

# Corpus based study of verbs *explain* and *clarify* as an example of assistance in pedagogical settings

Maja Séguin

*Evangelical Theological Seminary, Osijek, Croatia*

## Abstract

The use of synonymy and near-synonymy allows us to differently express similar ideas and meanings, as well as perspectives. However, their use and nuances may be unclear to language learners, such as the selected case of the verbs *explain* and *clarify*. This paper examines the usage of the two verbs by looking into corpus data and uses Sinclair's methodological procedure as an alternative to dictionary references. Also included is a discussion of aspects and criticisms of corpus linguistics, mentions (and uses of) computer technologies for the analysis of language by discovering usage patterns, significant exceptions and semantic prosody, and exploring whether using corpora in the classroom would be beneficial to language learners.

**Key words:** synonym; near-synonymy; corpus; corpus linguistics; British National Corpus (BNC); Google Books corpus.

## 1. Introduction

When using language, we face the possibility of expressing ourselves in different ways but portraying the same or similar meaning. Synonymy and near-synonymy are communicative tools which allow for the presentation of alternate perspectives of the same idea (Divjak, 2010:2). They are different words, belonging to the same part of speech, which can be used in various contexts, leaving the meaning of the sentence intact. To understand the shades of meaning similar lexemes produce, as well as to make appropriate communicative choices and to help language learners, further investigations into usage of similar-in-meaning words should be conducted. This paper examines the usage of two synonymous verbs, *explain* and *clarify* after noticing several EFL students using the terms interchangeably in an academic context and ignoring the potential for nuances that could exist. This analysis of the two verbs is intended to serve as an example of the analysis a language learner could apply to gain a deeper understanding of which (near-)synonym is most appropriate for a given context.

Corpora are used for distributional and collocational patterns, seeking meaning and frequency differences in various contexts, to suggest accurate selection when communicating, as well as to elicit implications for language instruction. The chosen verbs share the same basic dictionary definition, “to make clearer or easier to understand” (Cambridge Advanced Learner's Dictionary, 2005; The American Heritage English as a Second Language Dictionary, 1998; Longman Dictionary of American English, 1983), but their usage seems to differ when it comes to what can be *clarified* or *explained* (e.g. *clarify* cannot replace *explain* in *explain your blackened eye*, *explain your noisy behavior*). Insight into corpus data provides more information on tendencies than looking up the definition in dictionaries that can help non-native speakers.

This paper starts by presenting the theory used in section 2. The general steps of the analytical methodology used form section 3. Section 4 combines the results of the analysis with discussion. The conclusion in section 5 summarizes findings as to the finer aspects of the usage of the two verbs.

## 2. Theory

This literature review summarizes definitions and previous works on synonyms, corpus linguistics, collocates and colligates, pedagogical aspects and Sinclair's methodological procedure as used in this paper.

### 2.1. Synonyms

Perfect synonyms are words that convey identical meaning (Taylor, 2002: 265) and could thus be used interchangeably. Some consider such words to be rare while others question if these exist in this absolute sense (Inkpen & Hirst, 2006: 223). If this is true, then dictionaries of synonyms are actually lists of near-synonyms. Near-synonyms are words that have a highly similar, not identical, meaning and “a low degree of implicit contrastiveness” (Cruse, 1986:266), that is, they do not tend to exclude one another, such as *warm* and *hot* might (Taylor, 2002: 266).

The choice of near-synonym will depend “on the context and on the nuances that need to be conveyed” (Inkpen & Hirst, 2006: 224). While some (Divjak, 2008: 22) look for objective means to determine if words are near-synonyms, this paper remains on the level of introspection that a teacher of English as Second Language could rely on in a classroom setting.

## 2.2. Corpus

Texts making up a corpus, chosen by external criteria, are pieces of language meant to represent language variety (Sinclair, 2005a), as well as choices and competence of the language users documented in texts, but also the impact of social circumstance on communicative patterns (Andersen, 2010: 548). Computer technology has allowed extensive and rapid analysis of language use by handling large collections of texts as well as more easily identifying linguistic features such as “association patterns” (Biber et al, 2000: 4) and repetitions (Hunston, 2002: 123). It is concerned with frequent and typical occurrences in language, as well as relations between instances and norms (Stubbs, 2001: 151).

Scholars believe a corpus gives an insight into language and believe it to be a reliable guide as opposed to using native speaker intuition (Hunston, 2002: 20). Hunston states that corpora ‘lead to new descriptions of a language,’ and as such can assist in pedagogical settings when making teaching materials, syllabus and revising methodological approaches (Hunston, 2002: 137).

Although corpus linguistics has become a crucial methodological tool extensively accepted in the field (McEnery & Hardie, 2012: 226), the creation and use of corpora have provoked criticism and “Mount Everest questions” (Biber & Finegan, 1991: 205), as well as differing opinions on the usefulness of corpora even among the advocates (McEnery et al, 2006: 131). Widdowson (2000: 6) points out that corpus linguistics does not provide an insight into the language users or a measure of their competence, as it only displays textual instances of meaning-making, lacking the complexity of other linguistic, non-linguistic and contextual factors. Additionally, he notes that it lacks a defined pedagogical approach to be appropriate for classroom use. As it is, corpora reflect decontextualized language, which should be contextually made appropriate for learners to consider it real and motivating (Widdowson, 2000: 7). This individual subjectivist approach, stressing the importance of learning language through dialogue and context, sees an issue of missing out the elements critical to meaning which may not be reproduced in quantitative evaluations, questioning the learning outcomes from words that are removed from their context and whose meaning is constantly evolving. Another criticism involves the danger of using corpora by focusing on “surface forms” (Ädel, 2010: 49), only the word or phrase, which then the analyst needs to connect to meaning presuming s/he holds the knowledge and views language objectively. Some have raised the question of incompleteness of computer corpora, indicating that there may be words or phrases that are imaginable and possible, but not recorded in the corpus (Flowerdew, 2011: 32). Hunston (2002: 160) also explains the difficulty of compiling, and therefore using, comparable corpora when there

is no access to older texts. Besides the problem of representativity, Emons (1997: 66) sees a problem with grammatical mistakes potentially interpreted as the dialectics of variants of native speakers or not detected at all by corpus linguistic methods. When it comes to spelling variations, analysts face difficulties in the frequency and collocation calculations (Hunston, 2002: 160).

Regardless of debates and skepticism, use of corpora can be considered as a generally accepted and important linguistic tool as evidenced by its rapid development in semantic research (Glynn, 2014: 1) and fast growth as a methodological discipline in linguistics generally (Gries, 2015: 1). Sinclair (2005b: 101) warns “to avoid perfectionism,” as there is no ideal corpus and “the results of corpus research so far are indicative of patterns and trends, of core structures and likely contributions to theory and description, but they are not yet definitive.” The benefit of a corpus usually depends on the researcher’s intent, and as corpora do not inevitably provide descriptions, the analyst is there to draw explanations. Researchers’ task is to notice repetitions as well as implicit meaning, using methodologies that need to be explained by steps found between observation and interpretation (Hunston, 2002: 123). And yet, the ideal text analysis system would be completely automated, consisting of both linguistic and non-linguistic knowledge of a native speaker (Bell, 1991: 296).

### **2.3 Collocates and colligation**

Corpus has shown itself crucial in statistically measuring collocations, or tendencies of lexical words to co-occur. Observing the frequencies, noticing the regularities and comparing the co-occurrences is a part of the methodology. The observed data can provide information on noticeable concepts, such as change in frequency, differences and change in language use (Hunston, 2002: 120). Collocations can also identify lexical sets (Halliday, 1966: 156). For a reliable measure of collocation, it is not sufficient to know an absolute count, as much as there needs to be evidence for certainty of co-occurrence (Hunston, 2002: 72). This evidence is provided by statistical calculations or association measures (AMs) including commonly used methods such as log-likelihood, Mutual Information, t-score, and others (Gries, 2015: 95). While the use of such AMs may lead to a skewing of results and other problems, for example if dealing with sequences of words (Gries, 2015: 95), it can also be successful under certain circumstances such as if working with higher frequencies (Gries, 2015: 105). While a list of collocates by itself cannot show the association to meaning, necessary interpretation can reveal attitudes (Hunston, 2002: 120).

Colligation is the relation between words at the grammatical level, such

as co-occurrence of a word or phrase with grammatical classes or grammatical words (McEnery et al, 2006: 11). It is harder to identify colligation than collocation in a corpus that has not added grammatical or “part of speech” (POS) tagging although this can be added to an untagged text by tools using sophisticated statistical models and other rules to “achieve fairly accurate tagging” (Rayson et al, 2007: 2). These lexical and grammatical elements are crucial for analyzing the meaning of a word that co-occurs with them (Sinclair, 1991: 108).

## **2.4 Semantic preference, semantic prosody, frequency distribution**

As collocations give an insight into attraction between specific word forms, colligation into co-occurring grammatical classes, semantic preference shows textual coherence by observing collocate lists for attraction between words and certain semantic fields (Stubbs, 2009: 125).

Stubbs, using phrasal verbs (1995: 2) also presents Sinclair’s observation that there are words that co-occur with other words creating negative or positive meaning, calling this collocational phenomenon ‘semantic prosody.’ This refers to the phenomenon of a word or a phrase adopting the connotation of its setting carrying a hidden message or attitude (Hunston, 2002: 141). It represents the motivation and it is often not retrievable from conscious knowledge.

Hunston (2002: 161–163) notes how distribution, the uneven frequency of many words across registers and grammatical features including verb forms, can indicate different meanings. For example, negative forms and the present tense occur more frequently in conversations than in written academic prose, the past tense more in fiction, infinitive patterns more in expository prose and fiction than in conversation.

Stubbs (2001: 153) states that the impact of the previously mentioned semantic patterns must be further explored as individual speakers do not control them. However, they are interrelated and have an effect and influence on linguistic features and preferences.

## **2.5 Pedagogical aspects**

While corpora can be used indirectly in the process of creating educational materials, it is also possible to have learners use the material directly (Corino & Onesti, 2019: 2). Data driven learning (DDL), used to refer to the application of tools of corpus linguistics by language learners, has “every student a Sherlock Holmes” (Johns, 1991 in Corino & Onesti, 2019: 2). In such cases, learners are guided to query corpora for the purpose of learning

or as a reference resource (Cobb & Boulton, 2015: 480).

Although “quite demanding on the learner” (ibid), it has been stated that this helps students learn language in context, increases their autonomy and equips them with a number of skills (Gilquin & Granger, 2010: 1). Studies have shown that DDL and corpora usage generally show large positive effects (Boulton & Cobb, 2017: 378). Learners’ direct use of computer technologies over the usage of handouts has shown even greater effects (Boulton & Cobb, 2017: 372).

## ***2.6 Sinclair’s corpus analysis and patterns***

Identifying patterns in text usage is done by looking for words and structures frequently occurring together and defining a meaning (Hunston & Francis, 2000: 37). Latent patterns, as referred to by Sinclair and Coulthard, are patterns that “tend to remain unobserved when the same words or phrases are met in their normal contexts” (Hunston, 2002: 9). Sinclair (2004: 17–18) describes six types of patterns. He considers patterns based on variations of a lemma (‘move’ vs ‘moving’) and of word class (‘combat’ as noun vs verb), “privileges of occurrence or restrictions in group structures” (Sinclair, 2004: 18) and traditional categories (for example when ‘of’ is used as preposition). Finally, there are subliminal “aura of meaning” (ibid) patterns (such as ‘set in’ tending to present negative situations such as bad weather) and there are collocations. He (2004: 10) also cautions regarding theories related to patterned data, that “prioritizing some patterns” (ibid) could obscure others. Corpus Pattern Analysis (CPA) considers that patterns are not likely to lead to an understanding of meaning if verbs and nouns are taken as words in isolation (Hanks, 2004: 88). Rather, the analysis of words in the context of lexical sets can reduce the confusion for non-native speakers (Hanks, 2004: 91).

It has been argued (McEnery et al, 2006: 147) that Sinclair’s corpus analysis provides a method. Stubbs (2009: 115) summarizes Sinclair’s method of analysis as searching for patterning across large collections of long and authentic texts using computer technology. Such technology uses quantitative, observational techniques. Concordance analysis is of special interest as it is the most familiar of all corpus linguistic methods and is useful in investigations of individual word-forms. Sinclair (2003: xvi–xvii) describes his procedure as following a number of steps: initiate (identify first patterns from repeated words or grammatical word classes), interpret (form a hypothesis for the pattern), consolidate (gather additional evidence), report (write out the hypothesis), recycle (find additional patterns), result (list all hypothesis), and repeat (look at another sample).

Sinclair (ibid) considers it unwise to analyze all instances if there is a

larger number of results and that computers should do the “routine work” even if software cannot take over and “complete an exhaustive description of your data” (2003: xviii). Sinclair has no explicit step applying searches against the complete corpus. While a few lines may prove a pattern exists, it will not show statistical relevance.

### 3. Methodology

The primary corpus used in this paper is the 100 million-word British National Corpus (BNC) composed of 90% written and 10% spoken language samples, and consisting of various registers and contexts. Analysis was performed using the BNCweb tool hosted by Lancaster University.

The Google Books corpus with 361 billion words (Michel et al, 2011: 176) was used for a high-level overview of the popularity of the terms through time. To have a better comparison to the BNC results, the British English subset was chosen. Although the Google Books corpus is larger, the inability to see words in context and general tooling limited it to this secondary role.

The first step was to obtain a general picture of the usage of the two verbs, using the Google Corpus n-gram viewer, giving the relative frequency through time and BNCweb’s distribution function across different metadata categories and sub-corpora.

The second step was to find patterns for the verb *clarify* regardless of tense. This followed Sinclair’s model for discovering usage patterns, with 100 random lines as the sample, but focusing on patterns that could help define what is the object of the verb. Possible patterns in the sample were identified as individual hypotheses for further investigation. Where Sinclair’s ‘reporting’ step calls for an “explicit, testable version” of each hypothesis, search queries matching the patterns were defined to verify patterns’ relevance compared to all uses of the verb in the corpus. Finally, as an extension of Sinclair’s process, collocations from the corpus were retrieved to see if there are significant exceptions.

During this step, identification of patterns within the sample relied on human observation, while counts for collocations and colligations were retrieved through BNCweb tooling. While identifying patterns, possible semantic prosodies were noted. When working with the small samples, the results were ordered based on raw frequency. When working with the full corpus, log-likelihood was used primarily to verify statistical relevance, while Mutual Information (MI) was used for additional validation.

The third step was to find patterns for the verb *explain* compared with *clarify*. This was mainly a repetition of the second step but for the verb *explain* regardless of tense. Before looking for new patterns however, those

that existed for *clarify* were examined if also relevant for *explain*.

## 4. Results and patterns

Following the steps outlined in the methodology, results are presented and discussed. In this section, 'popularity' is used to represent the relative frequency per million words and 'sample' refers to the 100 lines retrieved for each verb (see appendix 1 and 2). References to individual results in the samples are put in square brackets, with the result number prefixed by a C or E referring to *clarify* and *explain* respectively (ex: [C1], [E99]).

### 4.1. General usage of the verbs

The Google Books (British English) corpus shows that the usage of *explain* was roughly ten times more popular than *clarify* around the year 2000. Both words have had their relative usage increase in the last century: *explain* by over 50%, *clarify* by over 1000%. The change in popularity of the words over time does not bias the further results considering the BNC's focus on the late twentieth century (BNC 2007) where the ratios were more stable.

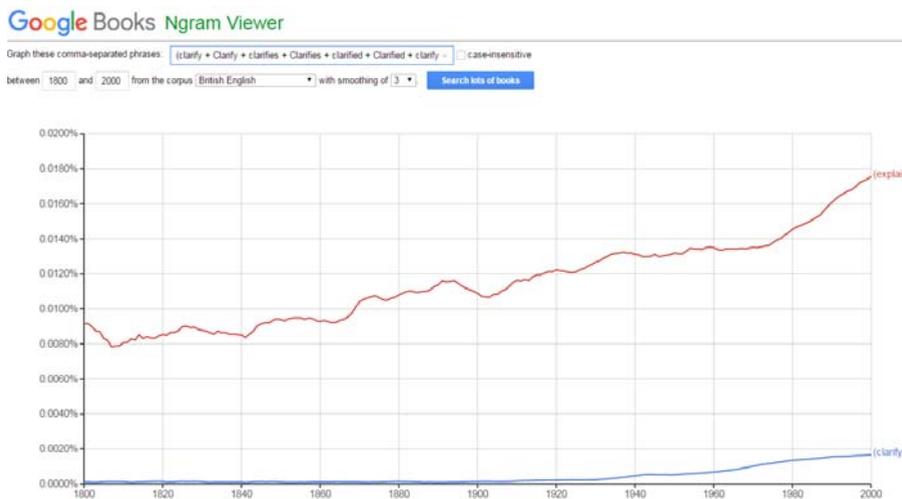


Figure 1: Graph of relative usage of *explain* (red) and *clarify* (blue)

Statistics from BNC's distribution function show more detailed information. Overall, the verb *explain* (in all forms) occurs 18651 times (relative frequency: 189,71 times per million words), while *clarify* occurs only 1435 times (14,6 times per million words), 13 times less.

Of the distribution across various types of written texts (Table 1), both

words are most popular in academic prose. The distribution between fiction and verse shows a contrast, this has the second highest popularity for *explain*, having nearly the same rate as for academic prose, while this is the second least likely place to find the verb *clarify*, over seven times less than academic prose. The reason for this contrast becomes clearer from further searches into the corpus and is visible in results presented later.

Table 1: Distribution by type of written text (relative frequency per million words)

Type of text	{clarify/V}	{explain/V}
Academic prose	28.96	280.2
Fiction and verse	3.72	224.23
Newspapers	8.29	142.37
Non-academic prose and biography	16.38	188.43
Other published written material	12.11	162.24
Other spoken material	26.07	148.32
Spoken conversation	1.42	48.65
Unpublished written material	13.43	153.15

A gender bias can be observed in the usage of *clarify* while not so with *explain* (

Table 2). Reasons for this were not immediately apparent nor further explored. The relative frequency of the two verbs per authors' age groups (

Table 3) reinforces that *explain* is a common word and thus used and known at a young age, while *clarify* is rarer.

Table 2: Distribution by gender of author of written text (relative frequency per million words)

Gender	{clarify/V}	{explain/V}
Male	28.96	280.2
Female	3.72	224.23
Mixed	8.29	142.37

Table 3: Distribution by age-group of author of written text (relative frequency per million words)

Age group	{clarify/V}	{explain/V}
0 - 14	0	167.9
15 - 24	3.69	224.85
25 - 34	10.59	213.5
35 - 44	10.59	220.61
45 - 59	10.1	210.91
60+	9.17	190.78

The popularity of the verbs between written and spoken texts (Table 4) shows different tendencies: *clarify* has a similar popularity, while *explain* is nearly twice as popular in the written than in the spoken form. A similar gender bias (Table 5) as for written texts is visible for the spoken texts for *clarify*, while an unexpected gender bias is found with *explain*, where it is used nearly twice as often. The age distribution shows a tendency that middle age groups use these verbs the most (Table 6).

Table 4: Distribution by type of text (relative frequency per million words)

Type of text	{clarify/V}	{explain/V}
Written	14.42	199.41
Spoken	16.04	107.78
<b>Overall</b>	<b>14.6</b>	<b>189.71</b>

Table 5: Distribution by speaker in spoken texts (relative frequency per million words)

Gender	{clarify/V}	{explain/V}
Male	14.95	126.47
Female	10.94	71.11

Table 6: Distribution by age-group of the speaker in spoken texts (relative frequency per million words)

Age group	{clarify/V}	{explain/V}
0 - 14	0	33.75
15 - 24	0	65.61
25 - 34	11.6	102.63
35 - 44	20.45	80.87
45 - 59	16.48	167.24
60+	0.88	28.13

The general information collected from the corpus does not answer the questions about usage, although it raises some questions. The first is why would 'explain', compared to 'clarify', be so much more popular in written fiction - the later analysis of the corpus provides a possible reason. The

second is the gender bias for which no attempt is made to find an answer in this paper.

#### 4.2 Using the corpus in explaining *clarify's* patterns

The first pattern that appears in the 100-line sample of '{clarify/V}' is how 'help' frequently appears to the left. Table 7 lists other verbs within a close distance of *clarify* in the sample,<sup>1</sup> and Table 8<sup>2</sup> lists collocates from the full corpus. Although the verb 'be' in its various forms is the most frequent collocation, its use as an auxiliary verb does not make for a special pattern, shown by the fact that 'help' is the strongest collocation based on statistical ranking in the corpus.

Trying to understand what can/cannot be clarified based on this leads to the idea that 'help' could be indicating a lack of confidence that the object will be completely understood after the clarifications. Several other verbs that are likely collocates ('seek', 'try', 'aim', 'attempt') fit that role. It could be that the object being clarified is either complicated or vague in such a way that even after the given clarification it may not be fully understood by the listener. A second pattern that appears, regardless of tense is that *clarify* is frequently (30 times) followed by the article 'the' immediately to the right hinting to a frequent usage of the verb in an active form with a simple term - article, adjective(s), noun - immediately following it which introduces the object under clarification.

Table 7: *Clarify*: collocated verbs from sample ordered by collocate frequency

Lemma (verb)	Observed collocate frequency
be	17
help	11
may	6
have	6
can	5
will	5
need	4
must	3
should	2

<sup>1</sup> Used the saved search results from '{clarify/V}' thinned to 100 results; used the collocations tool on this limited set to pull up collocations on POS-tags, range 3 left - 1 left, minimum frequency (collocate, node & collocate) were set to 2. The class filter was set to filter for verbs.

<sup>2</sup> Used search query '{clarify/V}'; used the collocations tool on this full set to pull up collocations on POS-tags, range 3 left - 1 left, minimum frequency (collocate, node & collocate) were set to 2. The class filter was set to filter for verbs.

Table 8: *Clarify*: collocated verbs from corpus ordered by log-likelihood

Lemma	Observed collocate frequency	Log-likelihood value
help	76	452.1587
can	72	167.3514
should	44	132.1665
will	69	128.5352
be	297	117.6842
need	32	108.3004
seek	17	80.7928
try	23	72.1474
may	31	71.9683
design	10	44.106
aim	8	38.3964
attempt	8	37.5633
have	94	35.1077
could	25	34.6621
like	12	29.3362

A count of POS tags (from BNC) immediately following the verb<sup>3</sup> (Table 9) confirms this is in fact the most popular form.

Table 9: *Clarify*: colligates from sample ordered by collocate frequency

Lemma (verb)	Observed collocate frequency
AT0	33
DT0	12
NN2	8
PRP	7
CJC	6
DTQ	5
PUN	5
DPS	3
AVQ	3
CJT	2
AV0	2
CJS	2
NN1-VVG	2

General determiners (DT0) fit the simple term pattern being clarified apart from the few cases when it is used alone (ex: "... to *clarify this*." [C44])

<sup>3</sup> Used the saved search results from {clarify/V} thinned to 100 results; used the collocations tool on this limited set to pull up collocations on POS-tags range 1 right - 1 right, minimum frequency (collocate, node & collocate) were set to 2.

where it refers to preceding clauses with the element to be clarified. This does not break the pattern if we consider punctuation or a conjunction encountered before a noun would indicate the end of the clause.

Likewise, nouns (NN2 and NN1) also fit the simple terms pattern if we consider articles and determiners as optional. Together, these three types of words immediately following the verb match more than half of the lines of the sample (57).

Seeking to identify the topics as simple terms that are typically clarified, Table 10 lists the nouns that repeat themselves in the sample at least twice and already certain meanings could be assumed (shown as grouping).

Table 10: *Clarify*: collocated nouns from sample ordered by collocate frequency

Lemma (SUBST)	Observed collocate frequency	Grouping
issue	6	Situation
problem	5	Situation
position	4	Situation/mental
role	3	Purpose
thing	2	Other
nature	2	Purpose
objective	2	Purpose
aim	2	Purpose

Turning to the full corpus with this theory of simple terms being frequent and that the more popular nouns in those terms could potentially be grouped, a search was made for nouns following the verb *clarify* with optionally POS previously identified in between that apply to the simple terms<sup>4</sup>. This search returned 824 hits representing ~57% of the occurrences of the verb.

As having the noun tag as part of the query precludes the use of the collocations tooling, and lacking control over the greediness modifiers in BNC web<sup>5</sup> a simplified query<sup>6</sup> was used resulting in 804 results on which the collocations tool was applied. Again, grouping was attempted for the top collocates what led to ‘situation’ and ‘purpose’ fitting to most of the statistically strongest collocations. The results in Table 11 give a first

<sup>4</sup> The query used was {clarify/V} (\_{ADV} | \_{ART} | \_{ADJ} | \_DT0 | \_DPS)\* \_{N}

<sup>5</sup> Based on possibilities described in The CQP Query Language Tutorial (Evert: 2009) a query [(lemma="clarify.\*") & (pos="V.\*")] [(pos="A.\*")+; set MatchingStrategy longest; was attempted with BNCweb (complex query), however this leads to a parsing error. No error is encountered if set MatchingStrategy is removed. No search modifiers are mentioned in the syntax for simple searches to control greediness.

<sup>6</sup> The query used was {clarify/V} (\_{ADV} | \_{ART} | \_{ADJ} | \_DT0 | \_DPS)

impression as to what topics can be ‘clarified’.

Table 11: *Clarify*: collocated nouns from corpus ordered by log-likelihood

Lemma (SUBST)	Observed collocate frequency	Log-likelihood value	Grouping
position	59	440.6632	Situation/mental
point	49	290.23	Purpose
issue	39	257.6768	Situation
situation	26	166.8745	Situation
role	19	108.0521	Purpose
meaning	14	97.6994	Purpose
thinking	11	85.5814	Purpose
nature	15	82.6285	Situation
law	17	79.2765	Other
understanding	11	73.3106	Purpose
distinction	9	63.7125	Situation
confusion	8	62.7296	Situation/mental
problem	17	60.6905	Situation
relationship	12	59.98	Situation
policy	13	51.5144	Situation

For verification, the top noun collocates immediately following *clarify* were also retrieved and presented in Table 12.<sup>7</sup> The only new (interesting) noun that appears in that list is ‘butter’ (as the most frequent and most likely collocate) which fits a secondary definition of making a liquid clear.

<sup>7</sup> Query used: {clarify/V}; used the collocations tool on the full set of results to pull up collocations on lemma with the range 1 right - 1 right, minimum frequency (collocate, node & collocate) were set to 3. Lemma class was filtered for SUBST.

Table 12: *Clarify*: collocated nouns in R1 from corpus ordered by log-likelihood

Lemma (SUBST)	Observed collocate frequency	Log-likelihood value	Grouping
butter	13	135.7011	Other
issue	10	48.4678	Situation
responsibility	4	18.4942	Purpose
matter	5	17.8759	Situation
thing	7	15.9343	Other
objective	3	15.1082	Purpose
aim	3	14.932	Purpose
point	5	12.3706	Purpose
problem	4	7.2989	Situation

### 4.3 Using the corpus to help in clarifying explain

In its sample, *explain* does not show the same trends as *clarify*. First, no special pattern appears based on verbs preceding *explain* (Table 13<sup>8</sup>) - the only preceding lexical verb repeating itself is ‘try’ but appearing only twice. Second, although *explain* is sometimes followed immediately by ‘the’, it is less frequent than for *clarify* (10 times vs 30 times).

Table 13: *Explain*: collocated verbs from sample ordered by collocate frequency

Lemma (verb)	Observed collocate frequency
be	24
have	8
can	6
could	3
try	2
may	2

A pattern that does come up frequently is ‘reported speech’ - 23 of the sample’s lines contain quoted speech (Appendix 3). Elements from the sample identifying quoted speech are the presence of quotes to the left or right within a near distance, a colon immediately following the verb, a proper name immediately before or after.<sup>9</sup> A query for such lines returns

<sup>8</sup> Used the saved search results from ‘{explain/V}’ thinned to 100 results; used the collocations tool on this limited set to pull up collocations on POS-tags range 3 left - 1 left, minimum frequency (collocate, node & collocate) were set to 2. Class filter was set to verb.

<sup>9</sup> Query used: ( {explain/V} \: | {explain/V} \_NP0 | \_NP0 {explain/V} | \_PNP {explain/V} | \_PUQ \*\*\*\* {explain/V} | {explain/V} \*\*\*\* \_PUQ )

7935 hits. If *explain* is used for quoting speech frequently, this can explain the difference in the utilization in novels (fiction and verse) between *explain* and *clarify*.

Using a table of colligate (POS items) immediately following the verb shows different patterns than for *clarify*. Table 14 shows the top 5 colligates for *clarify* and for *explain* as comparison.

Table 14: *Explain* and *clarify* - top 5 colligates compared

Tag	<i>explain</i>		<i>clarify</i>	
	Rank	Collocate frequency	Rank	Collocate frequency
PUN	1	22	7	5
PRP	2	19	4	7
AVQ	3	11	9	3
AT0	4	10	1	33
CJT	5	9	10	2
AT0	4	10	1	33
DT0	7	4	2	12
NN2	N/A	0	3	8
PRP	2	19	4	7
CJC	12	2	5	6

Looking at the colligates for *explain* from the sample, the most frequent one - punctuation - mostly comes from lines with quoted text. The prepositions 'by' and 'to' are often used in sentences with a passive voice while other prepositions (in, within, since, as, on) even if used in a sentence with an active voice, make it such that the object is further off.

The Wh-adverbs (AVQ) - why and how in the sample - and the conjunction 'that' act as the object to be explained, with a fuller and more complex description generally following to the right. 'The' introduces relatively simple objects very close to the verb still on the right.

From the sample, it could be concluded that more complex topics are frequently the object of *explain* introduced by 'that', 'how' and 'why'.

Using the top colligates (Articles - AT0, Wh-adverbs AVQ, subordinating conjunction *that* - CJT, prepositions - PRP) immediately following *explain*<sup>10</sup> in the sample, combined with the analysis above as to which side of the verb would the object of *explain* be found, leads to no new major patterns in terms of collocates (tables with the top 15 collocates per POS tags are available in

<sup>10</sup> Query used: {explain/V}; used the collocations tool on the full set of results to pull up collocations on POS-tags with the range 1 right - 1 right; minimum frequency (collocate,node & collocate) were set to 1. Opened up list of lines matching specific POS tags, and used the collocate tool on those subsets of the results.

appendix #4): based on log-likelihood and mutual-information statistical rankings, no terms eclipse or come close to those of ‘that’, ‘how’ and ‘why’ in such conditions.

Finally, the top general collocates sorted by log-likelihood (Table 15) show the same results as above, with the addition of ‘what’ following the verb used to refer to a more complex clause as the object being described.

Table 15: *Explain*: general collocations from corpus ordered by log-likelihood

Lemma	Observed collocate frequency	Log-likelihood value
why	1652	10181.7484
how	1041	4019.2333
that	1910	2668.2146
the	6047	1428.2652
what	746	1253.0517
term	330	1142.3733
by	922	808.7184
:	569	657.3993
behavior	113	399.636
phenomenon	72	372.3533
'	917	349.8931
everything	117	348.8142
to	2326	335.8554
difference	117	335.2195
he	1207	276.3146

## 5. Conclusion

This paper explored the usage of two synonymous verbs through their usage in sampled texts. This was possible due to the BNCweb tooling. While some features would have been nice to have – such as control over the matching strategy/greediness or position searches combined with grammatical tags, which are available with other tools, these are not considered likely to have impacted the results significantly. This paper does not present an exhaustive list of patterns and identifies certain differences (the distribution between genders) which could require further investigation in future work. Certain possible patterns were identified and demonstrated to be applicable to numerous texts.

Although dictionaries state that the function of the chosen verbs is to make something easier to understand, they have other uses: *explain* is used frequently to report speech and *clarify* is used in some rare cases in the sense of making a liquid clear. When they are used in a context of understanding,

*explain* is used frequently with complex clauses as object, introduced by key words such as 'why' and 'how', 'that' and 'what'; *clarify* is used frequently with simple clauses as object ('position', 'point'), but those terms refer to abstract elements such as general situations, purpose or elements which are otherwise mental in nature.

Sinclair's process provided a good foundation for a general analysis discovering patterns, but stopping too soon in looking at sample lines could have missed rarer meanings which were revealed using further searches on and general collocations found within the corpus. Using corpora in the classroom would require advanced research techniques and extensive work on finding possible differences but would provide a potential of teaching preconstructed phrases and explore with learners the words in their context setting.

## References

- Ädel, Annelie (2010). *Using Corpora to Teach Academic Writing: Challenges for the Direct Approach. Corpus Based Approaches to ELT*. London, UK: Continuum.
- Andersen, Gisle (2010). How to use corpus linguistics in sociolinguistics. O'Keeffe, Anne, Michael McCarthy, eds. (2010) *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, 547–562.
- Bell, Allan (1991). *The Language of News Media*. Oxford: Basil Blackwell.
- Biber, Douglas, Susan Conrad, Randi Reppen (2000). *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- Biber, Douglas, Edward Finnegan (2014) [1991]. On the exploitation of computerized corpora in variation studies. Aijmer, Karin, Bengt Altenberg, eds. (2014) *English Corpuses Linguistics*. London: Longman, 204–220.
- Boulton, Alex, Tom Cobb (2017). Corpus use in language learning: A meta-analysis. *Language Learning*, 67(2): 348–393
- Cobb, Tom, Alex Boulton (2015). Classroom applications of corpus analysis. Biber, Douglas, Randi Reppen, eds. *Cambridge Handbook of English Corpus Linguistics*. Cambridge: Cambridge University Press, 478–497.
- Corino, Elisa, Cristina Onesti (2019). Data-driven learning: a scaffolding methodology for CLIL and LSP teaching and learning. *Frontiers in Education* 4: 7.
- Cruse David Alan (1986). *Lexical Semantics*. Cambridge: Cambridge University Press.
- Divjak, Dagmar (2008). Ways of intending: Delineating and structuring near-synonyms. Gries, Stefan T., Anatol Stefanowitsch, eds. *Corpora in Cognitive Linguistics: Corpus-based Approaches to Syntax and Lexis*. Berlin: Mouton de Gruyter, 19–56.
- Divjak, Dagmar (2010). *Structuring the Lexicon*. Berlin, New York: De Gruyter Mouton.
- Emons, Rudolf (1997). Corpus linguistics: Some basic problems. *Studia Anglica Posnaniensia* 32: 61–68.
- Evert, Stefan (2009). *The CQP Query Language Tutorial*. (www) <http://cwb.sourceforge.net/temp/CQPTutorial.pdf> (15 June 2017)

- Flowerdew, Lynne (2011). *Corpora and Language Education*. Basingstoke: Palgrave Macmillan.
- Gilquin, Gaëtanelle, Sylviane Granger (2010). How can data-driven learning be used in language teaching. O’Keeffe, Anne, Michael McCarthy, eds. *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge, 359–370.
- Glynn, Dylan (2014). Techniques and tools. Corpus methods and statistics for semantics. Glynn, Dylan, Justyna Robinson, eds. *Corpus Methods for Semantics. Quantitative studies in polysemy and synonymy*. Amsterdam: John Benjamins, 1–6.
- Gries, Stefan (2015). Some Current Quantitative Problems in Corpus Linguistics and a Sketch of Some Solutions. *Language and Linguistics* 16(1): 93–117.
- Halliday, Michael A.K. (1966). Lexis as a linguistic level. Bazell, Charles E., John C. Catford, Michael A. K. Halliday, Robert H. Robins, eds. *In Memory of J. R. Firth*. London: Longman, 148–162.
- Hunston, Susan, Gill Francis (2000). *Pattern Grammar: A Corpus-Driven Approach to the Lexical Grammar of English*. Amsterdam: John Benjamins Publishing Company.
- Hunston, Susan (2002). *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Inkpen, Diana, Graeme Hirst (2006). Building and Using a Lexical Knowledge Base of Near-Synonym Differences. *Computational Linguistics* 32: 223–262
- McEnery, Tony, Andrew Hardie (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- McEnery, Tony, Richard Xiao, Yukio Tono (2006). *Corpus-Based Language Studies: An Advanced Resource Book*. London, New York: Routledge.
- Michel, Jean-Baptiste, Yuan Kui Shen, Aviva P. Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, et al. (2011) Quantitative analysis of culture using millions of digitized books. *Science* 331(6014): 176–182.
- O’Keeffe, Anne, Michael McCarthy (eds.) (2010). *The Routledge Handbook of Corpus Linguistics*. Abingdon: Routledge.
- Rayson, Paul, Dawn Elizabeth Archer, Alistair Baron, Jonathan Culpeper, Nicholas Smith (2007). Tagging the bard: Evaluating the accuracy of a modern POS Tagger on Early Modern English corpora. Corpus Linguistics Conference 2007 (CL2007), 27 July 2007 - 30 July 2007, University of Birmingham, UK.
- Sinclair, John (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, John (2003). *Reading Concordances: An Introduction*. London: Pearson Education.
- Sinclair, John (2004). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Sinclair, John (2005a). Corpus and text – Basic principles. Wynne, Martin, ed. *Developing Linguistic Corpora: A guide to Good Practice*. Oxford: Oxbow Books. 1–16. <http://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm> (17 March 2017)
- Sinclair, John (2005b). Appendix: How to build a corpus. Wynne, Martin, ed. *Developing Linguistic Corpora: A guide to Good Practice*. Oxford: Oxbow Books, 79–83. (10 April 2017) <http://ota.ox.ac.uk/documents/creating/dlc/appendix.htm>
- Stubbs, Michael (1995) Collocations and semantic profiles: On the cause of the trouble with quantitative studies. *Functions of Language* 2(1): 23–55. <https://www.>

- uni-trier.de/fileadmin/fb2/ANG/Linguistik/Stubbs/stubbs-1995-cause-trouble.pdf (4 April 2017)
- Stubbs, Michael (2001). Texts, corpora, and problems of interpretation: a response to Widdowson. *Applied Linguistics* 22(2): 149–172
- Stubbs, Michael (2009). The Search for units of meaning: Sinclair on empirical semantics. *Applied Linguistics* 30(1): 115–137.
- Taylor, John R. (2003). Near synonyms as co-extensive categories: ‘high’ and ‘tall’ revisited. *Language Sciences* 25(3): 263–284.
- The British National Corpus* (3<sup>rd</sup> version, BNC XML Edition) (2007). Distributed by Oxford University Computing Services on behalf of the BNC Consortium ([www](http://www.natcorp.ox.ac.uk/)) <http://www.natcorp.ox.ac.uk/> (April–July 2017)
- Widdowson, Henry G. (2000). The limitations of linguistics applied. *Applied Linguistics* 21(1): 3–25.

**Author’s address:**

Cvjetkova 32  
HR-31000 Osijek, Hrvatska  
e-mail: Maja.Seguin@evtos.hr

Received: February 18, 2020

Accepted for publication: July 1, 2020

## Appendix 1: 100-line sample for '{clarify/V}'

Table 16 - *clarify*, 100 random lines from corpus

1	A0J 1900	Occupational Therapist for the work. The Flow Chart may help to	clarify	the procedure. ' Making adaptations to your home may seem a
2	A10 1317	Concern believes that there is an urgent need for the Government to	clarify	the role of the health service in providing community care. Community
3	AAV 23	negotiate to obtain the best price. Their purchasing position will be	clarified	and the solicitors notified. Naturally the final decision is yours but
4	AJD 534	his audience. 'So I was delighted when last night he	clarified	the issue for the nation. As vicars say, I would
5	AMG 232	to a large extent an exercise in extrapolation. I attempt to	clarify	the roots of human cognition in their primordial states. An ethology
6	AMT 1193	what we mean by 'transcendence', so the first principle	clarifies	what we mean by 'immanence'. God is not to
7	AP1 316	invoice for the tape, but will cancel it once we have	clarified	the position. I look forward to receiving the details of your
8	B0H 1137	decided without troubling the whole Cabinet; and major issues to be	clarified	in order to save the time of the Cabinet. The method
9	B2D 746	1906 (as amended). Perhaps there is an argument for	clarifying	and formalising this co-operation. NON-LEGAL CONTROLS The Ombudsmen An avenue open
10	B2F 1000	routine to follow when setting an objective for an interaction: 1	clarify	the overall aim (ie the longer term end result).
11	B2M 986	were considered necessary to progress the study, ie one concerned with	clarifying	roles and information needs , and the second exploring and quantifying the
12	B7J 1925	, so that the causal influence of lead on IQ is not	clarified	in the experiment. However, in principle this would seem to
13	BLY 1441	on examination and selection in African schools has done so much to	clarify	problems and suggest solutions: 'If the Certificate of Primary Education
14	BME 301	and/or manager is a prerequisite for good communication. Instructions must be	clarified	and the instructor should probe when appropriate to ensure that his requirements

15	BMH 361	together still underpin modern legislation on obscenity. Lord Campbell, in	clarifying	the scope of the Act, stated that it applied ... exclusively
16	CAJ 747	which was last modified by the Court of Cassation in 1985 to	clarify	the Klaus Barbie case. The crime in question must be committed
17	CBT 3146	accept late partnership continuation elections in certain situations, which were recently	clarified	by Financial Secretary to the Treasury Stephen Dorrell. The election avoids
18	CBW 3339	£6.85, from HMSO. WESTMINSTER UPDATE Late filing of accounts dates	clarified	Following the introduction of civil penalties for late filing of accounts from
19	CBX 225	. In July, the Urgent Issues Task Force released a consensus	clarifying	the treatment of transfers from current assets to fixed assets. '
20	CDC 1401	of matters under discussion, and of asking questions which helped to	clarify	planning and decisions. I could not help feeling that if we
21	CGD 1345	Networks Project (Sinclair et al ., 1984) helped to	clarify	how informal_ADJ care networks function. It looked at formal and informal
22	CHT 392	in breach of proper professional conduct. This exercise should help to	clarify	the kind of theoretical and practical knowledge you need to update and
23	CJG 365	additional explanatory services and individual demonstration of processes may be needed to	clarify	work in mathematics and the sciences. For those pupils who possess
24	CMB 1616	's reforming zeal are as yet unknown. The White Paper will	clarify	many of these issues and will give a good idea as to
25	CMM 69	Before looking at the bloom shapes and forms, we need to	clarify	a few terms which are not as simple or straightforward as might
26	CN6 832	system, with its inconsistencies and confusion, conciliation services to help	clarify	arrangements in the best interests of children and encourage joint parental responsibilities
27	CR5 360	long established customs. A DoH spokesman said the new guidance will	clarify	where inspectors can play a legitimate role. He admitted drawing the
28	CRW 304	the phrase is to be infused with greater meaning, and to	clarify	whether what is meant is the best interests of all children in
29	CS2 22	action is right, if the latter it is wrong. To	clarify	this, Bentham lists seven so-called dimensions of pleasure and pain:

30	CSC 214	Software Inc, a noticeably popular fashion in start-ups this season.	clarify	, however, is positioning itself against the others as the most
31	EAJ 43	book. Oakeshott's objective in Experience and its Modes is to	clarify	the nature of philosophy. His method is to distinguish philosophy,
32	EEN 170	. But before looking in more detail at these developments we must	clarify	the focus of our argument. Robert Padgug has recently written that
33	EVS 1886	at the age of 34, in circumstances which will never be	clarified	. She had returned to El Salvador in January to collect evidence
34	EW6 401	. It is not possible in a modest book like this to	clarify	everything. For the moment just note that electron spin provides a
35	EX5 1506	automatically shares your viewpoint. To be persuasive, you need to	clarify	your position, make your assumptions explicit and your meaning as clear
36	F9T 2227	at some of the ideas associated with each approach may help to	clarify	further the differences between them (see Table 8.1). The
37	FA3 743	explored in Chapter 3. The model also goes some way to	clarify	the problem of the 'humanities'. It suggests that of
38	FA9 34	Understandable though this might be, it still leaves the problem of	clarifying	what is meant by 'social representation'. It will be
39	FA9 37	to undermine the revolutionary claims. Instead, it is hoped to	clarify	some central issues relating to the important, but difficult, concept
40	FA9 43	representations, in the particular sense of the term, will become	clarified	by investigating the son of socially shared beliefs which can be distinguished
41	FAF 1171	electorate. At the outset, however, it is important to	clarify	the distinction between marginal and average rates of tax. In public
42	FAM 214	i.e. national, local and school, and in an attempt to	clarify	them the following paragraphs examine each in turn. According to the
43	FRN 339	) remarks that these 1985 Guidelines have not been particularly successful in	clarifying	the issues, and vertical restraints remain an unsatisfactory area of US
44	FS6 964	often part of the actual disorder. Some examples might help to	clarify	this. If a child is born with some degree of mental

45	FSY 120	for $\beta$ -myosin heavy chain in some families with the condition has both	clarified	the nature of the genetic defect and simultaneously raised the prospect of
46	FTV 368	It is claimed that the meaning of existential assertions can be fully	clarified	only by clarifying the mode of being of the one who makes
47	FU3 862	between the different types of history books. E. CHECKLIST 2.	clarify	your understanding of the main differences between narrative, description and analysis
48	G0N 1284	10 Derek's previous experience of dealing with the police amounted to	clarifying	some technical points for the Fraud Squad when a client of Fithyan
49	G1F 1313	telling teachers how to teach? Or should policy concentrate more on	clarifying	the goals and outcomes of learning , and on providing the kind
50	G1H 388	Before discussing possible explanations of football hooliganism, it is necessary to	clarify	the different forms of behaviour incorporated within this term, where these
51	G3H 1502	put to the test. The Minister now has an opportunity to	clarify	that position. I assume that he has the information in front
52	GUJ 929	But most of these can be regarded as experimental works designed to	clarify	the problems that Picasso was facing in his painting. Picasso,
53	GVU 2118	, so much so that laws have to be passed 'to	clarify	' the original intention. Interpretations of statute are not unfamiliar to
54	GXG 237	Environmental Bureau, of which we are members, and aimed at	clarifying	our objectives for European Community action (a Directive?) on
55	GXG 1165	very useful if we could meet you to explore shared interests and	clarify	how best we can work together on this issue. In the
56	GXJ 1491	on the proposed arrangements is available for comment. Key Task 4	clarifying	the roles of general practitioners and primary health care teams Lead Role
57	H0Y 1022	authority" during lesson. 10 Use facial expressions and gestures to	clarify	meaning. Present new words in context. 11 Vary voice –
58	H7Y 1375	resemble the examples alongside the Irchester road, but here excavations have	clarified	aspects of their chronology and function. Work at the north end
59	H83 1501	sake. Lack of sufficient proportion, clarity or similarity means that	clarifying	and distinguishing the data is so fraught with excessive difficulty that interest

60	H88 1375	assess or evaluate mathematical attainment or progress in learning, we must	clarify	what we mean by 'doing mathematics'. Does it,
61	HAC 5896	Practical PC and I feel certain points about Windows Multimedia capabilities need	clarifying	. Although the soundblaster and Adlib cards are mentioned in the same
62	HB3 1482	on the way claims are handled. The following section attempts to	clarify	some of the more common areas where misunderstanding can occur. GENERAL
63	HC4 82	limited tolerances which are difficult to apply in practice; and to	clarify	which types of fish are covered by controls. A more thorough
64	HGY 3016	it nice that little Ellie was here just in time to help	clarify	things?' 'There's nothing to clarify. There was
65	HHV 18227	which will come before the House on Friday, but did not	clarify	whether he is in favour of a ban on fox hunting.
66	HHW 13153	policy framework and the arrangements that would apply to them. We	clarified	a number of points before the legislation was introduced. The Bill
67	HJ2 3443	; the latter, by analysing Chicken and Prisoners' Dilemma,	clarifies	the problem of voluntary co- operation to achieve valued goals. In showing
68	HJ2 4513	] The aim of this programme is to investigate, analyse and	clarify	the various linkages between government and industry that exists both in this
69	HKP 2944	in tariff bindings (ie undertakings not to increase tariffs) by	clarifying	the other duties and charges, in addition to ordinary customs duties
70	HL3 1262	Salvadorean government and the US administration itself, in its efforts to	clarify	the events. Cabinet change Oscar Santamaría left his post as Justice
71	HNW 2042	Primary teachers seem to have found the review of most value in	clarifying	aims and objectives (least, secondary), improving resources (
72	HPF 413	in turn below, but the objective of risk management should be	clarified	first: it should contribute to the business objectives and not be
73	HPU 165	discussion of some of the more directly relevant philosophical issues may help	clarify	matters. It is possible to distinguish, albeit rather grossly,
74	HRM 1970	the state bureaucracies which Weberians emphasise. The point can now be	clarified	by turning to another illustration, one based on a contemporary debate
75	HTV 1051	once again, the methodology of state-centrism serves to blur rather than	clarify	the issues. The national origin of the agents of media control

76	HU2 2686	effectively extracted from the cells. Further work is in progress to	clarify	this problem. Our results with western blotting contrast with those of
77	HU2 2879	with that in patients with established colorectal cancer, may help to	clarify	this issue. A prolonged colonic transit time may be expected to
78	HU4 4203	. The situation with respect to pressure changes and pain may be	clarified	by prolonged ambulatory measurements of colonic pressure activity, which allow the
79	HXT 1508	their populations. The purpose of this chapter is to try and	clarify	these issues by asking: (a) what is quality in
80	HXV 2372	aircraft whilst in flight. The position of aircraft has also been	clarified	at common law. The defendants took an aerial photograph of the
81	HY6 62	creative literature is the recognition of particular aspects of experience shaped and	clarified	, these authors are trying to share their understanding of the means
82	HY6 660	The end of the first half of The Form is concerned with	clarifying	the ordering of inner life – laying the fire – through an
83	HY6 1152	of the book, however, is occupied with the effort to	clarify	the process by which the reformation to the likeness of Jesus in
84	HYB 1366	possibilities of being clearer as to what we are talking about?	clarifying	what is meant by "spiritual" Table 6.1 suggests a way
85	J2K 23	is to have successfully communicated. A simple illustration may help to	clarify	the concept: it distinguishes between two kinds of "boos"
86	J44 135 Audio	no, that was, that was on the revenue Can I	clarify	that one? I think there's a bit of Right,
87	J6Y 859	avoid court proceedings. That the court has inherent discretionary jurisdiction to	clarify	the expert clause was stated in Royal Trust International Ltd v Nordbanken
88	J73 197	new rules to prevent such conduct in the future, or to	clarify	the situation. But it would have much more difficulty in applying
89	J9T 1037	[gap:name] raised last Thursday of Friday and which I thought I'd	clarified	Leeds position on. We we draw a clear distinction between provision
90	JAC 291 Audio	many days. We were talking about the new settlement. To	clarify	that we established yesterday that the new settlement will not necessarily be
91	JAD 1161	employment side. On Harrogate, erm I think the discussions have	clarified	a lot of the the sort of outstanding issues that that er

92	JK0 381 Audio	previous years accounts, yeah yes it is can we just erm	clarify	that Mr [gap:name] , of course it's dealing with the previous
93	JNN 146 Audio	remarks, it's hardly a question but perhaps you could just	clarify	Yes. my er thinking on on the [pause] on what you said
94	JSF 308	object was the same as Sir Thomas Bingham's, namely to	clarify	and strengthen the position of the auditor, not to change it
95	JSF 325	madam speaker, er Madam Deputy Speaker, if I seek to	clarify	er when an individual auditor is operating in his capacity as such
96	JSG 335	, what was needed was some amendment to the banking act to	clarify	that and that is precisely what [pause] er this order actually does and
97	JY7 1847	name.' It was, she knew, the moment to	clarify	things, to explain that Paolo was a company, not a
98	K2F 593	'Meetings of all the Palestinian institutions are starting in order to	clarify	and decide the prospects for the political action in the future,
99	KM6 23	just sort of clarify this before going on. Can I just	clarify	, what, what you're saying is that in a sense [pause]
100	KRL 521	? Well I think some of the big problems [pause] if we can	clarify	mechanical breakdown insurance. Now it is actually mechanical breakdown insurance.

## Appendix 2: 100-line sample for '{explain/V}'

Table 17 - *explain*, 100 random lines from corpus

1	A00 404	get home and has no money for transport. I try to	explain	that, unlike her home country, she will not be staying
2	A69 605	avoided. Helmut Schoeck, the professor of sociology at Mainz,	explains	the reasons why in a comprehensive analysis which restores the phenomenon of
3	ACK 1545	well. 'I'm used to Nigel now,' Elinor	explained	to me, 'but it's nice to have a change
4	AD2 988	is a large uncommitted population, this is probably not something which	explains	why more people do not join the Free Church. A second
5	ADR 1482	sex symbol was the Sydney fashion guru. Nicole Bonython. She	explained	: 'I was called in at the last minute to help
6	AEA 1221	adore.' 'I don't feel that you're really	explaining	yourself. I don't really understand. You love it here
7	AM4 181	a fear of stomach cancer. The first factor, the newspaper	explained	, was not important as no cases had occurred since 1972,
8	AM6 481	have in common is the requirement of 'aesthetic intention',	explains	: For – cutting straight through the recurring controversies about whether,
9	ASF 1030	either born or entrusted with his divine mission or died, are	explained	by al-Biruni (AD 973– c. 1050) in his great work
10	B08 1870	prepared by the company or initialling a few sheets of paper which	explain	the agreed terms. Employers often insist that departing executives must sign
11	B20 1554	everything. Now would you rather answer my questions, or try	explaining	yourself to Marshal Davout at headquarters?' The girl caught her
12	B25 200	this general picture (the theory) suggests certain ideas which may	explain	why non-Christians wish to have their babies undergo a ceremony in a
13	B2G 1233	, namely, the teaching of sexual magic, and this teaching	explains	, without exception, all the secrets of Nature, all the
14	B2J 80	observed increase in rank up to the medium-volatile stage could be adequately	explained	as the thermal result of former depth of burial; the higher
15	B3J 2949	' Nigger flushed, 'I was,' he replied and	explained	to Terry the reason for his incarceration. 'How's your
16	B7F 58	industries die they are usually operating near to peak employment,'	explains	James Wilson, the corporation's chief executive. 'The new

17	BMG 1110	how can the rare cases that have kept the story going be	explained	? The answer seems to be that a female cat sometimes experiences
18	BPF 869	in the red? 'Attitudes to money are important',	explains	Sarah Litvinoff, author of The Relate Guide to Better Relationships,
19	C8P 753	Catch the bomb Ask the children to sit in a circle and	explain	that the aim of the game is to toss the beanbag (
20	CBF 12543	we can – as normal as possible. 'You can't	explain	it. You can't tell them there is any logic to
21	CBU 708	to the cost of raising capital from shareholders. As Mr Lawrence	explains	, 'it's tilted the balance in favour of borrowing rather
22	CCD 290	that she was destined for high places. She could not have	explained	it but she no longer even thought of herself as Joan Halidon
23	CDP 523	the proposed operating centre was unsuitable on environmental grounds. The Council	explained	that the operating centre had no relevant planning permission nor had any
24	CFJ 1286	not like the sound of that. 'I mean,'	explained	Anabelle, 'I can't do it all alone. I
25	CG6 131	, rather than reinforcement by adults. While simple imitation fails to	explain	the occurrence of such novel utterances, a number of researchers have
26	CGF 413	one told about the Carib is clearly not adequate: it over	explains	what is really a rather small difference. Nevertheless the problem remains
27	CGF 595	the linguistic correlates of their roles. Sociolinguistics is most successful in	explaining	sex differences when it looks very carefully at the conditions of particular
28	CH2 10516	of men here or something?' Rita Loumansky, 71,	explained	: 'There are always more women than men here.'
29	CH5 5541	, ever since Allan Lamb, who plays for England, bravely	explained	to my colleagues what Wasim and Whacker – sorry, Waqar –
30	CHU 1159	oblige. Others, including the London councils, were sympathetic but	explained	that it would take months, not hours, to get the
31	CJX 2435	than you do. For reasons it would be too complicated to	explain	to you now, I believe that this may not be a
32	CLE 292	developed strategies comprising both labour market and political components how can we	explain	within the European context the greater emphasis upon the achievement of radical

33	CM5 1088	; and the remark made when I reported how I had initially	explained	my research aims to the union stewards – ‘You told the
34	CMH 710	because the ethologists are, eventually, going to have to start	explaining	why particular parts of the brain are involved in behaviours like sex
35	CN9 388	leave no doubt that the scale of these changes cannot be	explained	by any combination of relative adjustments within the context of post-war international
36	CRB 411	of them unsavoury, about his homosexual life in Cuba. Arenas	explained	that, although he was suffering from AIDS when he wrote the
37	CRS 1014	beginning to interview people in depth, I realised that, to	explain	this juxtaposition of apparently contradictory values within individuals, I needed to
38	CS3 1172	the absence of a strong socialist movement can to some degree be	explained	by (a) the ‘newness’ and apparent ‘classlessness
39	CTX 2013	in faxes, for example. The manual is well written and	explains	how the device works, but it doesn't explain how to
40	EA2 325	of 10314 civil servants between 1985 and 1988 with the aim of	explaining	socioeconomic gradients in morbidity and mortality. In addition to the established
41	EB2 195	a liberal theory of contract. Even if this flaw can be	explained	away, however, a further examination of the concept of protected
42	EDH 317	of the whole. According to the functionalist, then, to	explain	the existence of a social institution we must not look for purposive
43	EDU 274	of disciplines. So we have an operatic phenomenon that can be	explained	in good, old-fashioned aesthetic terms. At a time when the
44	EEC 585	the Criminal Justice Bill without any provision abolishing capital punishment and to	explain	in the Second Reading Debate that since the question of capital punishment
45	EER 387	the party a great deal of anxiety over my views’ he	explained	and added ‘but strangely enough with this book it has not
46	EF8 536	the type of phenomenon or the way in which it is being	explained	. In this chapter, we shall describe the phenomenon being explained
47	EFX 927	account of his expenses in a pocket notebook. One friend has	explained	how, in a conversation in which she was lamenting the cost

48	EW7 159	explication is sometimes a barrier to understanding. When the concept is	explained	by some writers, educational technology: ... can best be viewed
49	EWR 506	its aims and the most transcendental in its language. It also	explains	its oscillation between a concern with shrines and pilgrimages, with sanctification
50	F9W 477	the kind discussed in 2. Clearly, if we are to	explain	such interpretation we will need more than our list of cohesive devices
51	FP4 1340	characteristics. The residual, that part of the earnings gap not	explained	by personal attributes, is conventionally taken as an indicator of discrimination
52	FRL 1934	preference for negative additive conjunctions such as nor. Smith and Frawley	explain	this feature by suggesting that 'the high percentage of negative additive
53	FS0 1431	my decision as I understood his. I didn't have to	explain	. We talked a little more, while Terry and Tom prepared
54	FSA 16	by the empirical evidence, and that the earlier underpricing may be	explained	by the newness of the market and the inexperience of the participants
55	FSC 1054	Julia denied feeling the least bit tired and she pressed him to	explain	what he meant. 'It is difficult to justify when you
56	FSE 285	existed between them and other humans a coldness which could not be	explained	simply by xenophobia. 'You were connected with the original survey
57	G01 598	catwalk to the other. 'To prevent suicides,' Nicholson	explained	, standing beside her. 'Are there many attempts at suicide
58	G0H 639	Darwin were inspired by Lyell to look for geological changes that would	explain	the present distribution of species in Europe. But Forbes was unable
59	G0H 980	national responses suggest that we should be wary of any attempt to	explain	the popularity of evolutionism in terms of the self-evident superiority of Darwin
60	G0X 3093	Finn. 'That awful girl you took to Greece,'	explained	Lydia. 'Oh, she's around,' said Finn
61	G1R 413	by the Comintern. This neglect of the indigenous peasantry helps to	explain	the appeal of Maoism in Latin America and the over-estimation of the
62	GUF 3582	be a strange coupling, but Frances wouldn't refuse. He	explained	to Jacqui about the German measles. 'Oh no, for
63	GV8 3218	feeling like — maybe she'd give a chap a chance to	explain	that there's already a woman in his flat, a woman

64	GV8 3715	Innocent of what? She hadn't really given him time to	explain	, just looked at that horrible creature prancing about in the hall
65	GVK 335	Brezhnev's plan and they chose their arguments accordingly. One commentator	explained	that this plan would enable the Persian Gulf states 'to choose
66	GWM 436	this good-natured, reputedly moderate drinker merited such treatment has never been	explained	. Curtis Price asks: 'would Purcell have poked such cruel
67	GXJ 1194	members of staff from the Department of Highways will be present to	explain	the new road proposals in the Westburn/Baberton area. The Council would
68	H45 194	expressive nature. He would play us a popular classical piece,	explaining	what the compose was trying to get across to the listener.
69	H86 2618	small cigar. 'I'm not being inhospitable,' he	explained	, 'but we have to assume the chance of them having
70	H9A 733	a short run reputation effect. The answer to this question also	explains	why the bound derived by Fudenberg and Levine continues to hold.
71	HA4 1793	unromantic cold. 'We were worried about you,' Ianthe	explained	, sitting down on the edge of the narrow bed. '
72	HGD 756	. He flushed uneasily. 'What I mean is — she	explained	your circumstances quite openly and asked that if anything happened to her
73	HGR 2474	lexicon. However it is unlikely that a set of rules to	explain	how general language is processed will be available in the foreseeable future
74	HH2 805	resources. The object of the analysis is then to identify and	explain	in what circumstances the authoritarian direction of resources has advantage over market-type
75	HH3 1277	. The violent eruption which killed 1,700 people has never been satisfactorily	explained	. The official version, blaming the disaster on natural causes,
76	HH4 1404	. 'I've just remembered ... just realised,' she	explained	. 'So you really are Caro's stepbrother.' '
77	HHF 200	than filled, the more usual and cheaper method. A farmer	explained	that the area is subjected to floods sometimes, and it was
78	HHV 17276	. We will fight them. Mr. Geraint Howells Can the Minister	explain	why 10,000 full-time farmers left the land last year? There is
79	HJ3 5625	by a load of Job's Comforters telling you their troubles and	explaining	how hard and harsh life can be. Seek out the company

80	HLY 347	er, the new European directive almost ruined his holiday [pause] he then	explained	[pause] he booked up to take his family to EuroDisney. On his
81	HPB 114	the day. 'We did a lot of leafleting,'	explained	Alex. 'I was outside the hall leafleting when the debate
82	HS7 159	And so the catechist who talks but does not relate, who	explains	but does not feel is faced with a handicapped person like the
83	HSA 684	repeated between sips. The custom of cleaning the close had been	explained	to Madge on the day she moved in by a small woman
84	HTH 1190	surprise. They don't know we're here,' she	explained	. He gave her a sidelong glance. 'You've been
85	HU2 5230	metabolic effect of the tumours, how can such a phenomenon be	explained	? Upon consideration of the hypothesis that bile constituents and their intracolonic
86	HWF 6867	Interface for LIFESPAN. This manual has three main objectives: To	explain	in detail how a user begins work on LIFESPAN. To explain
87	HX4 551	, and franchiser companies to grow more rapidly, as Burger King	explain	. Burger King, the quality fast food hamburger chain, is
88	HX9 112	and bottom-up information given the experimental results. We are able to	explain	why certain search strategies (such as island-driving) are fundamentally misguided
89	JTE 322	Germany. And these'll be for the banquet and I'll	explain	the banquet later on as we go on. But this is
90	JY0 56	not – ' Juliet began. But she didn't have to	explain	. She'd apologised, hadn't she? She twisted out
91	K1D 1987	. Voice over For Flavio, the visit was a chance to	explain	how he's refused to allow blindness to get in the way
92	K55 1368	the town centre. Mr Milburn has contacted Durham County Council to	explain	the situation and push the case for the local people in the
93	K55 9399	RAF Leuchars in Scotland also raced to the scene. A spokesman	explained	: 'They are being called in because visibility is so bad
94	K5L 1683	per cent a year for the last five years. How to	explain	such results? Some attack the methodology of the American research,
95	K5P 411	the dissociation becomes slower at higher occupancy. This is difficult to	explain	since when a fully saturated complex is half dissociated it will have
96	K71 336 Audio	and the reason for me er tt coming today, as I	explained	on the telephone to you Mhm. is twofold really, one

97	K93 1326	last chapter the structure of the tone-unit was introduced and it was	explained	that when a tonic syllable is followed by a tail, that
98	K93 1696	clear that studying intonation in relation to discourse makes it possible to	explain	much more comprehensively the uses that speakers make of intonation. Practically
99	KLK 184	Erm Probably a little bit easier, certainly Mm. easier to	explain	. Okay let's have a look at cos X erm cos X
100	KPW 3	, you don't understand, you see I [unclear] Well,	explain	it then? You're so stupid! Go on, Charlotte

### Appendix 3: Quoted speech in 100-line sample for '{explain/V}'

Table 18: Quoted speech, *explain*

3	ACK 1545	well. 'I'm used to Nigel now,' Elinor	explained	to me, 'but it's nice to have a change
5	ADR 1482	sex symbol was the Sydney fashion guru. Nicole Bonython. She	explained	: 'I was called in at the last minute to help
6	AEA 1221	adore.' 'I don't feel that you're really	explaining	yourself. I don't really understand. You love it here
8	AM6 481	have in common is the requirement of 'aesthetic intention',	explains	: For – cutting straight through the recurring controversies about whether,
15	B3J 2949	' Nigger flushed, 'I was,' he replied and	explained	to Terry the reason for his incarceration. 'How's your
16	B7F 58	industries die they are usually operating near to peak employment,'	explains	James Wilson, the corporation's chief executive. 'The new
18	BPF 869	in the red?' 'Attitudes to money are important',	explains	Sarah Litvinoff, author of The Relate Guide to Better Relationships,
20	CBF 12543	we can – as normal as possible. 'You can't	explain	it. You can't tell them there is any logic to
21	CBU 708	to the cost of raising capital from shareholders. As Mr Lawrence	explains	, 'it's tilted the balance in favour of borrowing rather
24	CFJ 1286	not like the sound of that. 'I mean,'	explained	Anabelle, 'I can't do it all alone. I
28	CH2 10516	of men here or something?' Rita Loumansky, 71,	explained	: 'There are always more women than men here.'
45	EER 387	the party a great deal of anxiety over my views' he	explained	and added 'but strangely enough with this book it has not
57	G01 598	catwalk to the other. 'To prevent suicides,' Nicholson	explained	, standing beside her. 'Are there many attempts at suicide
60	G0X 3093	Finn. 'That awful girl you took to Greece,'	explained	Lydia. 'Oh, she's around,' said Finn
62	GUF 3582	be a strange coupling, but Frances wouldn't refuse. He	explained	to Jacqui about the German measles. 'Oh no, for
69	H86 2618	small cigar. 'I'm not being inhospitable,' he	explained	, 'but we have to assume the chance of them having
71	HA4 1793	unromantic cold. 'We were worried about you,' Ianthe	explained	, sitting down on the edge of the narrow bed. '

72	HGD 756	. He flushed uneasily. 'What I mean is — she	explained	your circumstances quite openly and asked that if anything happened to her
76	HHA 1404	. 'I've just remembered ... just realised,' she	explained	. 'So you really are Caro's stepbrother.' '
81	HPB 114	the day. 'We did a lot of leafleting,'	explained	Alex. 'I was outside the hall leafleting when the debate
84	HTH 1190	surprise. They don't know we're here,' she	explained	. He gave her a sidelong glance. 'You've been
90	JY0 56	not — ' Juliet began. But she didn't have to	explain	. She'd apologised, hadn't she? She twisted out
93	K55 9399	RAF Leuchars in Scotland also raced to the scene. A spokesman	explained	: 'They are being called in because visibility is so bad

## Appendix 4: Collocates close to *explain*, per colligate immediately following *explain*

The following tables represent the top 15 collocates which follow (range R1 – R4) or precede (L5 – L1) the verb. Although the range depends on where the topic being clarified is considered likely from analysis of the sample, in all cases the minimum frequency (both ‘node,collocate’ an ‘collocate’) are set to be at least 5. The collocates are based on the results found for the verb *explain* followed by the chosen colligate.

Table 19: Collocates ranked by log-likelihood following {explain/V} \_AT0

No	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Log-likelihood value
1	the_ART	6,040,293	634.514	3185	1271	5811.5261
2	of_PREP	3,040,508	319.396	1140	657	1320.2158
3	situation_SUBST	19,544	2.053	64	62	316.8933
4	nature_SUBST	17,880	1.878	56	47	272.4116
5	difference_SUBST	18,887	1.984	55	49	259.799
6	reason_SUBST	28,549	2.999	59	55	239.9279
7	phenomenon_SUBST	3,451	0.363	32	27	223.848
8	between_PREP	90,191	9.474	70	65	159.2891
9	origin_SUBST	4,697	0.493	25	23	147.431
10	behaviour_SUBST	12,676	1.332	31	26	135.9602
11	meaning_SUBST	7,904	0.830	27	19	135.8236
12	behind_PREP	19,056	2.002	31	31	111.9963
13	observed_ADJ	480	0.050	12	10	107.7464
14	procedure_SUBST	11,024	1.158	25	20	106.0235
15	apparent_ADJ	5,208	0.547	20	18	105.1545

Table 20: Collocates ranked by mutual information following {explain/V} \_AT0

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Mutual information value
1	flypaper_SUBST	14	0.001	5	1	11.7313
2	puzzling_ADJ	196	0.021	7	7	8.4093
3	observed_ADJ	480	0.050	12	10	7.8948
4	persistence_SUBST	460	0.048	8	8	7.3712
5	rationale_SUBST	583	0.061	9	9	7.1993
6	cross-section_SUBST	333	0.035	5	2	7.1592

7	discrepancy_SUBST	650	0.068	7	7	6.6797
8	phenomenon_SUBST	3,451	0.363	32	27	6.4639
9	paradox_SUBST	763	0.080	5	5	5.9631
10	emergence_SUBST	1,231	0.129	8	7	5.9511
11	reluctance_SUBST	1,013	0.106	6	6	5.8172
12	origin_SUBST	4,697	0.493	25	23	5.663
13	reasoning_SUBST	952	0.100	5	5	5.6438
14	occurrence_SUBST	1,343	0.141	7	6	5.6328
15	incidence_SUBST	1,761	0.185	8	8	5.4345

Table 21: Collocates ranked by log-likelihood following {explain/V} \_AVQ

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Log-likelihood value
1	why_ADV	48,563	4.058	1475	860	14739.5718
2	how_ADV	98,967	8.270	889	567	6648.4717
3	he_PRON	1,198,025	100.115	213	173	97.2402
4	should_VERB	108,970	9.106	48	44	81.9634
5	come_VERB	144,726	12.094	54	49	77.9841
6	it_PRON	1,213,293	101.391	196	172	70.1384
7	have_VERB	1,316,636	110.027	199	174	58.7649
8	so_ADV	214,145	17.895	59	57	58.7532
9	work_VERB	62,972	5.262	29	27	51.5835
10	be_VERB	4,119,764	344.274	473	382	44.9011
11	they_PRON	842,089	70.370	133	121	44.4976
12	choose_VERB	16,437	1.374	14	11	39.7797
13	can_VERB	261,805	21.878	57	53	39.0554
14	particular_ADJ	28,628	2.392	17	16	37.4869
15	woman_SUBST	60,149	5.026	24	20	37.1377

Table 22: collocates ranked by mutual information following {explain/V} \_AVQ

N o.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Mutual information value
1	why_ADV	48,563	4.058	1475	860	8.5056
2	how_ADV	98,967	8.270	889	567	6.7481
3	universe_SUBST	2,582	0.216	6	3	4.7974
4	lifespan_SUBST	3,729	0.312	7	1	4.4895
5	cope_VERB	5,029	0.420	5	4	3.5726
6	choose_VERB	16,437	1.374	14	11	3.3494

7	works_SUBST	7,958	0.665	6	6	3.1735
8	occur_VERB	15,430	1.289	10	9	2.9552
9	individual_SUBST	13,365	1.117	8	8	2.8405
10	particular_ADJ	28,628	2.392	17	16	2.829
11	operate_VERB	10,162	0.849	6	6	2.8208
12	object_SUBST	9,571	0.800	5	4	2.6442
13	manage_VERB	12,775	1.068	6	6	2.4906
14	work_VERB	62,972	5.262	29	27	2.4623
15	should_VERB	108,970	9.106	48	44	2.3981

Table 23: collocates ranked by log-likelihood following {explain/V} \_CJT

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Log-likelihood value
1	that_CONJ	740,455	34.743	1321	776	7388.1286
2	have_VERB	1,316,636	61.778	217	182	239.5101
3	he_PRON	1,198,025	56.213	188	160	193.7491
4	be_VERB	4,119,764	193.305	403	330	181.5635
5	she_PRON	654,356	30.703	79	74	53.1799
6	would_VERB	278,523	13.069	45	42	47.6156
7	although_CONJ	42,701	2.004	17	17	42.7565
8	this_ADJ	452,756	21.244	55	55	37.3471
9	not_ADV	767,448	36.010	71	67	26.6584
10	there_PRON	245,770	11.532	33	31	26.5455
11	they_PRON	842,089	39.512	72	63	21.6367
12	it_PRON	1,213,293	56.929	88	83	14.6979
13	need_VERB	64,349	3.019	11	11	12.4946
14	only_ADV	126,118	5.918	16	16	11.6843
15	no_ART	140,366	6.586	17	16	11.434

Table 24: Collocates ranked by mutual information following {explain/V} \_CJT

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Mutual information value
1	that_CONJ	740,455	34.743	1321	776	5.2488
2	although_CONJ	42,701	2.004	17	17	3.0849
3	meeting_SUBST	21,209	0.995	5	5	2.3289
4	reason_SUBST	28,549	1.340	6	6	2.1632
5	since_CONJ	28,653	1.344	5	5	1.8949
6	need_VERB	64,349	3.019	11	11	1.8652

7	never_ADV	53,143	2.494	9	8	1.8517
8	have_VERB	1,316,636	61.778	217	182	1.8125
9	would_VERB	278,523	13.069	45	42	1.7838
10	must_VERB	69,754	3.273	11	10	1.7488
11	he_PRON	1,198,025	56.213	188	160	1.7418
12	live_VERB	31,923	1.498	5	5	1.739
13	already_ADV	33,223	1.559	5	5	1.6814
14	there_PRON	245,770	11.532	33	31	1.5168
15	only_ADV	126,118	5.918	16	16	1.435

Table 25: Collocates ranked by log-likelihood preceding {explain/V} \_PRP

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Log-likelihood value
1	be_VERB	4,119,764	426.888	1428	723	1539.0401
2	can_VERB	261,805	27.128	336	239	1082.0428
3	could_VERB	159,818	16.560	144	122	369.5159
4	as_CONJ	378,061	39.175	199	154	329.4942
5	partly_ADV	5,574	0.578	42	35	277.6848
6	best_ADV	7,924	0.821	39	35	225.0862
7	try_VERB	52,838	5.475	64	59	198.0217
8	this_ADJ	452,756	46.914	162	131	172.5308
9	may_VERB	112,405	11.647	74	56	149.3148
10	i_PRON	1,128,916	116.978	269	204	146.0021
11	to_PREP	2,593,462	268.734	473	365	130.0277
12	fully_ADV	8,788	0.911	26	25	124.2375
13	not_ADV	767,448	79.523	185	147	102.4233
14	phenomenon_SUBST	3,451	0.358	17	12	98.1131
15	will_VERB	328,997	34.091	99	77	81.6462

Table 26: Collocates ranked by mutual information preceding {explain/V}  
\_PR

No.	Lemma	Total No. in whole BNC	Expected collocate frequency	Observed collocate frequency	In No. of texts	Mutual information value
1	satisfactorily_ADV	541	0.056	8	6	7.1569
2	patiently_ADV	497	0.051	5	5	6.6012
3	discrepancy_SUBST	650	0.067	6	5	6.4771
4	variance_SUBST	697	0.072	6	3	6.3764
5	gastrin_SUBST	600	0.062	5	4	6.3295
6	partly_ADV	5,574	0.578	42	35	6.1842
7	partially_ADV	1,284	0.133	7	7	5.7173
8	adequately_ADV	1,134	0.118	6	6	5.6742
9	phenomenon_SUBST	3,451	0.358	17	12	5.5711
10	best_ADV	7,924	0.821	39	35	5.5698
11	yield_SUBST	1,637	0.170	6	2	5.1445
12	readily_ADV	2,772	0.287	10	8	5.1216
13	fully_ADV	8,788	0.911	26	25	4.8355
14	easily_ADV	9,625	0.997	17	17	4.0913
15	curve_SUBST	3,288	0.341	5	1	3.8753