational *all*-cleft constructions are also found in the dataset. The examples given below exemplify the specificational and predicational ones, respectively:

(29) a. All that remained of any use was *the desk*. (COCA 2017 FIC)
　　 b. All that's the latest is *covered 24 hours a day on CNN*. (COCA 1991 SPOK)

In (29a), the sentence can be construed as a sentence "*(Only) the desk remained of any use*", which evokes the exhaustive reading. Meantime, in (29b), this kind of reading is unobservable, since it is interpreted as a sentence like "*All the latest things are covered on 24 hours a day on CNN*". The distribution is given in Table 4.

Table 4. Distribution of the Predicational and Specificational Copula (Raw Freq.)

Reading type	Specificational	Predicational	Total
Raw Freq.	590	7	597

Given the distribution and the examples, we can assume that the construction can be construed both in its specificational and predicational reading, even if the latter cases are not so frequently found.

5. Discussion

5.1 The embedded verb lexeme *do*

We have seen that the construction can license a gap licensed by the verb lexeme *do*, and it is to be filled by a VP[bse] post-copula element. These cases can be a counterexample that weakens the relativization-based approach suggested by Tellings (2020).

(30) a. [All that [it has to be]] is *a law*. (COCA 2018 ACAD)
b. It has to be *a law*.

According to Higgins (1979), the pivot in the pseudo-cleft constructions serves as a focus. That is, the pivot receives a focus if we reorder the *all*-cleft constructions into the typical sentence. In (30a), the pivot *a law* serves as a focus in the given context and can be safely reordered into a typical sentence as given in (30b). We may assume that the verb lexeme *do* is an emphatic one, but a sentence cannot posit the emphatic *do* in some cases. Consider the following example:

(31) a. John DID leave.
b. *John did *leave*. (Wilder 2013)

In (31), we can observe that the emphatic *do* can put a focus on itself but cannot on its complement; In (31a), *do* does not simply emphasize the action denoted by the lexeme *leave*. Rather, it emphasizes "the state of affairs expressed in one proposition is in contrast to or in opposition to a state of affairs to which it is being compared" (Soppelsa 1980). In this regard, the verb *did* in (31b) seems to be abundant thus the sentence is infelicitous (Soppelsa 1980, Wilder 2013).

Now, compare the following examples in (32) with (30-31):

(32) a. All we can *do* is *try to do our best*. (COCA 1990 FIC)
b. *We can DO *try to do our best*.

Given the assumption of Higgins (1979), we can assume that the pivot *try to do our best* receives the focus.

If we reorder the sentence as we did in (31), we would get the sentence in (32b), which is infelicitous since the emphatic *do* is licensed not to emphasize its complement. Instead, we propose that this is the realization of the construction-level interaction of the specificational *all*-cleft constructions and the *do-be* construction, such as given below:

(33) **English *do-be* constructions:**
a. [THING-CL The thing I'm doing] is *trying to learn from my mistakes*.
b. [ALL-CL All the government does] is *send out checks*.
c. [WH-CL What you have to do] is *get ready*. (Flickinger and Wasow 2013)

This interaction between constructions cannot be analyzed without considering the network system of constructions (Goldberg 2006). One restriction on the *all*-cleft constructions interacting with the *do-be* construction is that the pre-copula *do* can be licensed either its finite or infinite form, but the following post-copula verb requires to be in its base form or to match the infinite form of the preceding *do*.

5.2 Semantic representation of all

We have checked that *all*-cleft constructions can be used both in specificational and predicational ways as the authentic corpus data shows. However, we also found that the exhaustiveness can only be observed in the specificational ones, as observed in other literature (Homer 2019, Tellings 2020). This pragmatic property of the *all*-cleft constructions parallels that of the pseudo-cleft constructions, as noted by Higgins (1979).

He asserts that specificational pseudo-cleft constructions take three characteristic pragmatic factors: presupposition, focus, and assertion. This is because specificational pseudo-cleft constructions posit a focus on its pivot in the matrix clause and induce the embedded clause to be presupposed in the given discourse (Erteschik-Shir 1986, Collins 1991, Lambrecht 2001).

(34) A: What did you need?
    B: What I need is *a sheet of paper and a pencil.*
    a. PRESUPPOSITION: 'speaker needs *x*'
    b. FOCUS: 'a sheet of paper and a pencil'
    c. ASSERTION: '*x* = a sheet of paper and a pencil' (Lambrecht 2001: 5)

Here, the pseudo-cleft *what I need* serves as a pragmatic presupposition and introduces a variable *x* that is construed as a direct object of the embedded verb *needs*. Then, the construction posits a focus on the pivot. Finally, the pragmatic assertion links the variable licensed by the presupposition and the focus. Granting the observation, we can sketch the semantic representations and their information-structure of specificational *all*-cleft constructions as described below (cf. Higgins 1979, Lambrecht 2001):

(35) **Specificational all-cleft construction:**
    All I need is an address. (= 'I need nothing but an address')
    a. $\iota x[\text{thing}(x)\ \&\ \text{need}(i, x)]$: an address
    b. PRESUPPOSITION: 'I need something x'
      FOCUS: 'an address'
     ASSERTION: 'x = an address'

This exhaustiveness can be semantically represented with the iota operator ι, expressing the uniqueness of the bound variable (cf. Cann 1993, Kim In progress). Thus, the sentence presupposes that '*I need something x*,' which is unique, and the pragmatic assertion links the variable with the pivot '*an address*.' Finally, the specificational *all*-cleft construction in (35) induces the meaning of a sentence like '*I need nothing but an address.*'

The other type of *all*-cleft construction we could observe was the predicational one. Before we investigate the *all*-clefts, consider the predicational pseudo-cleft construction first (Lambrecht 2001):

(36) a. What he bought is *expensive*.
    b. PRESUPPOSITION: 'He bought some thing *x*'
      FOCUS:?'expensive'
     ASSERTION: ?'*x* = expensive'

As we can see, the predicational pseudo-cleft cannot be analyzed with Higgins (1979)'s pragmatic analysis. This is because the predicational pseudo-clefts serve as "pseudo-adjectives" (Lambrecht 2001: 494). In this regard, things would be quite different in the predicational *all*-cleft constructions.

(37) **Predicational *all*-cleft construction:**
    All that we see is noble and good and pure.
    (= '*We see a set of something x and all the members of such set are noble and good and pure.*')
    a. $\forall x[thing(x)\ \&\ see(we, x)\ \&\ noble(x)\ \&\ good(x)\ \&\ pure(x)]$
    b. (PREDICATE)-FOCUS: 'All x that we see'
    PREDICATE: 'noble, good and pure'

Here, *all* is construed as the universal quantifier, unlike that in the specificational *all*-cleft construction. Thus, instead of addressing a presupposition, we propose that the predicational *all*-cleft constructions posit a predicate in the post-copula position and it predicates the pre-copula elements, which denote a set of variables bound by the universal quantifier. This assumption concords with the function of other types of predicational pseudo-cleft construction, which can cover the relations between pseudo-cleft constructions and the types of copula.

5.3. Two types of copula *be*

We have shown that *all*-cleft constructions in English can be construed not only in its specificational reading but also predicational one. This property can be observed throughout all the pseudo-cleft construction family (among others, Higgins 1979). In pseudo-cleft constructions, the matrix clause needs to take a copula *be* as its syntactic head. Further, it is widely assumed that it is the type of copula that decides the type of pseudo-cleft constructions (for more detail, see Higgins 1979).

(38) a. What John ate for lunch was a banana and a strawberry.
    b. What John ate for lunch was tasty. (Homer 2019)

All-cleft constructions also need a copula *be* to be its matrix clause head (Kay 2013, Homer 2019). Given this syntactic parallelism, we can assume that it is the type of copula that decides the type of *all*-cleft constructions. Thus, we would like to propose a construction-based approach on the *all*-cleft constructions (Kim In progress).

(39) a. Specificational Copula ($\uparrow$ *cop-lxm*)

$$\begin{bmatrix} spec\text{-}cop \\ \text{ARG-ST} \quad \langle \text{XP}[\iota x[\text{P}(x)]], \text{YP}_y\rangle \\ \text{SEM} \quad\quad identify\_rel(x, y) \end{bmatrix}$$

    b. Predicational Copula ($\uparrow$ *cop-lxm*)

$$\begin{bmatrix} pred\text{-}cop \\ \text{ARG-ST} \quad \langle \text{XP}, \text{YP}\rangle \\ \text{SEM} \quad\quad [\text{YP}(\text{XP})] \end{bmatrix}$$

The two copula constructions given in (39) are to be the heads of each type of *all*-cleft constructions. The specificational copula in (39a) identifies the relation between the denoted entity in an *all*-cleft and post-copula, focus-bearing pivot. The iota operator in its specifier constituent implicates that the entity denoted by an *all*-cleft is unique, thus it can account for the exhaustiveness of the construction. The predicational copula, on the other hand, simply links the relation between pre- and post-copula element, where the latter predicates the former as a pseudo-adjective. Granting this construction-based view, we can notate the feature structures of the specificational and predicational *all*-cleft constructions as given below, respectively:

(40) a. All you need *is* love.

$$\begin{bmatrix} \textit{spec-cop} \\ \text{ARG-ST} & \langle \text{NP}[\iota x[\textit{need}(\textit{you}, x)], \text{NP}_y] \rangle \\ \text{SEM} & \iota x[\textit{need}(\textit{you}, x)]: \textit{love}_y \end{bmatrix}$$

　　b. All you need *is* in this box.

$$\begin{bmatrix} \textit{pred-cop} \\ \text{ARG-ST} & \langle \text{NP}_{[all]}, \text{PP}_{[in]} \rangle \\ \text{SEM} & \forall x[\textit{need}(\textit{you}, x) \ \& \ \textit{in}(x, \textit{box})] \end{bmatrix}$$

The feature structures above notate the argument-structure (ARG-ST) of each copula. Again, what is important here is the semantic (SEM) property of each; whether they specify the relation between the denoted entity of *all*-clefts and the pivot or simply the post-copula element predicates the set denoted by an *all*-cleft.

This non-derivational, construction-based view on the *all*-cleft constructions has two advantages in general. First, since non-derivational, the analysis can properly cover not only the specificational all-clefts with canonical pivots, but also idiosyncratic post-copula elements, VP$_{[bse]}$ or AP post-copula elements, for instance. Second, as it is construction-based, the analysis can account for the constructional properties of the all-cleft constructions, that are shared with the other members of the pseudo-cleft construction family.

5.4. Smallness reading

One of the unique semantic properties of the construction is the smallness effect, implying that something "*isn't much*" (Kay 2013; Tellings 2020). In the dataset, we have found several examples with the smallness reading. One common factor they display is that the pivots somehow take a contrastive focus in the given discourse. Compare the following sentences:

(41) a. All I got is [$_{FOC}$ *the garage*].
　　　(= 'I got nothing more than the garage' $\Rightarrow\!\!\!\!/$ The garage isn't much)
　　b. All I got is [$_{FOC}$ *the garage*]. You got [$_{CONT.FOC}$ *the whole rest of the damn house*].
　　　(= 'I got nothing more than the garage' $\Rightarrow$ The garage isn't much)

As given, the sentence in (41a) does not imply that *the garage isn't much*. Of course, it is certain that this is not the only factor for an *all*-cleft to implicate the smallness reading. However, as one of them, we assume that this is because it does not take any contrastive focus. This assumption can be supported by the example in (41b). Unlike (41a), we can observe that the pivot *the garage* takes a contrastive focus that was evoked by the focus of the following utterance, *the whole rest of the damn house*.

5.5. On connectivity effects

The last issue is the connectivity effects. If we assume that the embedded pre-copula clause induces a series of possible situations, and the situation is specified by the specificational copula, it would not be surprising that the construction displays such peculiar syntactic effects. For example, consider the following examples partially repeated from (12):

　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　　584

(42) a. All the president$_i$ wanted was [to succeed *himself$_i$*].   [Binding condition A]
 ⇒ { s_0: *want* (*i*, *succeed* (*i*, *i*))}
  b. All I said (that) I had*n't* done, was [read *any* books about syntax].   [NPI]
 ⇒ { s_0: *say* (*i*, ¬*read* (*i*, *b*))}

In (42a), the embedded clause *the president wanted* induces a series of possible situations that would be a complement of the verb *wanted*. Then, the specificational copula licenses the pivot *to succeed himself*. Since the complement of the verb *succeed* and the subject of the embedded verb *wanted* represent the same person *the president*, the complement of the verb *succeed* is realized in its reflexive form. In (42b), the embedded clause *I said I hadn't done* induces a series of possible situations which are negated by the negation *not*. In a similar manner to the previous one, the specificational copula licenses the pivot, but the predicate itself is inside of the negation, the entity *b* requires to be realized in the form of NPI (Ladusaw 1980, Giannakidou 2002). If we assume this kind of set-based approach, we can predict the instances of the *all*-cleft constructions without assuming any derivational operations.

6. Conclusion

This study starts by introducing key properties of English *all*-cleft constructions and questioning the validity of the derivation-based approaches. The derivation-based approaches might fail to catch how the construction can induce its predicational meaning, or even fail to predict one of the major constructs of the specificational reading of the construction, where the embedded gap-licensing verb is the verb lexeme *do*, and the post-copula element is a nonfinite verb. In order to provide a robust, plausible analysis of the construction, we conducted a corpus investigation to look into the linguistic patterns of the *all*-cleft constructions with authentic usages from the COCA. The results revealed that the derivational approaches cannot predict the various usages of the constructions (i.e., predicational *all*-clefts, the interaction with the *do-be* construction). Based on the corpus investigation, we first assumed that the derivation-based approaches cannot properly cover the constructs of the *all*-cleft constructions, and discussed how and why such properties can be displayed in terms of syntactic, semantic, and pragmatic perspectives.

References

Bonelli, E. T. 1992. 'All I'm saying is...': The correlation of form and function in pseudo-cleft sentences. *Literary and Linguistic Computing* 7(1), 30-42.

Brezina, V. 2018. Statistics in corpus linguistics: A practical guide. Cambridge: Cambridge University Press.

Cann, R. 1993. *Formal semantics: An introduction*. Cambridge: Cambridge University Press.

Collins, P. 1991. *Cleft and pseudo-cleft constructions in English*. New York, NY: Routledge.

Davies, M. 2008. *The Corpus of Contemporary American English (COCA)*. Available online at https://www.english-corpora.org/coca/.

Declerck, R. 1983. Predicational clefts. *Lingua* 61, 9-45.

Delahunty, G. P. 1984. The analysis of English cleft sentences. *Linguistic Analysis* 13(2), 63-113.

den Dikken, M. 2006. Specificational copular sentences and pseudoclefts. In M. Everaert and H. van Riemsdijk, eds., *The Blackwell companion to Syntax*. 292-409. Malden, MA: Blackwell.

Erteschik-Shir, N. 1986. *Wh*-questions and focus. *Linguistics and Philosophy* 9(2), 117-149.

Fauconnier, G. 1975. Pragmatic scales and logical structure. *Linguistic Inquiry* 5(3), 353-375.

Flickinger, D. and T. Wasow. 2013. A corpus-driven analysis of the *do-be* constructions. *The core and the periphery: Standing on the shoulders of Ivan* A. *Sag*. 35-63. Stanford, CA: CSLI Publications.

Giannakidou, A. 2002. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins.

Goldberg, A. 2006. *Constructions at work: The nature of generalization in language*. New York, NY: Oxford University Press.

Gries, S. Th. 2019. 15 years of collostructions: Some long overdue additions/corrections (to/of actually all sort of corpus-linguistic measures). *International Journal of Corpus Linguistics* 24(3), 385-412.

Gries, S. Th. 2022. Coll.analysis 4.0. A script for R to compute perform collostructional analyses. Available online at https://www.stgries.info/teaching/groningen/index.html.

Higgins, F. R. 1979. *The pseudo-cleft construction in English*. New York, NY: Routledge.

Homer, V. 2019. That's all. In R. Stockwell, M. O'Leary, Z. Xu and Z. L. Zhou, eds., *Proceedings of the 36th West Coast Conference on Formal Linguistics*. 1-21. Somerville, MA: Cascadilla.

Kay, P. 2013. The limits of construction grammar. In G. Trousdale and T. Hoffmann, eds., *The Oxford handbook of construction grammar*. 32-48. Oxford: Oxford University Press.

Kim, J. B. In progress. *Cleft constructions in form and function mismatch in the English: A construction grammar approach*.

Kim, J. B. and L. Michaelis. 2020. *Syntactic constructions in English*. Cambridge: Cambridge University Press.

Ladusaw, W. 1980. *Polarity sensitivity as inherent scope relations*. New York, NY: Garland.

Lambrecht, K. 2001. A framework for the analysis of cleft constructions. *Linguistics* 39(3), 463-516.

R Core Team. 2020. *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Available online at https://www.R-project.org/.

Ross, J. R. 1972. Act. In D. Davison. and G. Harman, eds., *Semantics of natural language*, 70-126. Dordrecht, Netherlands: Reidel.

Soppelsa, E. F. 1980. Simple and complex contradiction and the used of emphatic *do*. In M. Henderson, ed., *Papers of the Mid-America Linguistics Conference*, 413-422.

Tellings, J. 2020. An analysis of *all*-clefts. *Glossa: A journal of general linguistics* 5(1), 1-25.

Wilder, C. 2013. English 'emphatic *do*'. *Lingua* 128, 142-171.

Examples in: English
Applicable Languages: English
Applicable Level: Tertiary