Wiliński, Jarosław

**Adjectives in extraposed constructions with that-clauses : a quantitative corpus-driven analysis**

Jarosław Wiliński

# Adjectives in Extraposed Constructions with *that*-clauses: A Quantitative Corpus-driven Analysis

**Abstract**

This paper employs the background assumptions of Construction Grammar (Goldberg 2006) and Frame Semantics (Fillmore 1982) as well as a quantitative corpus-driven method for investigating the reciprocal interaction between adjectives and the extraposed construction with that-clauses in American English. The method, referred to as the *attraction-reliance measure* (Schmid 2000; Schmid & Küchenhoff 2013), is applied to the determination of strongly attracted and repelled adjectives of the *it is ADJ that*-construction. On the basis of the data extracted from the academic sub-corpus of COCA, the paper indicates that some adjectives are more strongly attracted to this construction than others, and that the occurrence of certain adjectives in this construction is more significant than their use in different types of extraposed constructions. In addition, the findings seem to point to frame-semantic knowledge as a determining factor affecting the mutual association between adjectives and the construction under study.

**Keywords**

*Extraposition; attraction-reliance method; frame semantics; COCA corpus*

## 1. Introduction

The past thirty years have witnessed a striking proliferation of publications devoted to different aspects of extraposition in English (Huddleston 1984; Quirk *et al.* 1985; Seppänen, Engström and Seppänen 1990; Seppänen 1999; Biber *et al.* 1999; Kaltenböck 2000, 2003, 2004; Van Linden 2012). Some researchers

have focused on the comparison of extraposition and right dislocation (McCawley 1988; Collins 1994) as well as extraposed constructions and clefts (Pérez-Guerra 1998; Calude 2008), while others have explored the discourse function of extraposition (Gómez-González 1997; Herriman 2000a,b; Hoey 2000; Miller 2001; Hewings and Hewings 2002; Rowley-Jolivet and Carter-Thomas 2005; Kaltenböck 2005). Many research studies have also examined the use of deontic-evaluative adjectives in various types of extraposition constructions (Biber *et al.* 1999; Van Linden 2012), their clausal complementation (Mindt 2011), and different valency patterns (Herbst *et al.* 2004).

However, little attention has hitherto been paid to the quantification of adjectives in extraposed constructions, the statistical corroboration of their occurrence in academic discourse, or the empirical substantiation of previous hypotheses about their use. Hilpert's (2014) case study of adjectives occurring in the *it is ADJ to V*-construction and Wiliński's (2017) quantitative study of adjectives complemented by *to*-clauses in academic discourse are notable exceptions. Quantifying the data extracted from the BNC corpus, Hilpert (2014) found that there are groups of adjectives closely related semantically that are strongly attracted to the construction under investigation. These adjectives refer to different scales, such as importance (*important*, *essential*, *necessary*), possibility (*possible, impossible*), ease (*easy*, *difficult*, *hard*) and advisability (*advisable*, *better*, *best*, *wise*). On the basis of the data extracted from the academic sub-corpus of COCA, Wiliński (2017) in turn uncovered that extraposed constructions with *to*-clauses are used much more frequently than those with *that*-clauses in academic discourse (see Biber, Conrad and Reppen 1998: 74–75 for a similar observation), and that adjectives reflecting various semantic frames, such as IMPORTANCE (*important, critical, crucial, imperative, vital*), DIFFICULTY (*difficult, easy, hard*), LIKELIHOOD (*possible, impossible, likely, unlikely*), NECESSITY (*necessary, essential*), MENTAL PROPERTY ATTRIBUTION (*reasonable*, *wise*), and USEFULNESS EVALUATION (*useful, helpful, instructive*) are the most strongly attracted lexemes of the *it is ADJ to V*-construction.

Other studies that have considered the use of adjectives in extraposed constructions are restricted either in terms of their scope of interest or the data investigated (e.g. Herriman 2000a; Kaltenböck 2005; Kataari 2010). Kataari's (2010) study, for example, examined the occurrence of several epistemic, deontic, dynamic and evaluative adjectives in extraposed constructions and post-predicative ones followed by *to*-infinitives and *that*-clauses in a sub-corpus consisting of one million words derived from the BNC. Kataari (2010: 22, 26) established that adjectives expressing deontic modality (e.g. *essential, important, necessary, desirable*, or *useful*) in *it*-extraposition are much more frequently used with *to*-clauses, while those expressing epistemic modality (e.g. *clear, likely, true* or *unlikely*) are almost exclusively complemented by *that*-clauses (cf. Herriman 2000a: 592; Mair 1990: 25).

Kaltenböck's (2005) study, in turn, indicated that the matrix predicate of *it*-extraposition, which expresses some speaker's evaluation or stance, can carry

different meanings. Kaltenböck (2005: 138), however, only mentioned some semantic categories for adjectival matrix predicates governing extraposed subject clauses: existence of a state of affairs (e.g. *difficult, easy*), truth/transparency of a state of affairs (e.g. *clear, apparent*), likelihood/possibility (e.g. *likely, possible*) and value judgement (e.g. *good, nice*) for finite clauses; value judgement (e.g. *good, nice*), likelihood/possibility (e.g. *likely, possible*) and necessity/desirability (e.g. *necessary, desirable*) for infinitive clauses. These semantic categories are largely congruent with the ones determined by Collins (1994: 19) and Gómez-González (1997: 102–3, 2001: 272–73).

Since Hilpert's study and Wiliński's investigation were limited solely to adjectives occurring in extraposed constructions with *to*-infinitive clauses and others (e.g. Kaltenböck 2005; Kataari 2010) were not specifically designed to determine the most significant adjectives of the pattern complemented by *that*-clauses in scholarly prose, there is still a need for the quantitative identification of strongly attracted and repelled adjectives of the *it is ADJ that*-construction in American English and for the qualitative analysis of their usage in academic discourse, in view of the common occurrence of this construction in the academic register. Employing data retrieved from the academic section of the Corpus of Contemporary American English, the author thus attempts to identify both frequent and rare occurrences of certain adjectives in extraposed constructions with *that*-clauses, i.e. to determine those adjectives that are strongly and loosely associated with the pattern under investigation.

The remainder of this paper is organized as follows. Section 2 explains both the theoretical notions central to the semantic description of the adjectives occurring in the investigated construction, and the methodological assumptions underpinning the quantitative analysis of these adjectives. Section 3 describes the corpus, the data, and the tools, as well as the statistical procedure applied in this study. Section 4 defines the *it is ADJ that*-construction and discusses its function and usage. Section 5 integrates the findings of the quantitative analysis into a semantic description of the adjectives and explains how different semantic frames contribute to the constructional meaning. Section 6 provides concluding remarks and puts forward some proposals for future research.

## 2. Theoretical and methodological assumptions

This analysis adopts the terminology and the background assumptions of two theories, Construction Grammar (Goldberg 1995, 2006) and Frame Semantics (Fillmore 1982; Fillmore and Atkins 1992; Fillmore and Baker 2010). The former rests on the premise that grammar is a repertoire of simple and complex patterns, called constructions (cf. Goldberg 1995, 2006; Croft 2001), that are conventionalized pairings of form and meaning. Form in constructions can be applied to any combination of morphological, lexical, syntactic, or prosodic patterns, while meaning, in broad terms, would include lexical semantics, pragmatics, socio-cultural

aspects, and discourse structure. From this perspective, grammar consists of networks of overlapping and complementary constructions that serve as a basis for encoding and decoding all kinds of linguistic units. The latter, in turn, is founded on the assumption that concepts cannot be treated as being isolated, but are embedded in a vast body of knowledge. Hence, conceptual structures, referred to as semantic frames (Fillmore *et al.* 2003), provide a background of experience, beliefs, or practices essential for the understanding of the meaning of the word in question. In this view, the meaning of a word cannot be interpreted without access to all the background knowledge that is connected to that word. Thus, a semantic frame refers to a coherent structure of related concepts: without encyclopaedic information about all of them, we do not possess complete knowledge of any one of them. Each word is assumed to evoke a particular frame. For example, the words *John, to buy*, *a new car*, and *Brian*, in the sentence *John bought a new car from Brian*, should be understood with reference to the COMMERCIAL TRANSACTION frame (cf. Johnson *et al.* 2001) which encompasses background knowledge of a situation in which a buyer buys goods from a seller.

Frame Semantics laid the theoretical foundation for the FrameNet project, a lexical database of English in which the meanings of words are explained on the basis of semantic frames, descriptions of different types of events, relations, or entities and the participants in it. Each semantic frame also provides annotated examples that show the meaning and usage and combinatorial properties of a core set of the English vocabulary. Semantic frames (including their descriptions) and data described in this project have been employed by researchers around the world for a wide variety of purposes: for example, the database has been used as a dictionary of word senses and combinatorial properties as well as in applications such as information extraction, machine translation, event recognition, sentiment analysis, etc.

To date, however, the information provided in the FrameNet project has not been applied to any analysis of adjectives occurring in the *it is ADJ that*-construction. This study therefore is based on the following set of frames and their descriptions taken from the database: OBVIOUSNESS, LIKELIHOOD, IMPORTANCE, NECESSITY, CERTAINTY, STIMULUS FOCUS, COINCIDENCE, TRUST, LUCK, CANDIDNESS, and DESIRABILITY. The remaining frames (including their descriptions), applied in an analysis of semantic properties of adjectives, are created by the author himself: in other words, ADMITTING FACT, BEING NORMAL OR ACCEPTED, BEING UNAVOIDABLE, BEING CERTAINLY TRUE, and BEING POSSIBLY TRUE.

The method called the attraction-reliance measure (Schmid 2000; Schmid & Küchenhoff 2013) is used to capture, in quantitative terms, the mutual association between adjectives and the extraposed construction with that-clauses: in other words, to identify adjectives that are strongly attracted to or repelled from the it is ADJ that-construction. Attraction is expressed as the proportion in which a grammatical construction is filled by a particular lexeme, while reliance is the proportion in which a lexeme occurs only in a particular construction (cf. Schmid & Küchenhoff 2013: 548). In the case of the current study, the former proportion

can be understood as reflecting the Attraction exerted by the it is ADJ that-construction on the adjective, and the latter as reflecting the Reliance of the adjective on the construction under scrutiny. Attraction is computed by dividing the observed frequency of occurrence of an adjective in an extraposed construction by the total frequency of the construction in the corpus. Reliance, in turn, is calculated by dividing the frequency of occurrence of an adjective in an extraposed construction by the frequency of its occurrence in the whole corpus (cf. Schmid 2000: 54).

In this study, however, the second formula was slightly altered: *Reliance* was computed by dividing the frequency of occurrence of an adjective in the *it is ADJ that*-construction by its frequency of occurrence in other extraposed constructions, thereby excluding the frequency of the adjective in the whole corpus. This modified formula was employed to capture to some degree the intuition that some extraposed constructions, e.g. those complemented by *to*-infinitives, *for/to*-infinitives, or *wh*-clauses, can be more significant for certain adjectives than the construction under study. For example, Wiliński's (2017: 96) study, which was based on the same sub-corpus of COCA, revealed that the adjective *important* (1946 occurrences in the *it is ADJ to V*-construction) relies on the construction in question in a proportion of 71.62%, while *likely* (159 occurrences in the same pattern) in a proportion of 21.09%. This means that the first adjective occurs more frequently in the pattern complemented by *to*-clauses than in extraposed constructions with *that*-clauses (350 occurrences, see Table 1), *for/to*-infinitives, or *wh*-clauses, whereas the latter is used more commonly in the *it is ADJ that*-construction (537 occurrences) and less frequently in other patterns (58 occurrences), as the total frequency of the adjective *likely* in all extraposed constructions indicate (754 occurrences in the same academic sub-corpus).

The results of both measures were converted to percentages by multiplying the observed frequency of an adjective in the construction in each case by one hundred. These calculations were performed by means of Microsoft Excel spreadsheets. The percentages provided by both measures were taken as indicators of attraction and reliance: the larger the percentage, the stronger the attraction or reliance.

Strictly quantitative as this approach may appear, the results of the arithmetical calculations are evaluated qualitatively and subjectively. More specifically, adjectives with strong attraction to the *it is ADJ that*-construction are grouped according to the semantic frames they evoke on the basis of intuitive judgments.

## 3. Corpora, data, tools, and statistical procedure

The source of the data used in this investigation is the academic section of the Corpus of   Contemporary American English (COCA), covering the years between 1990 and 2012. More precisely, the data were retrieved from the downloadable version of COCA, i.e. the full-text data corpus purchased from Mark

Davies. This version contains more than 190,000 texts representing many different genres, and it is evenly divided in total size between spoken, fiction, popular magazines, newspapers, and academic journals.

The academic section includes approximately 81 million words derived from nearly 100 different peer-reviewed journals. These were selected to encompass the entire range of the Library of Congress Classification system: e.g. a certain percentage from philosophy, psychology, religion, world history, education, and technology.

MonoConc Pro, a concordance program, was applied to the retrieval of the observed frequencies from the corpus. Then, all the frequencies necessary to compute the strength of the reciprocal interaction between adjectives and the construction under investigation were entered into an Excel worksheet and subjected to Schmid's arithmetic measures of attraction and reliance. The percentages produced by these calculations were used as indexes of association strength, i.e., an adjective's strength of attraction to or repulsion from the *it is ADJ that*-construction: the higher the percentage, the stronger the attraction to and reliance on the investigated pattern.

The procedure adopted in this study involved three steps. The initial step was to calculate the observed frequencies. For the adjective *clear* in Table 1 (see Section 5)*,* for example, all occurrences of this adjective in the *it is ADJ that*-construction were first identified from the corpus, yielding 921. Then, the total frequency of the adjective (*clear*) in all extraposed constructions was determined, yielding 1069. Finally, the total frequency of the *it is ADJ that*-construction was worked out, giving 5391. These three figures were extracted from the corpus manually by reading concordance lines and counting all the occurrences of the adjectives under scrutiny.

The next step entailed computing measures of attraction and reliance. For this purpose, the frequencies mentioned above were entered into an Excel spreadsheet and examined by arithmetic tests. As Table 1 shows (see Section 5), the percentages obtained from the computation of attraction and reliance for the adjective *clear* are very high: 17.08% and 86.16%, respectively. This means that the adjective occurs in 17.08% of the uses of the construction in the corpus: in other words, *clear* is a highly significant, very strongly attracted lexeme of this construction. Furthermore, 86.16% of the occurrences of the same adjective are found in this construction, which means that *clear* only relies on other extraposed constructions in a proportion of 13.84%.

The final step was to arrange the adjectives according to their strength of attraction and to interpret the results qualitatively and subjectively. It was found that there are indeed adjectives that are significantly attracted to or repelled from this construction, and it was shown that the mutual attraction between a particular adjective and the construction is determined by frame-semantic knowledge.

## 4. *It*-extraposition with *that*-clauses

The term *it*-extraposition is applied to the syntactic process of moving (extraposing) an embedded clause (in this case, a *that*-clause) from its subject position to the right of the superordinate predicate while replacing it by the dummy pronoun *it*. Typical examples of subject *it*-extraposition with that-clauses are exemplified in (1), with the non-extraposed counterparts being provided in (2):

(1)   *it*-extraposition with *that*-clauses
    a.  It is clear *that John went to Warsaw*
    b.  It is surprising *that he didn't pass the exam*
    c.  It is obvious *that she committed this crime*

(2)   non-extraposition
    a.  *That John went to Warsaw* is clear
    b.  *That he didn't pass the exam* is surprising
    c.  *That she committed this crime* is obvious

Since there are obvious parallels between *it*-extraposition and non-extraposition in structural and logico-semantic terms (with both sharing structural similarities and expressing the same propositional meaning), many researchers (e.g. Rosenbaum 1967; Huddleston 1971; Emonds 1972) dealing with transformational-generative grammar would treat the sentences in (1) as syntactic derivations or transformations of the sentences in (2). Some corpus-based research (e.g. Francis 1993; Biber *et al.* 1999; Kaltenböck 2000), however, has revealed that it is highly doubtful whether sentences such as those in (1) are indeed transformations of the sentences in (2), as examples of non-extraposition are extremely rare in corpora. For example, the study conducted by Kaltenböck (2000: 158) indicates that *it*-extraposition substantially outnumbers its non-extraposed counterpart with a ratio of 1:7.8 in the British component of the International Corpus of English. Thus, it seems to be more reasonable to accept *It is clear that John went to Warsaw* as a construction (a pairing of form and meaning/function) in its own right, and to examine it accordingly, rather than to consider it as a version of something that hardly ever occurs.

    This study, therefore, treats the examples such as the ones in (1) as a type of the English *it*-extraposition construction, a partially lexically-filled pattern involving three fixed lexical items (it is […] that) and one flexible slot that can be filled by adjectives. This construction can be represented structurally and schematically as [*it is ADJ that-clause*], where an expletive subject *it* is followed by a copula, a predicative adjective, and a *that*-clause. The construction can be illustrated by the following sentences extracted from the corpus:

(3)   It is clear that the Greek government is still working on its position

(4)   It is essential that employers know what practical advice to give employees who have multiple workstations

(5)   It is possible that high human and mosquito population densities contributed to this outbreak

With regard to the discourse function of extraposed constructions such as those in (3), (4), and (5), the study carried out by Rowley-Jolivet and Carter-Thomas (2005: 51) revealed two primary uses of extraposition. First, it is used to "increase dynamism" by placing new pieces of information at the end of a sentence. Second, it allows a speaker/writer to express evaluative opinions in a "rhetorically effective way" by introducing evaluative comments at the beginning of a sentence. These findings are consistent with the results of the research studies conducted by Collins (1994), Gómez-González (1997), Herriman (2000a), Hoey (2000), and Hewings and Hewings (2002). A different view for the functional properties of extraposed clauses is expressed by Mair (1990: 39), Miller (2001), and Kaltenböck (2005), who found that extraposed clauses convey given as well as new information.

While the discourse function and the syntax of extraposed constructions have received systematic treatment in the literature, the role of the adjectives in various types of extraposed patterns has largely been disregarded. Thus, research into the occurrence of adjectives in *it*-extraposition with *that*-clauses deserves more attention. The valid argument for carrying out such research is that the meaning of each adjective is a relevant factor affecting the constructional meaning. For example, the adjectives *clear, essential,* and *possible* in (3), (4), and (5) contribute substantially to the understanding of the illustrative sentences by attributing the meanings of obviousness, necessity, and possibility to the constructions under study. Hence, the results of the quantitative analysis of the adjectives and their semantic description in terms of the semantic frames they evoke may allow us to uncover subtle distributional differences in their occurrence and to understand their role in the investigated pattern, as well as to deepen our understanding of the meaning and function of the construction.

## 5. Findings and discussion

The data under study were extracted from the academic sub-corpus by means of the concordancing program. This tool was used to search through the corpus for all the occurrences of adjectives in the search string (it is adjective …) as well as for the immediate context in which each instance occurred, creating a concordance. The corpus search for all adjectives used in this search string provided 18090 instances of extraposed constructions complemented by different clauses, e.g. *to*-infinitives, *for/to*-infinitives, *wh*-clauses, *-ing* clauses, or NP + relative clauses. Each concordance line was then manually inspected to identify all com-

binations with the relevant pattern: the *it is ADJ that*-construction. All false hits (i.e. all anomalous uses) were discarded from further analysis, and the observed frequencies of the remaining instances of adjectives in the construction under investigation were calculated manually by reading concordance lines.

The search for all adjectives in the pattern complemented by *that*-clauses yielded 5391 occurrences of the *it is ADJ that*-construction, while this construction turned out to contain 167 types of adjectives, out of which 66 occurred only once in the investigated construction. By contrast, Wiliński's (2017) study, based on the same sub-corpus, found 9834 occurrences of the *it is ADJ to V*-construction comprising 336 types of adjectives, out of which 139 occurred only once in this pattern. In other words, the occurrence of the former turned out to be far less frequent than the latter in the academic section of COCA. This section, however, will only report the findings for the 30 most strongly attracted and repelled adjectives of the *it is ADJ that*-construction, since it is impossible to present and evaluate the results for all these adjectives in the space here allotted.

The findings reveal that the extraposed construction complemented by *that*-clauses displays a marked tendency to co-occur with adjectives expressing obviousness, likelihood, importance, necessity, certainty, doubt, and unusuality. Table 1 below shows the results of the measures of attraction and reliance for the 30 most strongly attracted adjectives of the *it is ADJ that*-construction. It also provides the total frequency of this construction, and the observed frequencies of adjectives in the pattern under study and in all extraposed constructions.

**Table 1.** The results of attraction and reliance for the thirty most strongly attracted adjectives
**a** = Frequency of adjective (e.g. *clear*) in the *it is ADJ that*-construction; **x** = Total frequency of the *it is ADJ that*-construction; **e** = Total frequency of adjective (e.g. *clear*) in all extraposed constructions

| rank | adjectives | a | x | e | attraction | reliance |
|------|------------|-----|------|------|------------|----------|
| 1. | clear | 921 | 5391 | 1069 | 17.08% | 86.16% |
| 2. | possible | 715 | 5391 | 1986 | 13.26% | 36.00% |
| 3. | likely | 537 | 5391 | 754 | 9.96% | 71.22% |
| 4. | important | 350 | 5391 | 2717 | 6.49% | 12.88% |
| 5. | unlikely | 334 | 5391 | 423 | 6.20% | 78.96% |
| 6. | true | 300 | 5391 | 458 | 5.56% | 65.50% |
| 7. | imperative | 186 | 5391 | 283 | 3.45% | 65.72% |
| 8. | evident | 168 | 5391 | 203 | 3.12% | 82.76% |
| 9. | essential | 155 | 5391 | 426 | 2.88% | 36.38% |
| 10. | apparent | 153 | 5391 | 173 | 2.84% | 88.44% |
| 11. | obvious | 122 | 5391 | 153 | 2.26% | 79.74% |
| 12. | critical | 104 | 5391 | 227 | 1.93% | 45.81% |

| rank | adjectives | a | x | e | attraction | reliance |
|------|------------|---|---|---|------------|----------|
| 13. | doubtful | 93 | 5391 | 127 | 1.73% | 73.23% |
| 14. | probable | 75 | 5391 | 82 | 1.39% | 91.46% |
| 15. | conceivable | 69 | 5391 | 71 | 1.28% | 97.18% |
| 16. | significant | 65 | 5391 | 94 | 1.21% | 69.15% |
| 17. | noteworthy | 59 | 5391 | 65 | 1.09% | 90.77% |
| 18. | necessary | 57 | 5391 | 1017 | 1.06% | 5.60% |
| 19. | certain | 52 | 5391 | 62 | 0.96% | 83.87% |
| 20. | crucial | 51 | 5391 | 172 | 0.95% | 29.65% |
| 21. | surprising | 46 | 5391 | 70 | 0.85% | 65.71% |
| 22. | interesting | 46 | 5391 | 226 | 0.85% | 20.35% |
| 23. | understandable | 43 | 5391 | 54 | 0.80% | 79.63% |
| 24. | inevitable | 38 | 5391 | 50 | 0.70% | 76.00% |
| 25. | vital | 37 | 5391 | 94 | 0.69% | 39.36% |
| 26. | undeniable | 34 | 5391 | 35 | 0.63% | 97.14% |
| 27. | remarkable | 32 | 5391 | 48 | 0.59% | 66.67% |
| 28. | plausible | 31 | 5391 | 65 | 0.58% | 47.69% |
| 29. | striking | 29 | 5391 | 55 | 0.54% | 52.73% |
| 30. | arguable | 27 | 5391 | 28 | 0.50% | 96.43% |

As can be seen in Table 1, the results are sorted according to the measure of attraction. The top of the table is dominated by relatively frequent adjectives, such as *clear, possible, likely, important, unlikely, true, imperative, evident, essential,* and *apparent.* The most logical explanation for this is that the total frequency of these adjectives in the academic section of COCA overall may obviously wield significant influence upon the likelihood of their occurrence in this pattern. For example, *clear* (Attraction score 17.08%) and *possible* (Attraction score 13.26%) achieved much higher scores for attraction than *striking* (Attraction score 0.54%) and *arguable* (Attraction score 0.50%), since they occurred much more frequently in the pattern than *striking* and *arguable*, as illustrated in Table 1. By contrast, the list for reliance comprises much higher scores for less frequent and more specialized adjectives occurring in the construction, such as *conceivable* (Reliance score 97.18%), *undeniable* (Reliance score 97.14%), and *arguable* (Reliance score 96.43%), as the formula employed for the calculation of reliance takes the total frequency of each adjective in other *it*-extraposed constructions in the corpus into consideration. For example, although *clear* occurs much more frequently in the *it is ADJ that*-construction than *arguable*, the latter obtains a much higher score for reliance because its overall frequency of occurrence in all extraposed constructions in the corpus is much lower (28 occurrences). As a result, the semantic affinity between *arguable* and the construction under scrutiny also appears to be extremely strong (96.43%).

The first set of the most strongly attracted adjectives of the *it is ADJ that*-construction consists of adjectives evoking the OBVIOUSNESS frame, a semantic frame in which a phenomenon, i.e. the entity and facts that are undeniable, is portrayed with respect to the degree of probability that it will be perceived or apprehended cognitively, as in *It is* [*very*] DEGREE *evident* [*that most of the teachers share this view in their teaching duties*] PHENOMENON. Its leading collexeme, *clear* in rank 1, is accompanied by *evident, apparent,* and *obvious* in ranks 8, 10, and 11. The scores of attraction and reliance reveal that the adjective *clear* accounts for 17.08% of the uses of the construction in the academic section of COCA, and that 86.16% of uses of the same adjective are found in that particular construction. The adjective is thus attracted to the pattern in a proportion of 17.08%, and relies on the construction in a proportion of 86.16%. Hence, the adjective is the most significant lexeme for the construction. By contrast, the adjective *obvious* is a much less important slot filler for the construction (Attraction score 2.26%) and relies on this pattern to a lesser extent (Reliance score 79.74%).

The second group in the ranking is constituted by a range of adjectives expressing possibility and likelihood. Its main adjective, *possible* (Attraction score 13.26%), in rank 2, is followed by *likely, unlikely, probable, conceivable,* and *plausible* in ranks 3, 5, 14, 15, and 28, respectively. This set of adjectives can be understood with respect to the LIKELIHOOD frame, in which the likelihood of a hypothetical event (a state of affairs or occurrence) is assessed by a judge, as in *It is possible* [*that high human and mosquito population densities contributed to this outbreak*] HYPOTHETICAL EVENT. *Possible, likely* and *unlikely* obtained much higher scores for attraction (13.26%, 9.96%, and 6.20%) than *probable* (Attraction score 1.39%), *conceivable* (1.28%)*, and *plausible* (Attraction score 0.58%), since they occurred much more frequently in the *it*-extraposition with *that*-clauses. *Possible* and *plausible* obtained relatively low scores for reliance (36.00% and 47.69% respectively), as they occurred more frequently in other *it*-extraposed constructions, particularly in the pattern complemented by *to*-clauses.

The third category comprises *important, significant,* and *crucial* in ranks 4, 16 and 20, whose meaning can be interpreted with reference to the IMPORTANCE frame. This frame describes a situation in which the importance of a factor affecting the outcome of a certain undertaking is assessed by the speaker, as in the sentence *It is important* [*that you learn these electronic instruments*] FACTOR. *Important* occupies the highest position among the adjectives falling into this category, attracted to the construction in a proportion of 6.49% but relying on this pattern in a proportion of 12.88%, which means that the adjective is used more frequently (in a proportion of 87.12%) in other *it*-extraposed constructions. By comparison, *significant* and *crucial* are much less significant lexemes of the pattern (Attraction scores 1.21% and 0.95%) and also rely on the pattern to a small degree (Reliance scores 69.15% and 29.65%).

Another category of strongly attracted lexemes of the construction under investigation includes adjectives such as *imperative, essential, critical, necessary,* and *vital,* bearing a similar but not identical meaning to that conveyed by the

adjectives *important, significant* and *crucial*. These adjectives, denoting extreme importance or necessity, instantiate the NECESSITY frame, a schematic knowledge structure pertaining to a dependent state of affairs which has a requirement as a prerequisite for obtaining or occurring, as in *It is essential that* [*all government elements are working together in coordination*] REQUIREMENT, or in *It is critical that* [*students feel they will not be evaluated and corrected when making these attempts*] REQUIREMENT. *Imperative,* ranked seventh, is the most strongly attracted lexeme belonging to this category (Attraction score 3.45%). It is followed by *essential, critical, necessary,* and *vital* in ranks 9, 12, 18, and 25, respectively. Note that *necessary* relies on the *it is ADJ that*-construction to the lowest degree (Reliance score 5.60%) as compared to other lexemes in the ranking list, which means that this adjective appears more frequently in other extraposed constructions, e.g. in the *it is ADJ to*-construction.

Among the most significant lexemes in the list, there are also adjectives referring to the CERTAINTY frame and the STIMULUS FOCUS frame. The former, evoked by *doubtful* and *certain* in ranks 13 and 19, concerns a speaker's certainty about the correctness of beliefs or expectations, i.e. the mental content that the speaker is certain or uncertain about, as in *It is certain* [*that I have never heard of any other, in his time or later*] CONTENT. The second, invoked by *noteworthy, surprising, interesting, remarkable,* and *striking* in ranks 17, 21, 22, 27 and 29, provides the background information about a stimulus, i.e. a person, an object, an event, or a state of affairs, bringing about a particular emotion or experience in an experiencer, as in *Under these circumstances, it is surprising* [*that the Republicans were ever elected at all*] STIMULUS.

The next group of lexemes, occupying ranks 6 and 23, consists of adjectives reflecting two different semantic frames: ADMITTING FACT (*true*) and BEING NORMAL OR ACCEPTED (*understandable*). The first frame refers to a particular fact or statement acknowledged as correct or true by a speaker, i.e. to a particular piece of information evaluated for its accuracy or in accordance with the actual state of affairs, as in *It is true* [*that our democratic systems do not work exactly as the autocratic ones*] FACT. In the second frame, a state of affairs is evaluated as normal, natural, or accepted with regard to a particular situation or under certain circumstances, as in *Of course, it is understandable* [*that the patience of some important members of the Security Council is running out*] A STATE OF AFFAIRS.

Finally, the bottom of the ranking list contains adjectives evoking the following frames: BEING UNAVOIDABLE (*inevitable*), BEING CERTAINLY TRUE (*undeniable*), and BEING POSSIBLY TRUE (*arguable*). The first frame describes a situation that cannot be avoided or prevented by an agent under certain circumstances, as in *It is inevitable* [*that this political debate is going to emerge*] SITUATION. In the second frame, a certain message is judged by a speaker to be certainly correct or true, thus being impossible to question or incapable of being denied or disputed, as in *It is undeniable* [*that the impact on councils' budgets is significant*] MESSAGE. In the third frame, a certain content, which can be a proposition to be believed or a course of action to be taken, is judged possibly true on the basis of evidence that

there are sufficient reasons for believing that the statement is true or that many people would agree with it, as in *From a philosophical standpoint it is arguable* [*that no such positive ethical right exists*] CONTENT.

At the final stage of the exposition, it is also essential to point out uncommon and unusual occurrences of adjectives in the *it*-extraposed construction complemented by *that*-clauses, i.e. those adjectives that are not strongly attracted to the investigated construction and occur less frequently than others in the academic discourse. Table 2 below displays the results of the measures of attraction and reliance for the 30 most strongly repelled adjectives in the pattern. As can be observed in Table 2, adjectives such as *accidental, alarming, conspicuous, credible, daunting, fearful, fortuitous* and several others are strongly repelled lexemes, since their occurrence in the investigated construction is extremely rare and the scores resulting from the calculation of the measure of attraction are very low: 0.02%, in all cases.

These adjectives, however, obtained much higher scores for reliance (Reliance score 100%) than those mentioned in Table 1, which means that no other uses of these adjectives were observed in *it*-extraposed constructions complemented by different clauses. This, in turn, may suggest that each occurrence of these adjectives in the academic discourse should be considered highly anomalous and exceptionally rare. *Accidental*, for example, only accounts for 0.02% of uses of the pattern in the academic register but relies on the construction in a proportion of 100%, which implies that this adjective solely occurs in *it is ADJ that*-constructions and fails to appear in other extraposed patterns in the academic part of COCA.

The adjectives mentioned above can be interpreted with respect to a range of semantic frames. However, this analysis will only focus on describing several adjectives in terms of the semantic frames they evoke, largely because of the limitation of space. For example, *accidental* can be characterized relative to the COINCIDENCE frame, a frame in which a certain state of affairs, i.e. a state or event, is perceived as being causally independent of other states or events that might have been expected to be causally-related states of affairs, as in *It is accidental* [*that they take a coupe rather than another type of carriage to the theatre*] STATE OF AFFAIRS. *Alarming* and *daunting* can be interpreted with respect to the STIMULUS FOCUS frame (see above), while *credible* and *fortuitous* in terms of the TRUST frame and the LUCK frame, respectively. The TRUST frame refers to a situation in which a cognizer expresses an opinion about the reliability of the information provided by a particular information source, as in *While it is credible* [*that gender is not fundamental*] INFORMATION, *it is also true that men are less likely to initiate policies that affect women's lives directly.* The LUCK frame in turn describes a state of affairs that is judged as being good, or bad, with respect to background knowledge according to which this state of affairs was previously assumed to be unlikely to occur, as in *Therefore it is fortuitous* [*that environmental awareness, which strengthens the antiwar pressures for a more cooperatively conceived organization of life on earth, may at the same time reduce some of the hazards of concentration of authority*] STATE OF AFFAIRS.

By contrast, adjectives such as *explicit, imminent, lucky*, *promising, symbolic, terrible, amusing,* and *perplexing* are attracted in same proportion to the construction but rely on this construction to a lesser extent (Reliance score 50%) than those mentioned above. This means, in turn, that these adjectives appear to occur in different extraposed patterns, as can be inferred from Table 2 by comparing the observed frequencies of these adjectives in the *it is ADJ that*-construction with their total frequencies in all other extraposed constructions.

With regard to semantic properties of these adjectives, they reflect various semantic frames. The meaning of the adjective *explicit,* for example, can be relativized to the CANDIDNESS frame. In this frame a speaker conveys a certain message, the truth or sincerity of which is at issue, as in *In Article 34 of the UN Convention on the Rights of the Child, it is explicit* [*that children have the right to be protected from sexual exploitation*] ₘₑₛₛₐ𝓰ₑ. *Amusing* and *perplexing* invoke the STIMULUS FOCUS frame. In this frame a stimulus, e.g. an object or event, brings about a particular emotion or experience in an experiencer, as in *First, given his important role in solar-neutrino studies in the 1980s and 1990s, it is amusing* [*that Hans did not include a neutrino in any of the paper's nuclear re action equations*] ₛₜᵢₘᵤₗᵤₛ. *Terrible* in turn activates the DESIRABILITY frame concerning a certain evaluee being implicitly judged good or bad with respect to its quality (other instances of its type), as in *You know, you have a very great mind and it is terrible* [*that it would be wasted*] ₑᵥₐₗᵤₑₑ.

The most logical explanation as to why adjectives such as *daunting, hopeful, refreshing, lucky, terrible, amusing, uncanny,* or *perplexing* are among the most strongly repelled lexemes of this construction in the academic section of COCA lies in the nature of the academic register. First, academic discourse indicates a strong preference for formal words and grammatical structures over informal ones, whereas adjectives such as those mentioned above are more frequently encountered in informal speech. For example, as can be observed in COCA, the adjective *lucky* tends to occur more commonly in the spoken section (4845 tokens) than in the academic sub-corpus (only 726 tokens). Similarly, *hopeful* is also used more frequently in spoken discourse (1999 occurrences in COCA) than in academic prose (639 occurrences).

Second, adjectives such as *clear, possible, necessary, likely,* and many others are typically used in *it*-extraposition to perform a particular function in an academic context. While the *that*-clause conveys the propositional content and expresses some state of affairs, the matrix clause (e.g. *It is clear*) expresses some speaker's evaluation or stance (albeit not directly attributed to a specific speaker due to the use of the impersonal *it*), and as such serves the function of providing impersonal judgements about the importance, likelihood, or obviousness of a particular fact, statement, or a situation. By contrast, adjectives such as *daunting, lucky, terrible, fearful,* and *amusing* refer to emotional reactions that are brought about by the mention of a given state of affairs in everyday speech and in informal contexts. Apart from informal words expressing emotive reactions, the ranking list for the most strongly repelled adjectives also seems to contain more formal

adjectives such as *inarguable, inferable,* or *presumable.* A possible reason for their loose association with the investigated pattern, however, is that these tend to occur infrequently in academic discourse overall. A search for these adjectives in the academic part of COCA, for example, provided 14 occurrences of inarguable, 6 occurrences of *inferable*, and 11 uses of *presumable.*

**Table 2.** The results of attraction and reliance for the thirty most strongly repelled adjectives

| rank | adjective | a | x | e | attraction | reliance |
|------|-----------|---|------|---|------------|----------|
| 1. | accidental | 1 | 5391 | 1 | 0.02% | 100% |
| 2. | alarming | 1 | 5391 | 1 | 0.02% | 100% |
| 3. | altruistic | 1 | 5391 | 1 | 0.02% | 100% |
| 4. | conspicuous | 1 | 5391 | 1 | 0.02% | 100% |
| 5. | credible | 1 | 5391 | 1 | 0.02% | 100% |
| 6. | daunting | 1 | 5391 | 1 | 0.02% | 100% |
| 7. | fearful | 1 | 5391 | 1 | 0.02% | 100% |
| 8. | fortuitous | 1 | 5391 | 1 | 0.02% | 100% |
| 9. | hopeful | 1 | 5391 | 1 | 0.02% | 100% |
| 10. | inarguable | 1 | 5391 | 1 | 0.02% | 100% |
| 11. | incontestable | 1 | 5391 | 1 | 0.02% | 100% |
| 12. | indispensable | 1 | 5391 | 1 | 0.02% | 100% |
| 13. | indubitable | 1 | 5391 | 1 | 0.02% | 100% |
| 14. | inexplicable | 1 | 5391 | 1 | 0.02% | 100% |
| 15. | inferable | 1 | 5391 | 1 | 0.02% | 100% |
| 16. | presumable | 1 | 5391 | 1 | 0.02% | 100% |
| 17. | refreshing | 1 | 5391 | 1 | 0.02% | 100% |
| 18. | uncanny | 1 | 5391 | 1 | 0.02% | 100% |
| 19. | unimaginable | 1 | 5391 | 1 | 0.02% | 100% |
| 20. | unmistakable | 1 | 5391 | 1 | 0.02% | 100% |
| 21. | well-documented | 1 | 5391 | 1 | 0.02% | 100% |
| 22. | well-recognized | 1 | 5391 | 1 | 0.02% | 100% |
| 23. | explicit | 1 | 5391 | 2 | 0.02% | 50% |
| 24. | imminent | 1 | 5391 | 2 | 0.02% | 50% |
| 25. | lucky | 1 | 5391 | 2 | 0.02% | 50% |
| 26. | promising | 1 | 5391 | 2 | 0.02% | 50% |
| 27. | symbolic | 1 | 5391 | 2 | 0.02% | 50% |
| 28. | terrible | 1 | 5391 | 2 | 0.02% | 50% |
| 29. | amusing | 1 | 5391 | 2 | 0.02% | 50% |
| 30. | perplexing | 1 | 5391 | 2 | 0.02% | 50% |

## 6. Concluding remarks

In conclusion, the quantitative study of adjectives occurring in the *it is ADJ that*-construction has indicated that there are indeed adjectives that form strong or loose associations with the investigated pattern, and that there are either fairly regular (i.e. common or frequent) or relatively uncommon (i.e. rare and unexpected) occurrences of some adjectives in the *it*-extraposition complemented by *that*-clauses. In addition, the results seem to suggest that frame-semantic knowledge of adjectives is a significant factor directly affecting the meaning of this pattern. In other words, the meanings of adjectives and frame-semantic knowledge associated with them seem to contribute substantially to the meanings of the construction. For example, as can be noticed in section 5, the adjectives *clear*, *possible*, *likely*, *important, certain*, or *necessary* directly affect the understanding of the illustrative sentences by attributing the meanings of obviousness, possibility, likelihood, importance, certainty, necessity, or difficulty to the construction under study.

   The most obvious finding to emerge from this study is that there are several groups of adjectives that are closely related semantically being strongly associated with the *it is ADJ that*-construction. More specifically, it was uncovered that adjectives evoking different semantic frames, such as OBVIOUSNESS (*clear, evident, apparent, obvious*), LIKELIHOOD (*possible, likely, unlikely, probable, conceivable, plausible*), IMPORTANCE (*important, significant, crucial*), NECESSITY (*imperative, essential, critical, necessary, vital*), and CERTAINTY (*doubtful, certain*) are the most significant lexemes contributing to the meaning and function of the construction. Among the most strongly attracted lexemes of this construction we can also find adjectives expressing different forms or ways of evaluating facts, statements, or states of affairs: MENTAL STIMULUS (*surprising, interesting, remarkable, striking, noteworthy*), ADMITTING FACT (*true*), BEING NORMAL OR ACCEPTED (*understandable*), BEING UNAVOIDABLE (*inevitable*), BEING CERTAINLY TRUE (*undeniable*), and BEING POSSIBLY TRUE (*arguable*).

   The occurrence of these and other significantly attracted adjectives in this type of *it*-extraposition supports that *it is ADJ that-clause* is a primarily semantic construction that demonstrates a marked preference for adjectives expressing a subjective judgment about how obvious, necessary, important, possible, or true a certain fact, statement, or state of affairs is for a writer or speaker. Furthermore, the prevailing view about the function of the investigated pattern is also confirmed. As the illustrative examples in Section 5 show, the construction is regularly used in the academic register to convey new facts and states of affairs (cf. Kaltenböck 2005: 130–131) and to express the speaker's or writer's evaluative judgement in an indirect way by introducing the evaluative comments in the form of various adjectives followed by *that*-clauses and preceded by the copula *be* and dummy *it* (cf. Collins 1994; Herriman 2000a; Hoey 2000; Hewings and Hewings 2002; Rowley-Jolivet and Carter-Thomas 2005).

These findings also conform with those of earlier studies into the occurrence of adjectives expressing epistemic, deontic, and dynamic modality in extraposed constructions complemented by various clauses and showing different valency patterns (Mair 1990; Herriman 2000a; Biber *et al.* 1999; Kaatari 2010; Van Linden 2012). For example, both Mair (1990: 25) and Herriman (2000a: 592) noted that the adjectives expressing epistemic modality, such as *certain, clear, impossible, likely, possible, true,* and *unlikely,* occur relatively frequently in the it-extraposition with that-clauses and are almost exclusively complemented by *that*-clauses. Kaatari's small-scale study (2010), in turn, showed that epistemic adjectives (such as *clear, impossible, likely,* and *true*) constitute the most significant group among adjectives followed by *that*-clauses in academic discourse, and that the academic register is the register with the highest number of adjectives complemented by *that*- and *to*-clauses. Interestingly, the top of the ranking list in the current study is also dominated by epistemic adjectives (e.g. *clear, possible, likely, unlikely,* and *true*) expressing the likelihood of the content in the *that*-clause being true, i.e. "the speaker's opinion about the truth value of the extraposed clause" (Herriman 2000a: 587). In addition, the list contains deontic adjectives (e.g. *necessary, essential, critical,* or *imperative*) used to "indicate one's will on other agents" (Herriman 2000a: 585) and adjectives (e.g. *surprising* and *interesting*) expressing value judgements, in particular emotive reactions concerning the content in the complement clause (cf. Herriman 2000a: 586). These findings, thus, suggest that the *it*-extraposition with that-clause in the academic register co-occurs not only with epistemic adjectives but also with deontic ones and those pertaining to emotive reactions.

After a careful examination of the results, it can be concluded that a quantitative, corpus-driven approach can not only provide a detailed semantic description of the adjectives in the construction, but it can also allow a researcher to uncover minor distributional differences in the occurrence of these adjectives in the pattern under study. The quantitative method adopted in this study turned out to be a useful technique allowing for the determination of frequent and rare occurrences of adjectives in the extraposed construction with *that*-clauses, and hence it can be successfully employed for the investigation of adjectives in other types of *it*-extraposed constructions. Future research, therefore, might focus on comparing and contrasting adjectives found in a pair of *it*-extraposed constructions complemented by *to*-infinitive clauses and *that*-clauses. Such a quantitative analysis could reveal subtle distributional differences in the occurrence of adjectives in both constructions, and it may help us account for the existence and degree of semantic discrepancies between the two. Given that the current study was restricted to the academic register, it would also be interesting to examine the distribution of adjectives across different types of both written and spoken registers, in view of the possible existence of slight variations in their occurrence.

# References

Biber, Douglas, Susan Conrad, Randi Reppen (1998) *Corpus Linguistics: Investigating Language Structure and Use*. Cambridge: Cambridge University Press.

Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan (1999) *Longman Grammar of Spoken and Written English*. Harlow: Pearson Education.

Calude, Andreea S. (2008) Clefting and extraposition in English. *ICAME Journal 32*, 7–34.

Collins, Peter (1994) Extraposition in English. *Functions of Language* 1, 7–24. https://doi.org/10.1075/fol.1.1.03col.

Croft, William (2001) *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Emonds, Joseph (1972) A reformulation of certain syntactic transformations. In: Stanley Peters (ed.) *Goals of Linguistic Theory*. Englewood Cliffs, NJ: Prentice-Hall, 21–62.

Fillmore, Charles J. (1982) Frame Semantics. In: The Linguistic Society of Korea (eds.) *Linguistics in the Morning Calm*. Seoul: Hanshin Publishing Company, 111–137.

Fillmore, Charles J. and Beryl T. Atkins (1992) Toward a frame-based lexicon: The semantics of RISK and its neighbors. In: Adrienne Lehrer and Eva F. Kittay (eds.) *Frames, Fields and Contrasts.* Hillsdale, New Jersey: Lawrence Erlbaum Assoc, 75–102.

Fillmore, Charles J. and Collin F. Baker (2010) A frames approach to semantic analysis. In: Bernd Heine and Heiko Narrog (eds.) *The Oxford Handbook of Linguistic Analysis*. Oxford University Press, 313–340.

Fillmore, Charles J., Christopher R. Johnson, and Miriam R. L. Petruck (2003) Background to FrameNet. *International Journal of Lexicography* 16(3), 235–250. https://doi.org/10.1093/ijl/16.3.235

Francis, Gill (1993) A corpus-driven approach to grammar: principles, methods and examples. In: Mona Baker, Gill Francis and Elena Toginini-Bognelli (eds.) *Text and Technology: In Honour of John Sinclair*. Amsterdam: John Benjamins, 137–156.

Goldberg, Adele (1995) *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: Chicago University Press.

Goldberg, Adele (2006) *Constructions at Work. The Nature of Generalization in Language*. Oxford: Oxford University Press.

Gómez-González, María Á. (1997) On subject it-extrapositions: Evidence from present-day English. *Revista Alicante de Estudios Ingleses* 10, 95–107.

Gómez-González, María Á. (2001) *The Theme-Topic Interface. Evidence from English*. Amsterdam and Philadelphia: John Benjamins.

Herbst, Thomas, David Heath., Ian F. Roe, and Dieter Götz (2004) *A Valency Dictionary of English: A Corpus-Based Analysis of the Complementation Patterns of English Verbs, Nouns and Adjectives*. Berlin: Mouton de Gruyter.

Herriman, Jennifer (2000a) Extraposition in English: A study of the interaction between the matrix predicate and the type of extraposed clause. *English Studies* 6, 582–599. https://doi.org/10.1076/enst.81.6.582.9180.

Herriman, Jennifer (2000b) The functions of extraposition in English texts. *Functions of Language* 7(2), 203–230. https://doi.org/10.1558/lhs.v6i1-3.129.

Hewings, Martin and Ann Hewings (2002) It is interesting to note that…: A comparative study of anticipatory 'it' in student and published writing. *English for Specific Purposes* 21, 367–383. https://doi.org/10.1016/S0889-4906(01)00016-3.

Hilpert, Martin (2014) Collostructional analysis: Measuring associations between constructions and lexical elements. In: Glynn, Dylan and Justyna Robinson (eds.) *Corpus Methods for Semantics: Quantitative Studies in Polysemy and Synonymy*. Amsterdam and Philadelphia: John Benjamins Publishing Company, 7–38.

Hoey, Michael (2000) Persuasive rhetoric in linguistics: A stylistic study of some features of the language of Noam Chomsky. In: Susan Hunston and Geoff Thompson (eds.) *Evaluation in Text. Authorial Stance and the Construction of Discourse*. Oxford: Oxford University Press, 28–37.

Huddleston, Rodney (1971) *The Sentence in Written English*. Cambridge: Cambridge University Press.

Huddleston, Rodney (1984) *Introduction to the Grammar of English*. Cambridge: Cambridge University Press.

Johnson, Christopher R., Charles J. Fillmore, Esther J. Wood, Josef Ruppenhofer, Margaret Urban, Miriam R.L. Petruck and Collin F. Baker (2001) *The FrameNet Project: Tools for Lexicon Building* [Version 0.7]. Berkeley, CA: International Computer Science Institute.

Kaatari, Henrik (2010) *Complementation of Adjectives: A Corpus-Based Study of Adjectival Complementation by that- and to-Clauses*. A master's thesis. Uppsala University, Department of English.

Kaltenböck, Gunther (2000) It-extraposition and non-extraposition in English discourse. In: Christian Mair and Marianne Hundt (eds.) *Corpus Linguistics and Linguistic Theory*. Amsterdam and Atlanta: Rodopi, 157–175.

Kaltenböck, Gunther (2003) On the syntactic and semantic status of anticipatory *it*. *English Language and Linguistics* 7(2), 235–255. https://doi.org/10.1017/s1360674303001096.

Kaltenböck, Gunther (2004) *It-Extraposition and Non-Extraposition in English: A Study of Syntax in Spoken and Written Texts*. Wien: Braumüller.

Kaltenböck, Gunther (2005) It-extraposition in English: A functional view. *International Journal of Corpus Linguistics* 10(2), 119–159. https://doi.org/10.1075/ijcl.10.2.02kal.

Mair, Christian (1990) *Infinitival Complement Clauses in English*. Cambridge: Cambridge University Press.

McCawley, James D. (1988) *The Syntactic Phenomena of English*. Volume II. Chicago: The University of Chicago Press.

Miller, Philip H. (2001) Discourse constraints on (non)extraposition from subject in English. *Linguistics* 39(4), 683–701. https://doi.org/10.1515/ling.2001.028.

Mindt, Ilka (2011) *Adjective Complementation*: *An Empirical Analysis of Adjectives Followed by that-Clauses*. Amsterdam and Philadelphia: John Benjamins.

Pérez-Guerra, Javier (1998) Integrating right-dislocated constituents: A study on cleaving and extraposition in the recent history of the English language. *Folia Linguistica Historica* XIX, 7–25.

Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1985) *A Comprehensive Grammar of the English Language*. New York and London: Longman.

Rosenbaum, Peter S. (1967) *The Grammar of English Predicate Complement Constructions*. Cambridge, MA: M.I.T. Press.

Rowley-Jolivet, Elizabeth and Shirley Carter-Thomas (2005) Genre awareness and rhetorical appropriacy: Manipulation of information structure by NS and NNS scientists in the international conference setting. *English for Specific Purposes* 24, 41–64.

Schmid, Hans-Jörg (2000) *English Abstract Nouns as Conceptual Shells. From Corpus to Cognition*. Berlin and New York: Mouton de Gruyter.

Schmid, Hans-Jörg and Helmut Küchenhoff (2013) Collostructional analysis and other ways of measuring lexicogrammatical attraction: Theoretical premises, practical problems and cognitive underpinnings. *Cognitive Linguistics* 24(3), 531–577. https://doi.org/10.1515/cog-2013-0018.

Seppänen, Aimo (1999) Extraposition in English revisited. *Neuphilologische Mitteilungen* 100, 51–66.

Seppänen, Aimo, Clas Göran Engström and Ruth Seppänen (1990) 'On the so-called anticipatory It.' *Zeitschrift für Phonetik, Sprachwissenschaft und Kommunikationsforschung* 43, 748–776.

Van linden, An (2012) *Modal Adjectives: English Deontic and Evaluative Constructions in Synchrony and Diachrony*. Berlin and Boston: Walter de Gruyter.

Wiliński, Jarosław (2017) Normal and anomalous occurrences of adjectives in extraposed constructions with *to*-infinitive clauses: A quantitative corpus-based study. In: Jarosław Wiliński and Joanna Stolarek (eds.) *Norm and Anomaly in Language, Literature and Culture.* Frankfurt am Main: Peter Lang, 89–106.

## Sources and tools

The Corpus of Contemporary American English (COCA). The full-text data (1990–2012). Available from https://www.corpusdata.org/purchase.asp

The FrameNet project. Available from https://framenet.icsi.berkeley.edu/fndrupal/ (accessed on 20 April 2017)

MonoConc Pro (MP 2.2). Available from http://www.athel.com/mono.html

JAROSŁAW WILIŃSKI is an Assistant Professor at the Institute of Modern Languages and Interdisciplinary Studies at Siedlce University of Natural Sciences and Humanities, where he teaches courses in linguistics. He is the author of several articles and book chapters on applications of cognitive linguistics in onomasiological lexicography and on the use of quantitative methods in linguistic research. His areas of interest include cognitive and corpus linguistics, lexicology, phraseology, and onomasiological lexicography. His current research focuses on applications of cognitive linguistics in the compilation of thematic reference works and on a quantitative corpus-based investigation of grammatical constructions, collocations, and metaphors.

Address: dr Jarosław Wiliński, Department of English Language, Faculty of Humanities, Siedlce University of Natural Sciences and Humanities, Żytnia 39, 08-110 Siedlce, Poland. [email: jarwili@poczta.onet.pl]