



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

## Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in  
Lingue Moderne per la Comunicazione e la Cooperazione Internazionale  
Classe LM-38

Tesi di Laurea

# *Different forms, different meanings?*

*A lexicographic and corpus-based study of four sets of  
English near-synonyms*

Relatrice  
Prof. Sara Gesuato

Laureanda  
Denise Celeghin  
n° matr.1036490/LMLCC

Anno Accademico 2013/2014



Ai miei genitori, a Morris e a Marco  
per avermi sostenuta  
e per aver sempre creduto in me.



## Table of Contents

<b>Chapter 1: Introduction</b> .....	<b>1</b>
1.1 Introduction.....	1
1.2 Relevant concepts in lexical semantics.....	2
1.3 Synonymy: etymology and meaning.....	5
1.4 Synonymy and the English language.....	7
1.5 Goals of the study.....	8
1.6 Outline of the dissertation.....	9
<b>Chapter 2: Literature review</b> .....	<b>10</b>
2.1 Introduction.....	10
2.2 Linguists' views on synonymy.....	10
2.3 Studies in the field of synonymy.....	13
2.4 Studies on synonyms and near-synonyms in the English language.....	14
2.5 Studies on synonyms in other languages.....	20
2.6 Conclusion.....	23
<b>Chapter 3: Method</b> .....	<b>24</b>
3.1 Introduction.....	24
3.2 Selection of near-synonyms.....	24
3.3 Data collection.....	24
3.3.1 Dictionary data.....	25
3.3.2 Corpus data.....	26
3.4 Conclusion.....	28
<b>Chapter 4: Findings from lexicographic sources</b> .....	<b>29</b>
4.1 Introduction.....	29
4.2 Dictionary data.....	29
4.2.1 The nouns.....	29
4.2.1.1 <i>Assassin</i> .....	29
4.2.1.2 <i>Murderer</i> .....	32
4.2.1.3 <i>Killer</i> .....	33
4.2.1.4 Conclusion.....	36

4.2.2 The verbs .....	37
4.2.2.1 <i>To disturb</i> .....	37
4.2.2.2 <i>To bother</i> .....	40
4.2.2.3 Conclusion .....	42
4.2.3 The adjectives .....	43
4.2.3.1 <i>Compulsory</i> .....	43
4.2.3.2 <i>Obligatory</i> .....	45
4.2.3.3 <i>Mandatory</i> .....	46
4.2.3.4 Conclusion .....	48
4.2.4 The adverbs.....	49
4.2.4.1 <i>Maybe</i> .....	49
4.2.4.2 <i>Perhaps</i> .....	51
4.2.4.3 <i>Possibly</i> .....	52
4.2.4.4 Conclusion .....	54
4.3 Conclusion.....	55
<b>Chapter 5: Findings from corpus sources .....</b>	<b>57</b>
5.1 Introduction .....	57
5.2 The nouns .....	57
5.2.1 <i>Assassin</i> .....	57
5.2.2 <i>Murderer</i> .....	66
5.2.3 <i>Killer</i> .....	73
5.2.4 Conclusion .....	81
5.3 The verbs .....	83
5.3.1 <i>To disturb</i> .....	83
5.3.2 <i>To bother</i> .....	90
5.3.3 Conclusion .....	96
5.4 The adjectives.....	98
5.4.1 <i>Compulsory</i> .....	98
5.4.2 <i>Obligatory</i> .....	103
5.4.3 <i>Mandatory</i> .....	108
5.4.4 Conclusion .....	113
5.5 The adverbs .....	114
5.5.1 <i>Maybe</i> .....	115
5.5.2 <i>Perhaps</i> .....	118

5.5.3 <i>Possibly</i> .....	120
5.5.4 Conclusion .....	123
5.6 Conclusion .....	124
<b>Chapter 6: Conclusion .....</b>	<b>125</b>
6.1 Introduction.....	125
6.2 The findings .....	125
6.2.1 The nouns .....	126
6.2.2 The verbs .....	129
6.2.3 The adjectives.....	131
6.2.4 The adverbs .....	133
6.3 Strengths and weaknesses of the research .....	134
6.4 Recommendation for further work.....	136
6.5 Conclusion .....	137
<b>References .....</b>	<b>139</b>





# CHAPTER 1

## INTRODUCTION

### 1.1 Introduction

This dissertation is meant to provide a contribution to the field of lexical semantics in the English language, more specifically addressing some issues regarding synonymy and near-synonymy. As is well known, lexical semantics deals with the meaning of words and phrases, that is, the lexical units that make up the vocabulary of a language, independently of whether they are realized as single words or groups of words. It describes the entities that lexical units denote and the concepts they convey; the syntactic and pragmatic<sup>1</sup> similarities and differences among words belonging to the same semantic field; the semantic contribution made by individual lexical units to the overall meaning of the utterance in which they occur; and the semantic-structural relationships among words within a language (e.g. synonymy, antonymy, hyponymy, polysemy and others). More simply put, lexical semantics describes lexical meaning, which has been defined as : “[t]he meaning of a word considered in isolation from the sentence containing it, and regardless of its grammatical context” (*Oxford Dictionary* <http://oxforddictionaries.com>), “the equivalent to the commonly used, less technical (but ambiguous), term ‘word-meaning’” (Lyons 1995: 47) and “the meaning of individual words” (Kearns 2000: 3).

In particular, studies on synonyms and near-synonyms investigate the relation occurring between words or expressions seemingly having the same meaning. As will become clear in sections 2.2, 2.4 and 2.5, virtually all of these studies observe that there may be only few, if any, cases of absolute synonymy, that is, when all the meanings of two words are identical such that either term can be used indifferently in all contexts. Rather, these studies reveal that there are many cases of quasi synonymy, that is instances in which words which have similar, rather than identical meaning, and with different (i.e. complementary or partly overlapping) distribution in language use such that a speaker cannot freely interchange the near-synonyms without causing a change in meaning in the utterance in which they are used. Such studies are of interest to the linguistic community, because their findings help shedding light on aspects of the

---

<sup>1</sup> *Pragmatic* here means ‘with relevance to their typical context of use’, such as register, topic/field, geographic distribution.

semantic structure of a language which native speakers themselves might not necessarily be aware of, and thus they make it possible to account for and reliably predict in which contexts given terms are (to be) used. In addition, such studies can be helpful to foreign language learners, who may find it difficult to appreciate the different nuances of meanings of seemingly very similar terms, and who therefore might not know according to what criteria to prefer one over others in given contexts.

The above observation is particularly relevant to the English language. Because of its history, English is particularly rich in synonyms and near-synonyms of Germanic and Latin-Romance origin. Its vocabulary often appears to be much more varied than that of other languages, including Italian. But this richness may lead to confusion, in the sense that it may be difficult to assert to what extent (near-)synonyms are comparable in the way in which they are used.

Several studies have been carried out on (near-)synonymic terms in various languages, including English (see sections 2.4 - 2.5), but much more remains to be done. As the literature review (see chapter 2) will show, a set of near-synonyms normally have a common denotational meaning but also differ in terms of connotation, shades of meaning, idiomatic and stylistic usage. As a result, the choice of the word that perfectly fits the context requires great attention. In this dissertation, I aim to contribute to shedding light on the use of English near-synonyms by analysing some sets of semantically similar terms, representative of four parts of speech, namely the nouns *murderer*, *killer* and *assassin*; the verbs *to disturb* and *to bother*; the adjectives *compulsory*, *obligatory* and *mandatory*; and the adverbs *maybe*, *perhaps* and *possibly*, which, as far as I know, have not been previously investigated.

In this chapter, I will first give an overview of the basic notions which are commonly discussed within the field of lexical semantics and which are relevant to synonymy. Then I will define what synonymy is, introduce some of the publications which will be better presented in the second chapter and give a historical overview of vocabulary development in the English language. Finally, after having described the aim of this work, I will provide a brief outline of this dissertation.

## **1.2 Relevant concepts in lexical semantics**

To describe and classify the meaning of words, and detect their relationships, several concepts are used in lexical semantics, namely polysemy, homonymy, part-whole (or meronymy), presupposition, hyponymy, antonymy, semantic feature,

prototypes, denotation, sense and connotation. These are of central importance when examining similarities and differences among (near-)synonyms.

Both polysemy and homonymy deal with multiplicity of meaning, but while polysemy refers to the set of different meanings which are conveyed by the same word – and listed under the same lexical entry in dictionaries (e.g. *flight*: ‘passing through the air, power of flying, air journey, unit of the Air Force, volley, digression, series of steps’; Palmer 1981: 100), homonymy refers to several words with the same orthographic-phonetic shape, each with a different meaning, and which are listed under separate entries in dictionaries (e.g. *gay*: ‘lively, light-hearted, bright’; *gay*: ‘homosexual’; Saeed 1997: 65).<sup>2</sup>

Meronymy is “a part-whole relationship between lexical items” (Saeed 1997: 70), or better, “the semantic relationship between a lexical item denoting a part and that denoting the corresponding whole” (Cruse 1986: 159). For this reason, the semantic relationship can be identified with the sentence frame “X is part of Y”, e.g. *finger* and *palm* are meronyms of *hand* (Kearns 2000: 10).

Presupposition is a tacit, background assumption made about a state of affairs as implicitly triggered by a given term; for example, the sentence “*When did you stop smoking?*” presupposes that the addressee used to smoke, or similarly, “*The King of France is bald*” presupposes that France has a king (Palmer 1981: 5, 166).

Hyponymy is “a relation of inclusion” (Saeed 1997: 68) in the sense that the meaning of a hyponym is included in the meaning of a more general word, called *superordinate*; Palmer (1981: 85) gives the following examples: “*tulip* and *rose* are included in *flower*, and *lion* and *elephant* in *mammal* (or perhaps *animal*). Similarly *scarlet* is included in *red*”.

The term *antonymy* indicates the semantic relation of ‘opposition’, and it is the opposite of *synonymy*. There are different kind of oppositions, which are discussed, for instance, by Saeed (1997), Palmer (1981) and Cruse (1986). There are pairs of antonyms called *complementary pairs*, in which the “positive of one item implies the negative of the other” (Saeed 1997: 66) such as *dead/alive*, *pass/fail*. There are also antonyms, typically adjectives, in which the positive of one term does not necessarily imply the negative of the other one; these antonyms identify the extremes of a continuum of a quality that may exist to different degrees, such as the gradation of age,

---

<sup>2</sup> However, words currently perceived as homonyms by speakers might, in fact, historically originate from one and the same term, as is indeed the case with *gay*.

size and so on (e.g. *old/young*, *big/small*, *good/ bad*). Then there are antonyms which are characterized by a reverse relation of literal movement – or metaphorical evolution – in which one term describes movement/evolution in one direction and the other the same movement/evolution in the opposite direction, such as *push/pull*, *come/go*, *ascend/descend*. And finally there are converse antonyms, both of which describe the same relation between two entities, but from alternative points of view, such as *above/below*, *employer/employee*, *north of/south of*, *buy/sell*, *wife/husband*.

Besides describing the semantic relationships between words, lexical semantics also deals with the semantics traits of individual words. The semantic make-up of a word can be described in terms of its semantic features, prototype structure, denotational and connotational meanings.

From one perspective, words may be analysed in terms of their sense components, also known as semantic features. These are markers that express the meaning of lexical items and of structures containing them, and are “intended to reflect, in their formal structure, the structure of the concepts they represent” (Katz 1972: 101) or, in other words they express the existence or non-existence of semantic properties which contribute to define the meaning of a word: for example, words like *woman* vs *girl*, *sheep* vs *lamb* are identified by a marker which is a characteristic that represent the element preceded by a plus or minus sign: ‘Adult’ ([+Adult]) as opposed to ‘Non-adult’ ([-Adult]). So, the word *stallion* is [+EQUINE, +ADULT, +MALE], *lamb* is [+OVINE, +JUVENILE] and *pig* is [+PORCINE] (Kearns 2000: 10).

Prototype theory was first defined by Rosch (1978), who affirmed that people categorize items and concepts on the basis of a so-called prototype or ideal representation. According to this theory, items that instantiate all or the majority of the features of a category are considered as central members, and are called prototypes of that category. On the other hand, items that do not have in common the majority of such features, may still belong to the category without being prototypes, that is, they instantiate peripheral members of that category. The following examples clarify the definition: *chair* is more central member of the category *furniture* than *lamp*, and *sparrow* is a more typical member of the category *bird* than *penguin*. For this reason, the word *whale* is not typical of the category *mammal*, being far from the central prototype. The prototype theory is relevant to lexical semantics because it helped researchers in the study of vocabulary and mental lexicon; it is also important to show the relations occurring between items in categories.

Finally, the meaning of words can also be described in terms of their denotation and connotation. Denotation represents the connection between the linguistic expression and reality, that is, the literal, linguistic meaning of a term thanks to which entities can be identified in the world (Palmer 1981: 18-20; Lyons 1995: 78-82; Kearns 2000: 2-3): for example, when we look up the meaning of the word *snake* in monolingual dictionaries, we will find a definition that will help us to recognize its possible referent in given contexts (e.g. “a long limbless reptile which has no eyelids, a short tail, and jaws that are capable of considerable extension. Some snakes have a venomous bite”; *Oxford Dictionaries* definition <http://oxforddictionaries.com>). On the other hand, sense is defined as one of the linguistic meanings of a word, and in particular, the way in which a word or expression can be interpreted, independently of its denotational meaning. Finally, connotation represents the emotive or evaluative meaning indirectly conveyed by a word, which may be related to the characteristics associated to the items that a given word defines (Palmer 1981: 92). For example, the word *snake* suggests the connotation of something stealthy, dangerous, predatory, unethical.

So, as we have seen there are many elements which are involved in understanding the use of a word and defining its linguistic meaning(s). All these are to be taken into consideration when dealing with synonymy, because each can partly shed light on the similarities and differences between words with very similar meaning.

### **1.3 Synonymy: etymology and meaning**

The term (*near-*)*synonym* is usually used to mean words that have the same or similar meanings and, as we have already seen in the previous section, many notions enter into a description of a word’s meaning. All the facets of meaning outlined above contribute to determining the kind and amount of similarity between (*near-*)*synonyms*. In this section I will give an overview of the concept of synonymy, by drawing on definitions of this concept provided by semanticists and retrieved in dictionaries.

The online version of the *Oxford Dictionaries* (<http://oxforddictionaries.com/>) explains the origin of the word *synonym* as follows: it reaches late Middle English via Latin from Greek, where *sunōnumon* was a de-adjectival noun – or better, the neuter form of the adjective *sunōnumos*, but used as a noun – comprising the preposition *sun-* ‘with’ and the noun *onoma* ‘name’. So, the original meaning of *synonym* is something like ‘co-name’. The *Oxford Dictionaries* defines *synonym* as “a word or phrase that means exactly or nearly the same as another word or phrase in the same language”, and

exactly the same definition is found in the *Cambridge Dictionaries Online* (<http://dictionary.cambridge.org/>). Similar definitions are found in other monolingual dictionaries: the *Longman Dictionary of Contemporary English* (<http://www.ldoceonline.com/>) defines *synonym* as “a word with the same meaning as another word in the same language”; the *Merriam Webster Dictionary* (<http://www.merriam-webster.com/>) provides definitions for three senses, namely “(1) one of two or more words or expressions of the same language that have the same or nearly the same meaning in some or all senses; (2) a word or phrase that by association is held to embody something (as a concept or quality) <a tyrant whose name has become a *synonym* for oppression>; (3) one of two or more scientific names used to designate the same taxonomic group — compare homonym”. Similarly, the *Collins English Dictionary* (<http://www.collinsdictionary.com/>) includes more than one definition under the entry for *synonym*: “(1) a word that means the same or nearly the same as another word, such as *bucket* and *pail*; (2) a word or phrase used as another name for something, such as *Hellene* for a *Greek*; (3) (biology) a taxonomic name that has been superseded or rejected”. Overall, therefore, the above definitions suggest that *synonyms* can be defined as two or more words having the same or very similar meanings within the same language.

Of course, the notion of synonymy has been largely examined in lexical semantics by several scholars, whose theories I will investigate more thoroughly in the second chapter; here I will only introduce the definitions given by some of them on the topic. Palmer (1988: 88) defines *synonymy* as “sameness of meaning”, which can be identified by the use of substitution. However, there are no real synonyms, that is words having *exactly* the same meaning, but rather partial and near-synonyms, “expressions that are more or less similar, but not identical in meaning” (Lyons 1995: 60).

Cruse (1986) also studied synonymy considering cases of synonymy those in which “certain pairs or groups of lexical items bear a special sort of semantic resemblance to one another” (p. 265). Cruse focused his attention on the fact that some pairs of items are more synonymous than others because of the differences of overlap of their semantic traits and tried to identify absolute synonymy: he suggested a test for revealing the extent to which two words can be considered synonymous. This consisted in inserting them into sentential contexts and looking at possible semantic and syntactic similarities and differences; he found that the terms not always had the same

collocational patterns and acceptability and thus, he assumed that absolute synonymy was not always possible.

Saeed (1997: 65) defined synonyms as “different phonological words<sup>3</sup> which have the same or very similar meaning”. He gave only an overview of synonymy and briefly wrote on the influence of registers, regional differences, collocational and distributional restrictions when choosing a synonym to use.

As we have seen, both the dictionary definitions and the linguists’ definition considered agree on describing *synonymy* as a relation between different words that have more or less the same meaning.

#### **1.4 Synonymy and the English language**

The vocabulary of a language changes all the time. New words are added and others fall out of use depending on communication needs and goals of speakers. If a new word enters a language to express a concept for which another term is still available, then the opportunity for synonymy or near-synonymy arises. That is, if two words start or continue to co-exist side by side, they might either specialize, developing slightly different meanings, or they might be used interchangeably to convey the same meaning. This happened to English more than once. It has always accepted new words coming from other languages without discarding “previous” ones, as is reported in many books on the history of the English language, among these Strang (1971), Palmer (1981) and Danglli (2010).

The first important moment for the development of the British vocabulary happened in the 5<sup>th</sup>-6<sup>th</sup> centuries, when the Anglo-Saxons invaded Britain and introduced some basic, everyday terms like *man* and *woman*, which, later on, favoured the formation of new words in English language. In the 9<sup>th</sup>-10<sup>th</sup> centuries, invaders from Denmark and Norway brought into the country new words such as *scare* and *frighten*. Then followed the Norman invasion in 1066; the Norman invaders imported into Britain many words of French origin relevant to the administration, law, army and arts so that with the Norman aristocracy, French became the language of English courts, of the landed proprietors and of the upper classes. At this time, words were created containing combinations of Anglo-Saxon and Romance morphemes like *gentleman* (the first part of the word is French, whereas the second one derives from Old English). Another

---

<sup>3</sup> With phonological words are meant strings of sounds that behave as units for certain kinds of phonological processes, especially stress or accent.

important source of extra vocabulary was Latin, which spread widely in Britain in two historical moments, one being the advent of Christianity (6<sup>th</sup>-7<sup>th</sup> centuries) and the other the Renaissance (14<sup>th</sup>-17<sup>th</sup> centuries): words that were introduced during the former and the latter eras include *fraternal* and *corporeal*, respectively. Generally, words of Latin origin were used in more formal contexts, while native words in everyday circumstances. Finally a great contribution to the development of the language was given by Greek words which entered English through Latin: for example, the word *bishop*, which is taken from Latin but whose origin is in the Old Greek (*episkopos* ‘overseer’). In more recent times, as a result of Britain’s colonial expansion, English adopted many words from other languages such Dutch (*hoist*), American Indian (*skunk*), Australian aborigines (*kangaroo*), Indian (*bungalow*), Arabic (*average*) and Gaelic (*slogan*). These words introduced new concepts into the English culture and further enriched the English vocabulary.

As we will see in section 2.4, loanwords that entered the English language sometimes introduced “only” new terms for concepts or entities which English already had “names” for, and at other times, instead, they introduced both new terms and new concepts. In the former cases the opportunity for synonymy came up. Examples are: (a) *to begin* (which derives from Old English), *to commence* (from Old French), *to initiate* (of Romance origin); (b) *empty* (from Old English), *devoid* (from French) and *vacuous* (from Latin); (c) *brotherly* (from Old English) and *fraternal* (from Latin); (d) *bodily* (from Old English) and *corporeal* (from Latin) and others.

In languages such as English, the occurrence of synonyms reflects the occasions for linguistic contact created by historical events such as invasions, scientific developments and trade. However, how (near-)synonyms are being used in Present-Day English might or might not depend on the etymology of those words, which are thus worth investigating.

### **1.5 Goals of the study**

In this study I am going to analyse four sets of English (near-)synonyms which, to my knowledge, have not been previously taken into consideration by other linguists. The synonym sets I will analyse are likely to cause problems to Italian- learners of English as a foreign language because these terms have the same Italian translation equivalents and are defined in very similar ways in dictionaries. The terms in question are: the nouns *murderer*, *killer* and *assassin*; the verbs *to disturb* and *to bother*; the



adjectives *compulsory*, *obligatory* and *mandatory*; the adverbs *maybe*, *perhaps* and *possibly*. The aim of this work is intended to highlight similarities and differences between sets of near-synonyms by considering their dictionary definitions and corpus concordances, so that this may be of help to foreign language learners when they are in doubt about the right choice of words so as to avoid potential sources of error.

## **1.6 Outline of the dissertation**

Chapter two will present the literature review by reporting on studies in lexical semantics dealing with the notion of synonymy; in particular, these will show both the influence of context in speakers' choice of words and the rare occurrence of absolute synonymy across languages. Chapter three will motivate my choice of the synonyms to analyse and will outline the research method adopted. Chapter four will illustrate the findings from the dictionary survey. Chapter five will give an overview of the findings retrieved from the analysis of corpora. Chapter six will sum up and give an interpretation to the findings described in the previous chapters; this concluding chapter will point out the relevance of my work to the teaching and learning of lexicon, show the strengths and the weaknesses of this research and suggest possible future developments of this kind of research.

## CHAPTER 2

### LITERATURE REVIEW

#### 2.1 Introduction

In recent years an increasing number of studies have been carried out in the field of lexical semantics. In particular, some scholars have taken into consideration pairs of words with, apparently, very similar meanings in order to analyse their context of use and thus better define, describe and narrow down their specific meanings. In this chapter, I will report linguists' views on synonymy and near-synonymy, and then overview some studies carried out over the last 25 years which compare and contrast the meanings and contexts of use of pairs of words normally considered synonyms or near-synonyms.

#### 2.2 Linguists' views on synonymy

Many linguists focused their works on synonymy, which plays an important role in the field of lexical semantics. These linguists have tried to explore the complex facets of the semantic relations among the meanings of words and in particular, they tried to define synonyms and analysed the way in which synonymous words behave when inserted in sentential contexts. In this section I will give an overview of scholars' views on synonymy.

Cruse (1986) defined synonyms as lexical terms with a high degree of semantic overlap and a low degree of implicit contrastiveness adding that they share identical central semantic traits, namely those that determine the meaning of given lexical items. However, Cruse also stated that if absolute synonymy exists, it is extremely uncommon, and he predicted that in each pair of small set of near-synonyms, one of the terms would tend to fall in disuse or that some kind of semantic differences would develop. To better describe synonymy, the author discussed the following notions: the concept of scale of synonymity, propositional and expressive meaning, and finally propositional synonymy. As the name suggests, a scale of synonymity is a range of values or levels of synonymy, ordered from the lowest to the highest. At the lowest level, one can find words that have nothing in common, while at the highest, called absolute synonymy or zero synonymy, are placed words that have exactly the same semantic identity, that is meanings and collocational distributions. But Cruse observes that in many cases it is difficult to

establish where synonymy becomes non-synonymy; for example, in the following scale of synonymy, *rap* represents a quick, sharp knock made with knuckles, but also *thud* indicates a dull heavy sound: *rap:tap*, *rap:knock*, *rap:thwack*, *rap:bang*, *rap:thud* (Cruse 1986: 268). When dealing with the concept of synonymy, Cruse makes reference to two notions which outline the meaning of a word and identify its possible synonyms: such notions are the propositional and the expressive meaning. The idea of propositional meaning (also called *descriptive meaning*), which is strictly related to propositional truth condition, (see footnote 1) represents the meaning of an expression and in particular the relationship between the word and what it refers to / describes. It has to be distinguished from the expressive meaning, which expresses emotional impact to the speaker and is mostly expressed through expletives like exclamations, because the latter is tied only to the place and to the time of utterance, whereas the former is used also to refer to something which happened at another time. Hence, if two items are considered propositional synonyms, they must be identical with reference to propositional traits but they may be different in relation to their expressive traits. In particular, the concept of propositional synonymy indicates a truth-conditional relation<sup>4</sup> between synonyms: these kind of synonyms are syntactically identical (that is, they can be used in the same kind of sentences: declaratives, interrogative and so on) and when they are placed in exactly the same sentence, its truth conditions do not change; e.g. *He plays the violin very well* entails<sup>5</sup> and is entailed by *He plays the fiddle very well* (Cruse 1986: 88).

Palmer (1981) observed that synonymous words cannot be perfect synonyms on the grounds that two words with exactly the same meaning would not be able to survive in the same language. He listed several reasons that are responsible for lack of absolute synonymy: for example, near-synonyms may belong to different dialects of a language, e.g. farming terms vary depending on where the people who use them live (e.g. *cowshed* and *cowhouse* or *byre*, *haystack*, *hayrick* and *haymow*; palmer 1981: 89)<sup>6</sup>; or they may belong to different registers such as formal versus colloquial (e.g. *die* versus *pass away*); or they may also convey different emotive or evaluative meaning, although their denotational meaning is the same (e.g. *hide* and *conceal*, *politician* and *statesman*, *liberty* and *freedom* are said to imply different levels of approval or disapproval; in

---

<sup>4</sup> Truth conditions are properties of sentences. In particular, sentences can be true or false and truth conditions are the conditions under which a sentence can be said to be true or false.

<sup>5</sup> Entailment is used to “mean a relationship between sentences so that if a sentence A entails a sentence B, then if we know A we automatically know B. Or alternatively, it should be impossible, at the same time, to assert A and deny B.” (Saeed 1997: 4).

<sup>6</sup> Although he gives this example, Palmer (1981) does not mention which dialects those terms belong to.

particular, the second term of each pair being said to convey a more positive meaning than the other term). Palmer also observed that the use of synonyms may vary according to restrictions in their lexical collocation, that is, they occur only combined with other specific words. In order to identify real synonyms, Palmer proposed to check if candidates for synonymy are perfectly substitutable one for the other in the same sentential context or to contrast them by analysing their opposites.

Church et al. (1994) suggested a test of substitutability to help lexicographers to verify if given words are close synonyms, namely, if a word can be replaced by another one in the same context such that the meaning of the overall message does not change. The degree of substitutability of such words is high, and hence they have high possibilities to be synonyms. Church et al. also discussed *gradient synonymy*: this is a semantic relationship among semantically similar terms, or near-synonyms, located on a gradient scale of synonymy such that each near-synonym can find itself on a different step of the gradient scale. The concept of gradient synonymy is similar to the one proposed by Cruse on the scale of synonymy but with the difference that Church et al. made clearer the passages that let words move from synonymy to antonymy.

According to Lyons (1995), a distinction should be made between the concept of absolute synonymy and near-synonymy. Absolute synonymy is rarer than near-synonymy. It occurs when all the meanings of given terms are identical; when synonymous terms can be used indifferently in all contexts; and when they are semantically equivalent both in descriptive and non-descriptive terms, that is in terms of the state of affairs in the real world they describe, and in terms of the speaker's beliefs, attitudes and feelings they express. But absolute synonymy is not easy to be established because, even words that may be considered as having exactly the same meaning, may not represent absolute synonymy: for example, *big* has at least one meaning which is not shared with *large* (one would not say "I will tell my big sister" because it is an ambiguous sentence, "by the virtue of the polisemy of *big*"; Lyons 1995: 61-2) and they also cannot be used indifferently in all contexts because they have different idiomatic usages. On the other hand, near-synonymy is exemplified in words like *big* and *large*, *mist* and *fog*, *stream* and *brook* or *dive* and *plunge* which, even if they convey similar meanings, may differ in respect of the degree or nature of the feelings and attitudes expressed.

Saeed (1997) reconsidered what Palmer (1981) focused on, namely the rarity of absolute synonymy, and in particular he stressed the importance of registers, styles and

collocational restrictions in the selection of synonyms that, as he pointed out, can convey different attitudes of the speaker in relation to given situations (e.g. neutral, negative or positive). As an example, he mentioned words like *fuzz*, *flatfoot*, *pigs* or *the slime*, which convey negative speaker attitudes, whereas *cop*, which, like the other terms, occurs in informal use, seems to be a more neutral term.

In conclusion, the above semanticists observed that absolute synonymy is rare because when two semantically identical terms exist, they do not tend to survive within the same language. If two terms co-exist, there are two possible outcomes: either one disappears or one develops a new meaning. So-called synonyms more often tend to be “only” near- synonyms, that is, terms with slightly different meanings, especially denotational uses, or terms characterized by different linguistic or extra-linguistic contexts of use (e.g. register, syntactic frame, geographic distribution).

### **2.3 Studies in the field of synonymy**

Many scholars have investigated synonymy. Some have explored the semantic relation per se; others have compiled lists of synonyms; still others have developed conceptual or mathematical models so as to identify and outline semantic relationships between sets of (near-)synonyms. Here I am going to give a brief overview of some of these studies.

Brodda and Karlgren (1969) used algorithms, theorems and other mathematical models to identify sets of synonymous words and defined a numerical measure which showed the degree of connection/association between words. Similarly to Brodda and Karlgren (1969), Edmonds and Hirst (2002) developed a clustered model of lexical knowledge used in computational systems. A list of pairs of English synonyms was proposed by Wilding and Mohindra (1983); this list was created after having fifty College students complete a questionnaire in which they had to choose the most acceptable synonyms of the words given by the linguists; this research allowed the authors to highlight the preferred synonyms for each noun in 279 synonym pairs. Hudson et al. (1996) gave an exemplification of pairs of synonymous words which are, however, stylistically and syntactically different, in the sense that they belong to different levels of style and their syntactic constructions differ. Moss and Motta (2001) explained how to read Nicolò Tomaseo’s famous *Dizionario dei sinonimi della lingua Italiana* of 1830, which is a dictionary of Italian synonyms, by giving indications on how dictionary entries are set out (semantic frames, frame title, head-words,

equivalents, grammatical information, cross-referencing, examples of usage, notes, indexes). Inkpen and Hirst (2004) worked with algorithms to classify sentences according to the classes of distinction they express (i.e. denotational, attitude and style distinctions) and derived a lexical-knowledge-base from a dictionary of near-synonym discriminations.

As we have seen, some of the above mentioned studies are focused on mathematical processes, while others represent only lists of words. In the following sections I will review the work of some authors who investigated specific sets of (near-) synonyms.

## **2.4 Studies on synonyms and near-synonyms in the English language**

In every language it is possible to identify cases in which two or more words convey the “same” meaning or very similar meanings. Synonyms and near-synonyms are largely present also in the English language. In this section I will introduce the work of some linguists, who have examined specific sets of English (near-)synonyms, trying to identify to what extent they are similar in their use.

British English and American English are two varieties of the same language, but with considerable differences in their vocabulary. Filippov (1971) made an analysis of the coexistence and rivalry of American and British synonyms in British English as a result of American loans entering British English. According to Filippov, when an Americanism enters British English, and a word with the same meaning already exists, the outcome of this encounter may vary. First, the terms can coexist but with different levels of frequency of use, the Americanisms normally have a lower frequency of occurrence than the corresponding British synonym: e.g. BE (British English) *luggage* and AE (American English) *baggage*, BE *stones* and AE *rocks*, BE *tin* and AE *can*, BE *lorry* and AE *truck*, BE *Government* and AE *Administration*, BE *team* and AE *squad*, BE *autumn* and AE *fall*. However, there are a few cases in which the Americanism is more frequent than the British synonym, as in the case of AE *commuter* as opposed to *season ticket holder*. Second, two synonyms can differ stylistically, with the Americanism usually being used as a colloquial or slang term, and the British word as a neutral term: e.g. *intellectual* and AE *egghead*, *excuse* and AE *alibi*, *averse to* and AE *allergic*, *to advertise* and AE *to sell*. Third, they can also differ in lexical collocability, with some restrictions applying to the British English word: for example, the Americanism *merchant* replaced the traditional British *shopkeeper* and *dealer* in

particular combinations such as *coal merchant* or *wine merchant*. Finally, two synonyms can differ according to sociolinguistic values, which means that the assimilation of an Americanism into British English may vary according to the age of the person that uses it: e.g. the younger generation is more likely to adopt Americanisms, while the older generation prefers the traditional terms.

Also within the same variety of English, i.e. American English, (near-)synonyms abound. Among these Von Schneidemesser (1980) took into consideration the term *purse* and its synonyms, and by consulting the *Dictionary of American Regional English*<sup>7</sup> and by analysing informants' comments on the use of such synonyms, she found that *purse* and its synonyms *pocketbook*, *handbag*, *wallet*, *billfold*, *change purse* and *coin purse* varied in the geographic distribution of their use. With reference to *purse*, *pocketbook* and *handbag*, Von Schneidemesser highlighted that they were considered standard terms and therefore they had been reported from all parts of the country: in particular *purse* was more frequent on the West Coast and *pocketbook* on the East Coast, whereas *handbag* was more frequent in Maryland, and its use was related to the age of speakers (i.e. more common among people aged sixty or more years old). *Wallet* was more used in California and the Northeast area of the United States, generally conveying the meaning of 'leather billfold'; on the other hand, *billfold* was mapped outside the Northeast. With reference to its meaning, *purse* was considered by several informants as the oldest term referring to something that can be carried inside a *pocketbook* or a *handbag* and only in some cases did the informants consider its meaning as overlapping with that of *handbag*. Only *wallet* and *billfold* were usually considered as fold objects mainly used to contain money and smaller than the other objects. Finally, the frequency of use of *coin purse* and *change purse* was considered close behind that of *handbag* and the terms appeared to be used in the Upper Midwest area and in the South Midland area, respectively.

Gottschlank (1992) compared two English interrogative structures frequently used either to suggest something or inquire about something, namely *what about* and *how about*. In his research he involved college and high school students: he tested their ability as native speakers of English to select the best phrase to fill in the gaps in some sentences. According to the different interpretations given by the students to those sentences in terms of their inferred illocutionary force, it turned out that they made

---

<sup>7</sup> Von Schneidemesser does not give any publication date for the *Dictionary of American Regional English*.

distinctions between *how about* and *what about* and also that not all of them chose the same phrase in the same kind of utterance. The main differences that he noticed were that *what about* was mainly chosen when followed by the definite article, proper names, possessive and demonstrative pronouns, whereas *how about* was chosen when followed by the indefinite article, *some*, and non-finite clauses (e.g. infinitive or gerund). Moreover, *what about* was generally used as a reminder of something known to all interlocutors, whereas *how about* was used to provide new information or to express suggestions, persuasion or the invitation to do something. The two structures are also commonly used to make questions; in particular, since *how about* was used in questions expressing suggestions, which give options to the addressee, the most common answer to it was a yes/no answer. On the other hand, *what about* was used to ask for comments and opinions, so the most common response was the required information. Furthermore, the author analysed the use of the above-mentioned structures in literary texts and pointed out that in the literature examined, the choice between *how about* and *what about* was mainly determined by stylistic considerations and related to the structure of the plot; for example, in Scott Turow's novel *Presumed Innocent*, *how about* is used in the first hundred pages as a stylistic device to characterise the desultory approaches to identify the true murderer, but from chapter 23 on, *what about* is used to make inquiries and issue reminders. Still with regard to these texts, he also highlighted the difficulties encountered by translators when rendering these structures in German. So, Gottschlack showed that only sometimes was it possible to find overlaps in the dialogic and sentential distribution of these phrases and that the choice of *how about* rather than *what about* was used by the speakers to convey different nuances of meaning.

Church et al. (1994) studied the synonyms *ask for* and *request*, which they compared by using thesauri and corpora. They found several differences between the terms, by considering which verbs they could be replaced by, the direct objects they could take and their frequency of occurrence. With reference to the substitutable verbs, they listed *inquire*, *demand*, *claim*, *ask for* and *request*; moreover, *request* and *ask for* not only had a large overlap on direct objects, but also similar object distributions and meanings as well; both were accompanied with nouns of actions or states of affairs and a small number of nouns denoting human agents or agencies; finally, *request* is more frequent than *ask for* and it is more probable that the former can replace the latter than the contrary.



Atkins and Levin (1995) consulted electronic corpora and English monolingual dictionaries to analyse the verbs expressing the notion of shaking, namely *quake*, *quiver*, *shake*, *shiver*, *shudder*, *tremble* and *vibrate*. They examined their syntactic use, noticing how verbs normally considered intransitive, are instead also attested in transitive patterns, e.g. “An elevator shuddered him to the sixth floor” (Atkins, Levin 1995: 87). They also looked at the types of subject noun phrases co-occurring with these verbs and examined the definitions given by dictionaries for foreign language learners. They found that *quiver* is more likely to occur with reference to body parts, while *shiver* is used for people. Moreover *shudder*, *tremble* and *vibrate* were found to be used when talking about machineries; *quiver*, *shudder*, *tremble* and *vibrate* with reference to rooms and buildings; *vibrate* and *shudder* in relation to vehicles. Finally they also discovered that *shake* is the one with the most general meaning and for this reason the most frequent.

Duffley and Joubert (1999) analysed the verbs *intend*, *mean* and *propose* and their close synonyms *aim*, *plan* and *purpose* with reference to their syntactic construction with the gerund or the infinitive, which are the cause of problems of temporality and subject control. The authors found that the verbs “are the product of the interaction between the function of –ing and the lexical meaning of the main verb rather than being inherent ‘tense’ values of –ing”. Moreover, the solution to the problem of subject control was given by the two linguists, who defined it as a “logical property of certain types of utterances” (Duffley and Joubert 1999: 264).

Clift (2003) discussed the synonyms *actually* and *in fact* with particular reference to their occurrence in conversation. She noticed differences in the use of *actually* and *in fact* in terms of the position taken by each of them in a turn-at-talk and in terms of the composition of that turn. In particular, *actually* was found out to be used at the beginning or at the end of turns to indicate a change in the topic, a touch-off marker or the launch of a new story, while *in fact* appeared to be placed only at the beginning of a turn-constructive unit, creating a link with what had been said before.

Taylor (2003) used a one-million word corpus to analyse the distribution of the English adjectives *tall* and *high* in specific lexical collocations and sentential contexts. The author observed that the term *high*, mainly used when referring to physical bodies, buildings, constructions, clothing, topographical features and natural phenomena, is more frequent than *tall*, which is mostly used with reference to human beings but also plants and buildings. Taylor pointed out that both terms are related to the property of

vertical extent (which means that the upward extension of the entity being described is predominant over its other physical dimensions), but that only *high* can be used with reference to the vertical position (e.g. *high ceiling* implies that the ceiling is situated at a particular distance from the floor). Moreover, Taylor described *high* as the dominant term for the vertical dimension, in the sense that it is mapped onto a broad range of instances, whereas *tall* is considered the recessive term and is not mapped into a wide range of instances as *high* is.

Saeed and Fareh (2006) published an analysis on the contexts in which synonyms like *rob*, *steal* and *burglarize* may appear, and considered the different attitudes of interlocutors towards the situations in which these verbs were normally used, which could be either positive or negative. To this end, the authors took into consideration texts from newspapers like *The New York Times*, books on crime and criminal laws, magazines and electronic concordances. Saeed and Fareh identified a meaning shared by these synonyms, namely the performance of an illegal activity that involves depriving someone or something of something else, but also highlighted their different semantic behaviour. They noticed the following properties: (a) while the verbs *steal* and *rob* refer to thefts that can take place everywhere, either inside or outside a building, the verb *burglarize* narrows down its meaning to the description of illegal actions occurring only inside a house, a building or a secure place (this meaning shows up in the compound noun *house burglary*); (b) while *steal* can be used with direct objects identifying non-human, portable and concrete entities (or abstract entities with a metaphorical meaning), *rob* is followed by direct objects denoting human beings or places (houses or institutions) that are non-portable and concrete (if there is an abstract object a metaphorical meaning is conveyed), and *burglarize*, which is never used in association with a human being, indicates a theft that involves entering a building (e.g. “their house was burglarized last night”; Saeed-Fareh 2006: 330). Semantically, the authors noticed that the verb *steal* also appears to imply the theft of small things without force or violence, whereas *rob* and *burglarize* tend to co-occur with nouns denoting bigger and more valuable objects and with expressions relevant to the use of violence, force or threat. When the authors examined the connotative meaning of the terms, they found that both *rob* and *burglarize* are associated with a pejorative, negative connotation, while only *steal*, in its metaphorical use, can be connected with a positive meaning as in the sentence “She stole his heart” (Saeed-Fareh 2006: 333). Although the verbs are often mentioned as synonymous, this study revealed that this was not always

the case, and that they cannot be freely substituted for each other in the same context, each verb having its most typical context of use.

Gesuato (2007) took into consideration four pairs of English near-synonyms representative of four parts of speech, each of which has only one corresponding Italian equivalent (i.e. *island* and *isle*, *feeble* and *weak*, *gratefully* and *thankfully*, *to adore* and *to worship*). The author looked up dictionary definitions of each term of the above pairs, collected concordances of the terms from the *Collins-Cobuild Bank of English* corpus (Cobuild 1995), and compared and contrasted their frequency of occurrence and co-text of use. Both similarities and differences emerged. With reference to the nouns tested, she noticed that *island* can be used as the basic term with a more generic denotation, with a wide range of referents and as part of placenames, while *isle* is presented as a variant of the basic term appearing more frequently in placenames. Also, *island* was more frequent than *isle* and occurred in informal contexts; at the same time, both terms were often found in the singular form and in the written register. The analysis of the adjectives *feeble* and *weak* revealed that the former was more often used in attributive position, while the more frequent *weak* occurred in both attributive and predicative positions. Both adjectives appeared to preferentially co-occur with abstract rather than concrete nouns, and to convey the general meaning of ‘lacking strength’. With regard to the two adverbs examined, *gratefully* and *thankfully*, the author observed that the more frequent *thankfully*, of Germanic origin, was used as a clause adverb through which the speaker or writer gives a positive or negative evaluation of a global event or situation, while the less frequent *gratefully* more technically conveyed the meaning of ‘in a grateful way’ and was used to modify the verb phrase in a clause, that is, to express the manner in which a process was carried out. The verbs *to adore* and *to worship* also revealed different distributional patterns, and the author observed that only the former could occur in the past tense, while the latter was more commonly found in the simple present tense and never followed by a gerund. Furthermore, *to adore* turned out to be commonly used with subjects that denote single persons rather than crowds, whereas *to worship* tended to be associated with generic groups of people, thus signalling collective actions.

Cappuzzo (2010) analysed the medical terms *disease* and *illness* by taking into consideration information included in dictionaries and by consulting corpora. The dictionaries consulted presented the two words as synonyms used both in medical texts and in everyday language, *disease* having a more specific meaning and *illness* a more

general one. More specifically, in English-Italian bilingual dictionaries the terms were described as having the same meaning but in slightly different contexts, with *disease* referring to particular pathological conditions affecting specific parts of the body which may be analysed by physicians, and with *illness* being used as a generic term and usually when dealing with common language usage rather than with medical discourse (e.g. “She died after a long painful illness”, Cappuzzo 2010: 22). Monolingual dictionaries, on the other hand, gave a complete definition of *disease*, but only defined *illness* as a synonym of *disease* without explaining how to differentiate the two terms. Corpus data provided support for the information in the dictionary entries: the occurrences of *illness* were higher when indicating the condition of “not being well” and without referring to any specific pathological condition, while *disease* was more frequent when describing a specific pathological condition with reference to the specific organ damaged.

From these studies it emerged that synonyms not only convey different nuances of meaning, but also differ from the point of view of their collocational preferences, frequency, speakers’ attitude towards the situations being described, and the types of noun phrases they are usually combined with. Therefore, research suggests that so-called synonyms are hardly ever, if at all, freely interchangeable with each other<sup>8</sup>.

## 2.5 Studies on synonyms in other languages

In the following paragraphs I will report some studies carried out on near-synonyms in languages other than English.

Suàrez (1971) wrote an article in which he reported on a case of absolute synonymy. He conducted his research on an American group from the Province of Santa Cruz in Patagonia (Argentina), who speak the Tehuelche language. This tribe used various synonymous words to identify the same concrete object, and the informants confirmed that those words not only meant the same thing, but they also offered two or three words as corresponding to the Spanish equivalent<sup>9</sup>. More specifically, the informants used the synonyms indifferently when telling a traditional story, so that the author excluded the possibility of stylistic differences in the use of such synonyms. Moreover, Suàrez pointed out a tendency in the group to give

---

<sup>8</sup> Additional publications are available on the topic of English (near-) synonyms, but I was unable to access copies of them. They are: Y.B. Albader (2011), B. Levin, G. Song and B.T.S Atkins (1997), L.Lindvall (1986).

<sup>9</sup> Suarez (1971) do not give any illustrative example of Spanish equivalent words.

nicknames to individuals and to avoid the use of taboo words. Nicknames were given with reference to some moral or physical characteristics of the person being referred to, while the taboo rule forbade the use of name and nicknames of dead people for a year. So, because of the existence of the taboo rule, common nicknames temporarily fell into disuse and had to be replaced with words of related meanings for some time; the result was that new synonyms emerged, which could identify the same kind of characteristic or object as the taboo terms. The author therefore concluded that absolute synonymy exists in the Tehuelche vocabulary, in the sense that, after a one-year taboo period of lexical replacement, synonyms start co-exist together in the same language.

Sullivan (1987) examined six Russian synonyms (i.e. *vozle*, *bliz*, *okolo*, *vblizi*, *podle* and *u*) used to translate the word *near* by consulting standard reference sources and ordinary narrative texts. Sullivan based her study on the substitutability between the terms involved, and noticed that dictionaries defined each of the six terms with at least one of the other five, but also that, when such terms occurred in a context, their occurrences broadly varied. In fact, only in few cases did substitutability coincide with a full overlap in meaning, while in most cases it created semantic distinctions. In particular, their meanings overlapped when they shared given functions and varied when they were used according to different “geometric roles”: for example, Sullivan examined the terms *okolo*, *vozle*, *bliz*, *vblizi*, *podle* and *u* and found that it was possible to assign a label to each preposition, which identified the semantic role played by the referent of that term in sentences; among the geometric roles she used there were ‘latus’, which expressed the meaning of side, ‘proximate’, which referred to space proximity, the non-specific role of ‘part’ and others. The results showed that the main semantic overlaps involved the roles of ‘locus’ and ‘proximate’, shared by all the terms, and that when a role was added to the meaning of the word, the preposition changed. Finally, despite the partial overlap given by locus and proximate, Sullivan concluded that the prepositions are not fully synonymous.

De Jonge (1993) reported part of the study carried out by Lindvall (1986) on two Italian near-synonyms *parere* and *sembrare* ‘to appear, to seem’. For his research Lindvall used a relatively small corpus of 14 detective stories translated from English, and discovered that the two verbs are almost equally frequent, with *sembrare* focusing

on the entity involved, and *parere* on the activity of the entity. However, De Jonge claimed that “the illustrative examples cited by Lindvall generally lack context”<sup>10</sup>.

Niepokuj (1997) made a comparison of some Proto-Indo-European verbs of cutting by differentiating their meaning and looking for some common semantic values among them. In her paper, Niepokuj described the transitive action of cutting made by an agent with the use of an instrument and with some important effects on a patient. Each term taken into consideration by the author communicated specific notions. For example, a verb could refer to cutting limited to a particular time of the year (e.g. the Greek word *ámētos* was reminiscent of the action of reaping in the harvest-time) or could specify the size of the part cut (e.g. in the Old Norse language the word *klauf* referred to a hoof divided in two equal parts). Some verbs of cutting referred to destruction, especially when they denoted an action done with force and absence of control. Hence, what Niepokuj tried to highlight was the fact that it is difficult to determine synonymy among different verbs, since they can convey the same core meaning, but focus on different aspect of the event they encode, such as the instrument of the action or the patient affected by it.

Tsuji and Kageura (1999) compared 2,000 pairs of Japanese medical synonyms for various diseases, and showed how a language changes and establishes one dominant term, which then becomes more stable over time with respect to its synonyms. With their analysis, Tsuji and Kageura tried to identify what types of terms were more likely to become established main terms and which ones synonymous terms. In their analysis, they noticed that they had to take into consideration words as compounds with a word expressing the part of the body infected and another one for the figurative meaning. In order to outline the structural properties of terms, the authors divided words into grammatical and semantic categories: in the grammatical categories they first included constituent units that can be considered words by themselves (a verb and a noun which function as a verb; adjective and noun which function as an adjective and noun on their own) and secondly, units that cannot be considered words by themselves (a prefix, a suffix and others); in the semantic categories they made a distinction between conceptual categories (represented by individual units independent of their use, which means that these words convey the most important element to identify the disease) and the functional role of the units within the terms (that is the morphology of the disease

---

<sup>10</sup> With reference to this work, I could only partially use it because some lines were unreadable.

used figuratively, for example, by using the name of the person who made the disease known). The results of their research led to the conclusion that among synonyms in disease names, speakers usually preferred to employ disease terms which conveyed transparent information, namely which describes or identify the disease itself rather than its discoverer. Moreover, Tsuji and Kageura found that the main terms were mainly of Western-language origin, while synonyms terms were taken from Chinese.

The above studies show that (near-) synonymy is a cross-linguistic and widespread phenomenon. There may exist even six or ten words to say “the same thing”, but they may differ for some semantic nuances of meaning such that speakers use them freely and interchangeably.

## **2.6 Conclusion**

To sum up, in this chapter I have outlined some linguists’ descriptions of synonymy and near-synonymy, and some case studies of (near-)synonyms in the English language and in other languages. As we have seen, in order to carry out their studies, these scholars compared definitions of near-synonyms in dictionaries, made large use of corpora and also considered the terms as they normally appeared in conversations, newspapers and books, dealing with synonyms in their context of production/use. Although the way in which they carried out their studies varies, there is a common theme that keeps their ideas together. For all of them the influence of the context is a relevant aspect to be taken into consideration when analysing synonym pairs. Their findings generally point in the same direction, namely that absolute synonymy is rare and it would be a mistake to say that two synonyms are completely equivalent from every point of view.

## CHAPTER 3

### METHOD

#### 3.1 Introduction

In this chapter I will first explain the reasons that led me to choose some specific (near-)synonyms as the object of study for this dissertation, namely the nouns *murderer*, *killer* and *assassin*; the verbs *to disturb* and *to bother*; the adjectives *compulsory*, *obligatory* and *mandatory*; and the adverbs *maybe*, *perhaps* and *possibly*. Then, I will outline my research method, which comprised two phases: (1) a survey of the information provided about the above terms in mono- and bi-lingual dictionaries, and (2) the collection and analysis of corpus concordances of the above terms.

#### 3.2 Selection of near-synonyms

There are three main reasons why I chose to analyse the sets of near-synonyms mentioned in the introduction. First of all, as far as I know, no publications have taken into consideration the terms I will focus on. Therefore, it is useful to try and determine to what extent the terms can be used interchangeably in the same or similar contexts. Moreover, these specific words do not belong to particular disciplinary fields, and are rather part of everyday language. Therefore, research findings on them can be useful to all language users, rather than only practitioners in given fields, and particularly useful to foreign language learners, who are likely to encounter these terms or be required to use them, on multiple occasions. Finally, on a personal level, I have often run into these pairs or sets of (near-)synonyms without knowing in what respects, if any, they differed from one another, or what determined the choice of one rather than the other in given texts, and thus without knowing how to correctly use them. Through this study, therefore, I hope to be able to find a partial explanation to my questions so as to satisfy my curiosity as a learner of English as a foreign language.

#### 3.3 Data collection

The first phase of the data collection aimed at raising my own awareness of what is known, or claimed to be known, about the above-mentioned terms and to systematically organize this type of knowledge. Given that no lexicological studies are available on those terms, the only sources of information I could consult were



dictionaries. The second phase of the data collection aimed at exploring the usage of the above terms in texts, and more specifically, the lexico-grammatical and semantic patterns associated with these near-synonyms, so as to check the validity of, and maybe complement, the information retrievable from the dictionaries I consulted. To this end, I consulted large-scale corpora, comprising both transcripts of spontaneously produced speech and spontaneously produced written texts.

### 3.3.1 Dictionary data

In the first phase of my research, I decided to look up the definitions of the terms under study as they are found in monolingual, bilingual and synonym dictionaries. The twofold goal was to become aware of the range of meanings that these terms are said to have, and then, by comparing such definitions with one another, to identify semantic similarities and differences between the same terms. That is, I wanted to draw a preliminary semantic profile of the terms in question on the basis of readily available lexicographic information.

The dictionaries I consulted were chosen on the basis of their ease of accessibility both on the Internet and at my University library. They comprise English monolingual dictionaries, Italian-English/English-Italian bilingual dictionaries and dictionaries of English synonyms. In particular, I took into consideration an accredited dictionary reporting information on the etymology and first recorded uses of words, namely the *Oxford English Dictionary (OED)*; (<http://www.oed.com/>) and another etymological dictionary, namely *The American Heritage Dictionary of Indo-European Roots* (2000), which glosses the roots of words of Indo-European origin. I also looked up a few monolingual English dictionaries: the *Macmillan Dictionary online* (*MacMillan*, <http://www.macmillandictionary.com/>), the *Oxford Dictionaries online* (*Oxford Dict.*, <http://oxforddictionaries.com/>), the *Longman Dictionary of Contemporary English online* (*Longman*, <http://www.ldoceonline.com/>), the *Longman Language Activator* (*Longman-Activator*, 1993), the *Merriam Webster Dictionary online* (*Merriam-Webster*, <http://www.merriam-webster.com/>), the *Collins English Dictionary online* (*Collins*, <http://www.collinsdictionary.com/>) and the *Collins Cobuild Advanced Learner's English Dictionary* (*Cobuild*, 2003). I also examined the following bilingual dictionaries: the *Grande dizionario Hazon di inglese: italiano-inglese, inglese-italiano* (*Garzanti*, 2010), the *Dizionario Hoepli Inglese: inglese-italiano, italiano-inglese* (*Hoepli*, 2010), and the *Zanichelli: Il Ragazzini 2013 (Il Ragazzini*,

<http://dizionarioonline.zanichelli.it/dizionariOnline/#ragazzini>). Finally, I looked up the following dictionaries of synonyms: the *Webster's New World Dictionary of Synonyms* (*Webster's Dict.*, 1984), the *Cassell's Modern Guide to Synonyms & Related Words* (*Cassell's Dict.*, 1971), the *Thesaurus.com* (<http://thesaurus.com>) and *A Dictionary of English Synonyms and Synonymous Expressions* (ADESSE, 1938).

After selecting the dictionaries to consult, I proceeded to retrieve from them four main types of information about the terms considered, namely their etymology, their core meanings in English, their translation equivalents in Italian and their (near-) synonyms.

Therefore, to learn about the etymology of the terms I consulted the *OED* and *The American Heritage Dictionary of Indo-European Roots* (2000); the former is a historical dictionary that quotes the first attested uses of terms and also describes their semantic evolution over time and the latter traces the origins of Indo-European words. Then, to identify the current meanings of the terms I consulted Present-Day English dictionaries which are mainly addressed to foreign language learners. To determine the Italian translation equivalents of the terms under study, I consulted some Italian-English/English Italian dictionaries and compared the Italian translation equivalents given there with the meanings given in the Present-Day definitions previously gathered. Finally, to compile a list of the (near-)synonyms that the terms have, I consulted some synonym dictionaries so as to detect the nuances of meaning, if any, that make these terms differ from the other near-synonyms.

### **3.3.2 Corpus data**

In the second phase of my research, I adopted a corpus<sup>11</sup>-driven approach. In particular, I consulted a few large electronic corpora so as to retrieve instances of the actual use of the terms under study, and then identify their frequency of occurrence, their preferred combinatorial patterns and their register/genre preferences, if any.

The corpora I consulted represent the British and American varieties of English, and are accessible on the Internet. In particular, I collected data from four main corpora: the *British National Corpus (BNC)*, which is made up of 100 million words collected from written and spoken samples in the British English language from 1980s to 1993

---

<sup>11</sup> The term *corpus* (plural *corpora*) refers to an electronic database of authentic language data, that is a collection of written and spoken texts, gathered according to a pre-specified set of criteria so as to represent one or more language varieties, registers or genres.

and which I consulted both through its main website and via the *Sketch Engine* platform (*BNC*: <http://corpus.byu.edu/bnc/>; *Sketch Engine*: <http://www.sketchengine.co.uk/>); the *Corpus of Contemporary American English*, which is a monitor corpus made up of 450 million words collected from written and spoken American English texts from 1990 to the present (20 million words are added each year) and which I consulted through the interface available on the website (*COCA*: <http://corpus.byu.edu/coca/>); and the *Collins Wordbanks online*, which is a 455-million word database of mostly British, but also American English, containing both written and spoken material, mostly from the early 2000's. I consulted the *Collins Wordbanks online* through the Collins website (*CWo*: <http://wordbanks.harpercollins.co.uk/auth/>).

Through the consultation of corpora I wanted to retrieve information about the terms under study which was not likely to be available in the dictionaries, namely their frequency of occurrence in written and spoken contexts, the frequency of occurrence of their various word-forms, and also their collocational and colligational patterns.

To find out about the frequency of occurrence of the terms under study, I consulted the *BNC* and the *COCA* websites. Their built-in software interface easily allows the interested user to type in a given term, and to see, by means of automatically generated tables and charts, its overall frequency of occurrence – both in number of occurrences and normalized per million words (pmw) – and its distribution across different registers/genres and in its various word-form.

Then I looked at the collocational and colligational patterns of the terms under study, that is the lexical and grammatical relations between the terms and their neighbouring words. In particular, I consulted the *BNC* through *Sketch Engine* and the *CWo*, which were the only corpora that let me save the data retrieved; I randomly selected and saved 200 concordances<sup>12</sup> of each term under study from the *BNC* and 200 more from the *CWo* (in the text I will refer to them as *my datasets*): the 200 concordances of each verb randomly represented all the word-forms of the verb in question, while the 200 concordances of each noun comprises 100 instances of its singular and 100 instances of its plural word-form. Each list of concordances was sorted alphabetically once by the word to the left and once by the word to the right of the term in question in order to analyse the immediate phraseological and semantic patterns of use of the terms under investigation.

---

<sup>12</sup> A concordance is a line of text showing the occurrence of a term in a corpus, presented within the immediate co-text that it occurs in.

Finally, I compared the definitions given in the dictionaries with the findings that emerged from the corpus analysis in terms of collocational and colligational patterns so as to verify if and to what extent the former accurately represent the actual use of the terms as attested in spoken and written text production.

### **3.4 Conclusion**

To sum up, in this chapter I have described the way in which I carried out my research on the terms chosen as the object of study. In particular, I looked up their definitions in monolingual, bilingual and synonym dictionaries and then, after having outlined the main features common to each sets of words, I examined the frequency occurrences and the lexico-grammatical patterns that the words had in the texts in which they occur, by collecting and examining corpus data. In the following chapters, I will present the findings for each term, as collected through the examination of dictionaries (chapter 4) and of corpus data (chapter 5).

## CHAPTER 4

### FINDINGS FROM LEXICOGRAPHIC SOURCES

#### 4.1 Introduction

This chapter will report on the findings of the consultation of both general dictionaries and dictionaries of synonyms, for the purpose of collecting some preliminary information on the sets of English near-synonyms chosen as the object of this study: the nouns *murderer*, *killer* and *assassin*; the verbs *to disturb* and *to bother*; the adjectives *compulsory*, *obligatory* and *mandatory*; and the adverbs *maybe*, *perhaps* and *possibly*. In particular, the focus will be on a comparison of monolingual, bilingual and synonym dictionaries' definitions of those terms in order to give an overview of what is known of their meanings and to identify, if possible, their distinctive senses and semantic nuances.

#### 4.2 Dictionary data

In the following sections I will introduce and examine the definitions and part of the examples of the above sets of synonyms as found in the relevant entries of monolingual, bilingual and synonym dictionaries. The goals are to highlight the extent to which their meanings appear to overlap and to provide a preliminary illustration of the ways in which they are used. First, I will give a brief etymological overview of the terms, as retrievable from the *OED*, and then, I will overview their definitions, their Italian equivalents, and the near-synonyms given in the monolingual, bilingual and synonym dictionaries.

##### 4.2.1 The nouns

In the following sub-sections I will introduce and give some dictionary definitions of the nouns *assassin*, *murderer* and *killer*.

##### 4.2.1.1 *Assassin*

*Assassin* is not a word of Indo-European origin, and so there is no mention of it in *The American Heritage Dictionary of Indo-European Roots* (2000). Etymological information about it is found the *OED*, which says that the word *assassin* comes from the Arabic language (i.e. *haššāšīn* and *haššīšīūn*, the plurals of *haššāš* and *haššīš*), in

which it indicated a hashish-eater and was applied to the Ismaili sectarians, “who used to intoxicate themselves with hashish or hemp, when preparing to dispatch some king or public man”.

Echoes of the original meaning of the term are found in the definitions of the *OED*, the *Oxford Dict.* and the *Merriam-Webster*. These dictionaries explicitly refer to the past uses of the term and point out, in particular, that the word *assassin* was in use in the time of the Crusades, when, after having used hashish, certain Muslim fanatic groups were sent by their sheikh (the “Old Man of the Mountains”) on suicidal missions to murder the Christian leaders. The *OED* also reports that the term “retains so much of its original application as to be used chiefly of the murderer of a public personage, who is generally hired or devoted to the deed, and aims purely at the death of his victim”; the example it gives is a quote of the British historian and politician T.B. Macaulay (1855): “Barclay's assassins were hunted like wolves by the whole population”.

The monolingual dictionaries give similar definitions of the term *assassin*. Here, *assassin* is defined as someone who murders a politically important person, and relevant examples are provided (e.g. “Kennedy’s assassin is assumed to have been Lee Harvey Oswald” *Longman*; “John Wilkes Booth was the assassin of Abraham Lincoln” *Merriam-Webster*). Only the *Longman-Activator* (1993) also defines an assassin in more general terms, that is, as someone paid to kill someone, in line with the following example: “Although the assassins were never caught, it is commonly believed that they were working for the government” ( p. 725).

Some of the dictionaries consulted define the verb *to assassinate*<sup>13</sup> rather than the noun *assassin*. For instance, the entry for *assassin* given in the *Cobuild* (2003) only specifies that the term derives from the verb *to assassinate*, whose referent is an agent (“a person who assassinates someone”). The entry for *to assassinate* specifies that the verb is used with reference to someone important being killed or to killing as a political act. Similarly, the *Webster’s Dict.* (1984) specifies that the verb *to assassinate* “implies the sudden killing of a politically important person by someone hired or delegated to do this”, and the *Cassell’s Dict.* (1971) provides a similar definition of *to assassinate*, which includes two of its near-synonyms, namely *to murder* and *to kill*: “a specific form of *murder* in which someone *kills* a public figure, usually a political leader, for whatever reason” (e.g. “a televised debate on whether there had been a conspiracy *to*

---

<sup>13</sup> In some dictionaries, the entry is the bare infinitive *assassinate*, but I will use the *to*-infinitive throughout, independently of the specific choices of the various dictionaries.

*assassinate* President Kennedy”; p. 314). Finally, the bilingual dictionary *Il Ragazzini* (2013) also specifies that both the noun *assassin* and the verb *to assassinate* are used to refer to the assassination of an important person for political or religious reasons (e.g. “[t]he attempted assassination of the Pope: il tentato assassinio del Papa”).

All the bilingual dictionaries consulted associate the word *assassin* with the Italian translation equivalent *assassin*, to which the *Garzanti* (2010) adds *sicario*. Apart from the *Garzanti*, the other bilingual dictionaries also give examples of the translation equivalent mentioned: the *Hoepli* (2010) focuses on the hiring of an assassin in the sentence “[t]he assassin was a hired killer: l’assassino era un killer a pagamento” (p. 31) and the *Il Ragazzini* writes “Benazir Bhutto has been assassinated”, in line with the meaning of the verb *to assassinate*.

All the dictionaries of synonyms consulted list the following as the synonyms of *assassin*: *killer*, *murderer* and *slayer*. Some provide additional synonyms: the *ADESSE* (1938) gives synonyms such as *bravo*, which indicates a hired assassin, and *thug*, which is also used to indicate a criminal, a brutal ruffian or a thief. The *Thesaurus.com* also adds *butcher*, *clipper*, *dropper*, *eliminator*, *enforcer*, *guerrilla*, *gun person*, *hatchet person*, *hit person*, *liquidator*, *piece person*, *plugger*, *soldier*, *torpedo* and *trigger person*.

Other synonyms of the word *assassin*, which were not included in the synonym dictionaries, are given in the monolingual dictionaries. In there we can find words such as *death squad*, *lynch mob*, *hit squad*, *serial killer*, *hired gun* (*MacMillan*), *homicide*, *manslayer* (*Merriam-Webster*) and *executioner* (*Collins*). Within the above mentioned terms, only the words *killer*, *murderer*, *homicide*, *slayer* and *manslayer* encode a general meaning, comparable to that of *assassin*; the other near-synonyms listed can be applied to more specific fields (such as the military field); also, not only may they indicate specific types of killers, but they may also specify different ways of killing (e.g. *hit person*) or refer to a group of people rather than to a single individual (e.g. *death squad*). Therefore, these near-synonyms may be considered as hyponyms of the term *assassin*; for example, the word *guerrilla* is used for “a member of a military group that is not official and usually wants to change a political situation (...) [by using] unexpected attacks in small groups” (*MacMillan*).

As we have seen, even if not all dictionaries give exactly the same definitions or translation equivalents of the word *assassin*, overall, they all appear to describe this term as one that identifies a person who commits a violent attack against an important

political person, which leads to that person's death, and the most common Italian translation equivalent mentioned is the word *assassino*. As regards synonyms, the ones which are shared by both the synonym dictionaries and monolingual dictionaries are *killer*, *murderer* and *(man)slayer*.

#### 4.2.1.2 Murderer

*The American Heritage Dictionary of Indo-European Roots* reveals that the word *murderer* derives from the root \**mer-* (with the meaning of 'to die'), which can be found in words somehow referring to death and to human beings as mortal. The *OED* writes that the term dates back to the 12<sup>th</sup> century, when it was recorded in the Anglo-Norman variant forms *murdreour*, *murdrer*, *murdreere*, *murthrur*, *mourdrer*, *mourdreere* and in Old French as *mordreur*<sup>14</sup>.

Only the *OED* gives some obsolete meanings of *murderer*: (1) "[a] traitor"; (2) "[g]unnery. A small cannon or mortar; (later) esp. such a weapon used on a warship to repel enemies attempting to board"; (3) "[a] dagger, a knife"; (4) "[h]airdressing. A kind of knot used to gather curls"; (5) "[f]ishing. A device used for catching cod". Apart from hairdressing, the other definitions have some common traits, that is, they all refer to an object or an instrument mainly used as weapon for killing.

The meaning of the word *murderer* that is currently in common use, which refers to people's behaviour, is defined in similar ways by both the monolingual dictionaries and the *OED* (e.g. "a person who commits murders" *Oxford Dict.*; "someone who murders another person" *Longman*); relevant examples are: "[t]he *murderer* was sentenced to life in prison without the possibility of parole" (*Merriam-Webster*), "[o]ne of these men may have been the murderer" (*Cobuild*; p. 941). According to the *Longman-Activator*, the *Webster's Dict.* and the *Cassell's Dict.*, the action of murdering someone is done deliberately and especially after having planned it (e.g. "[h]e was put in a special high security jail for convicted murderers and other violent criminals"; *Longman-Activator*, p. 725). The *Cassell's Dict.* also adds that "[s]ometimes the word can refer to a brutal *killing*, as in war" (e.g. "unprovoked aggression in which one nation set out to *murder* the citizens of an adjoining state"; p. 314). Finally, *murderer* can also

---

<sup>14</sup> No Present-Day English gloss for the variant forms is explicitly provided either in the *OED* or in *The American heritage dictionary of Indo-European roots* (2000).



be used as a hyperbole, to “point to the mishandling of anything”, as in the sentence: “expressionless actors who *murder* their lines” (p. 314).

The bilingual dictionaries give only two Italian equivalents for the word *murderer*, namely *assassino* and *omicida*, and the only example given is found in the *Hoeppli* dictionary (“[t]he murderer was electrocuted yesterday: ieri l’assassino è stato giustiziato sulla sedia elettrica”; *Hoeppli*, p. 450).

According to the synonym dictionaries consulted, synonyms for the word *murderer* are *assassin*, *butcher*, *slaughterer*, *slayer*, *cut-throat*, *killer* and to these, the *Thesaurus.com* adds *enforcer*, *hit person*, *hit-and-run*, *homicide*, *trigger person*, while the *ADESSE* also includes *blood-shedder*. Therefore, there appear to be many near-synonyms of *murderer*, some of which conveying meanings at the same level of generality as *murderer*, while the others are characterized by more specific semantic nuances which make them more suitable for given contexts: for example, *slaughterer* can be used to refer to someone who kills animals in addition to people, while *murderer* cannot be used when referring to the killing of animals; *slayer* refers to someone who kills someone else in a violent way, while *murderer* does not include in its meaning how violent the criminal may be in carrying out the action.

To sum up, 1) the monolingual dictionaries and the *OED* similarly define a murderer as someone who commits a previously planned act of killing; 2) the bilingual dictionaries give only two possible Italian equivalents of *murderer*, namely *assassino* and *omicida*; 3) all the dictionaries list the words *assassin* and *killer* as synonyms of *murderer*.

#### **4.2.1.3 Killer**

*The American Heritage Dictionary of Indo-European Roots* states that the root of the word *kill*, which the noun *killer* derives from, is *g<sup>w</sup>elə-*, which conveys the general meaning of ‘piercing’. In particular, when the root is suffixed in the form *g<sup>w</sup>l-yo*, it gives origin to the Middle English verb *killen*, which probably derives from the Old English *\*cyllan*, and which meant ‘to strike, to beat’. Also, according to the *OED*, the word *kill* seems to derive from the German word *quälen* (whose root is *\*kuljan*, very similar to *\*cyllan*), meaning ‘to be cruel to/torture/torment/plague somebody’.

The *OED* also generally defines the noun *killer* as someone who or something that kills. Similarly, the *Cobuild* reports that *killer* can be generally used when referring “to

something that causes death or is likely to cause death as a killer” (e.g. “heart disease is the biggest killer of men in most developed countries”; p. 791).

The monolingual dictionaries also provide more specific definitions of *killer*. For instance, the *OED* and the *Collins* write that the word is used when a cow, sheep or another animal is killed for food, specifying that this meaning is mainly used in colloquial Australian and New Zealand Englishes. Also, the *OED*, the *MacMillan*, the *Oxford Dict.*, the *Longman* and the *Collins* specify that the term *killer* lists a meaning which is typical of informal language, that is ‘a formidable impressive, taxing, exhausting, boring or difficult thing/activity or person’ (e.g. “his new novel is a killer” *Oxford Dict.*) and it is frequently applied to popular music with the meaning of “very effective; excellent, sensational”. Finally, the *OED* gives additional, specific definitions which are not given in the other dictionaries: therein we find that *killer* can indicate “[a] name of the grampus, *Orcinus orca* [namely the killer whale], and other ferocious cetaceans of kindred genera” and that it can be used figuratively, although the dictionary does not give any relevant example of the term used in such way. In addition, the *OED* also uses *killer* to refer to something used to kill animals such as “[a]n effective angler’s bait”; “a. A club of hard wood for killing fish with; b. A contrivance for killing a large ferocious animal (e.g. a wolf, a shark); also, an explosive implement for the painless killing of cattle, horses, etc.”; other definitions describe the term as a neutralizing agent (e.g. “[a]n agent used to neutralize the active property of anything, to neutralize a colour, to remove spots or stains, prevent pitch-stains on pine-boards, or the like”).

Moreover, as in the cases of *assassin* and *murderer*, the *Webster’s Dict.* and the *Cassell’s Dict.* give some descriptions for the verb *to kill* rather than for the noun *killer*. In particular, the former indicates that *to kill* means “to cause the death of in any way and may be applied to persons, animals or plants”, while the latter reports that among the words referring to the taking of lives, “*kill* is the most general word, applying to any kind of death-dealing activity” (e.g. “a drought that *killed* our fruit trees; an insecticide to *kill* cockroaches; two people *killed* in a car accident; a madman who threatened to *kill* me; soldiers *killed* in action”; p. 314). Finally, the *Cassell’s Dict.* also indicates that the verb can be used “where no life is actually lost” (e.g. “a veto that *killed* the revolution”; “their decision to *kill* the news story after it has appeared in the early edition”; p. 314).

The main Italian translation equivalents for the term *killer* given by the bilingual dictionaries are *assassino*, *omicida* and *sicario*. The *Il Ragazzini* gives ten glosses of the

term *killer* used as a noun rather than as an attributive adjective: (1) “uccisore; omicida; assassino”; (2) “sicario, killer”; (3) “animale che uccide”; (4) “cosa, malattia, ecc., che uccide; causa di morte: *Measles used to be a big killer*: il morbillo un tempo era spesso causa di morte”; (5) “(fam.) cosa (o esperienza, ecc.) assai faticosa; faticata”; (6) “(fam.) cosa o persona straordinaria; cannonata; schianto; fine del mondo; figata”; (7) “(fam.) barzelletta, battuta, ecc., assai divertente”; (8) “(USA) (timbro di) annullo postale”; (9) “(pallavolo) schiacciatore”; (10) “(Austral.) bestia da macello”. Therefore, from these glosses, it appears that this dictionary considers more senses of the word *killer* than the monolingual dictionaries, since there are some translation equivalents that do not correspond to any of the definitions given in the monolingual dictionaries, such as numbers 8 and 9.

Among the synonym dictionaries consulted, only the *Thesaurus.com* gives some synonyms for the term *killer*, namely *assassin*, *cut-throat*, *exterminator*, *gunman/woman*, *gunperson*, *hit-man/woman*, *hunter*, but also *butcher*, *executioner*, *hit person*, *slayer* and *soldier*, which are shared with the nouns *assassin* and *murderer* (exceptions are *exterminator* and *gunperson*, which are only given as synonyms of the word *killer*). To these synonyms, it is possible to add *slaughterer*, *liquidator*, *terminator* and *genocidaire*, which are found in the *Collins*. As we can see, the near-synonyms listed only identify individuals rather than groups of people, unlike some of the synonyms given for *assassin*. In addition, some of them (e.g. *hunter*, *butcher* and *liquidator*) are more specific in meaning than *killer* and also have meanings that are not shared with *killer*: for example a hunter is “someone that hunts”; a butcher is “a person whose trade is cutting up and selling meat in a shop” and a liquidator is “a person appointed to wind up the affairs of a company or firm” (*Oxford Dict.* definitions).

In conclusion, on the basis of the dictionaries consulted, the term *killer* appears to define a person, an animal or thing that kills, and its Italian corresponding words are *assassino*, *omicida* and *sicario*. *Killer* can also be used to specify a quality of something and, in particular, something out of the ordinary because of its extremely positive or extremely negative qualities (e.g. an amazing form of entertainment, a very funny joke or a very boring, difficult and exhausting task). The synonym dictionaries list the word *killer* as a near-synonym of *assassin* and *murderer* and to this, they add some other synonyms: *executioner*, *hit person*, *soldier*, *butcher*, *cut-throat* and *slayer*. Such synonyms share some characteristics with the terms under focus in this study: they stand for someone who can kill or hit another person or animal in some specific ways. But

they also have some distinctive features: they can specifically refer to persons who have to carry out a task because of their job: for example, a butcher slaughters animals to sell them as food, while a hunter kills animals to get their skin for clothing or their meat for food, and so on.

#### 4.2.1.4 Conclusion

The nouns *assassin*, *murderer* and *killer* are generally used to identify someone who kills another person. In particular, *assassin* makes reference to a violent act committed against a politically important person and, while *assassin* and *murderer* cannot be used when the act of killing is committed against an animal, *killer* can. The Italian term *assassino* is used as a translation equivalent for all these nouns; *omicida* is given only for *murderer* and *killer*, and finally *sicario* is given as the equivalent of *killer* and *assassin* (this last equivalent of *assassin* is given only by the *Garzanti*). From the analysis of the dictionaries of synonyms, it appears that *assassin*, *murderer* and *killer* are near-synonyms of one another and also that other near-synonyms such as *slayer* and *butcher* are associated to the three nouns.

In conclusion, the data retrieved from the consultation of dictionaries in the dictionary consultation appear to show that *assassin*, *murderer* and *killer* are not freely interchangeable because, despite the fact that the three of them denote people that kill and that they have the same Italian translation equivalent (*assassino*), they also have other nuances of meaning that render them suitable to specific contexts, not shared with the others. In particular, only *assassin* is used when the victim is an important political figure, and only *killer* can be used when animal are killed. These findings are summarised in Table 4.1.

Table 4.1 *Assassin, Murderer, Killer*

Word class: nouns				
		<i>ASSASSIN</i>	<i>MUDERER</i>	<i>KILLER</i>
<b>Meanings</b>	Act of killing someone	+	+	+
	To kill a political person	+	-	-
	To kill animals	-	-	+
	Out of the ordinary event	-	-	+
<b>Italian equivalents</b>	<i>assassino</i>	+	+	+
	<i>omicida</i>	-	+	+
	<i>sicario</i>	+	-	+
	<i>cannonata</i>	-	-	+
<b>Synonyms</b>	<i>Slayer, butcher</i>	+	+	+
	<i>Bravo, thug</i>	+	-	-
	<i>Slaughterer</i>	-	+	+
	<i>Hunter, liquidator</i>	-	-	+

(The plus and minus signs indicate presence vs absence, respectively, of a given feature.)

#### 4.2.2 The verbs

In the following sub-sections, I will introduce and give dictionary definitions of the verbs *to bother* and *to disturb*.

##### 4.2.2.1 *To disturb*

*The American Heritage Dictionary of Indo-European Roots* indicates that the root of the verb *to disturb* is *(s)twēr-*, which appears as a zero-grade affix in the form *\*turbā-*, meaning ‘tumult’. According to the *OED*, the etymology of *disturb* goes back to the Middle English forms *destorben*, *destourben*, with cognate forms attested in Old French *destourber* and in Old Spanish *destorbar*. These go back to Latin *disturbare*, which derives from *dis-* ‘utterly’ + *turbare* ‘disturb’.

In the *OED*, the verb *to disturb* is said to mean “a. to agitate [mentally] and destroy (quiet, peace, rest); to break up the quiet, tranquillity, [calmness] or rest of (a

person, a country, [mind,] etc.); to stir up, trouble, disquiet” and “b. to throw into a state of physical agitation, commotion or disorder”. The *OED* also defines an additional sense of the verb as “[t]o interfere with the settled course or operation of; to put out of its course; to interrupt, derange, hinder, frustrate” and it adds definitions of by now obsolete meanings: “a. to deprive of; to drive, turn, or draw away from, by disturbance; b. Law. To deprive of the peaceful enjoyment or possession of”.

When referring to the present-day meaning of *to disturb*, the monolingual dictionaries give definitions which can be divided into a few groups. The first group of meanings refers to the act of making something move so that it changes its position or shape (e.g. “The items on her desk had been *disturbed*”; *Merriam-Webster*; “A soft breeze gently disturbed the surface of the pool”; *MacMillan*). The second group of meanings mentions doing something that destroys or interrupts the quiet or peace of a place or a situation (e.g. “Not even a breath of wind disturbed the beautiful scene”; *MacMillan*; “The noisy lawnmower *disturbed* their sleep”; *Collins*). The third group of meanings refers to the action of interrupting someone so that they cannot continue doing what they were doing (e.g. “The thieves fled when they were disturbed by a neighbor”; *Longman*; “I hope I’m not disturbing you”; *Cobuild*, p. 411). The fourth group of meanings explains disturbing as an action that makes someone feel upset or worried (e.g. “Ministers declared themselves profoundly disturbed by the violence”; *MacMillan*; “I am disturbed by the document I have just read”; *Oxford Dict.*). To these, only the *MacMillan* adds another meaning which refers specifically to animals: “to frighten wild animals or birds so that they run away”, and the *Longman* adds the following definition: “to change a normal situation in a way that causes problems” (e.g. “[m]y hormone balance is disturbed by my pregnancy”).

Apart from giving synonyms of the verb *to disturb*, the *Webster’s Dict.* and the *Cassell’s Dict.* also define the verb: the former highlights that *to disturb* “implies the unsettling of normal mental calm or powers of concentration by worry, interruption, etc.” (e.g. “to disturb one’s train of thought”; p. 71), and the latter describes *to disturb* as a verb being more intense than its synonyms (*to harass*, *to pester*, *to plague*, *to trouble*, *to worry*) when “suggesting specifically, at its most extreme, mental derangement” (e.g. “the mentally disturbed delinquent”; p. 57). As we can see, these descriptions of the term *to disturb* overlap with the meanings given in the monolingual dictionaries; in particular, they refer to the acts of interrupting someone and creating agitation or mental insanity.

The most common translation equivalent of the verb *to disturb* given by the bilingual dictionaries is *disturbare*. The dictionaries also give other translation equivalents, which appear to match the senses described in the definitions given by the monolingual dictionaries.

Two translation equivalents are *muovere* and *spostare*, which refer to the first group of meanings in the monolingual dictionaries: ‘make something move’ (e.g. “The footprints in the sand had not been disturbed: le impronte sulla sabbia erano intatte”; *Hoeppli* p.194). An additional verb listed is *turbare*: ‘to destroy or interrupt the quietness’ (e.g. “To disturb the peace: turbare l’ordine pubblico, la quiete pubblica”; *Garzanti*, p. 352). We also have *interrompere*, which corresponds to the third group of meanings I found in the monolingual dictionaries: ‘interrupting someone so that they cannot continue what they were doing’ (e.g. “They disturbed a burglar, who fled through the back door: hanno disturbato un ladro, che è scappato dalla porta sul retro” *Il Ragazzini*). Another one is *preoccupare*: ‘make someone feel upset or worried’ (e.g. “Recent political developments have disturbed us: i recenti sviluppi politici ci hanno preoccupati”; *Hoeppli*, p. 194). To all these possible translation equivalents, only *Il Ragazzini* adds the gloss ‘far paura a (un animale)’, which clearly echoes the definition of the *MacMillan* on the annoyance given to animals (e.g. “Do not disturb nesting birds: non disturbare gli uccelli che covano”).

The (near-)synonyms of *to disturb* are given by both the synonym dictionaries and the monolingual dictionaries. In particular, the synonyms of the verb *to disturb* retrieved in the *Cassell’s Dict.* are *to harass*, *to pester*, *to plague*, *to trouble*, *to worry*, while the *ADESSE* provides a list of sets of near-synonyms without, however, explaining the differences that characterise each set of synonyms: “(1) [*To*] *agitate*, [*to*] *shake*, [*to*] *stir*; (2) [*To*] *disarrange*, [*to*] *derange*, *disorder*, [*to*] *confuse*, [*to*] *upset*, [*to*] *unsettle*, [*to*] *throw into confusion*, [*to*] *put into disorder*; (3) [*To*] *molest*, [*to*] *annoy*, [*to*] *disquiet*, [*to*] *distract*, [*to*] “fuss”, *discontent*, [*to*] *perturb*, [*to*] *discompose*, [*to*] *vex*, [*to*] *ruffle*, [*to*] *worry*, [*to*] *plague*, [*to*] *trouble*, [*to*] *incommode*; (4) [*To*] *interrupt*, [*to*] *impede*, [*to*] *hinder*”. The *Collins* gives [*to*] *interrupt*, [*to*] *trouble*, [*to*] *bother*, [*to*] *startle*, [*to*] *plague*, [*to*] *disrupt*, [*to*] *put out*, [*to*] *interfere with*, [*to*] *rouse*, [*to*] *hassle*, *inconvenience*, [*to*] *pester*, [*to*] *intrude on* and also [*to*] *upset*, [*to*] *concern*, [*to*] *worry*, [*to*] *shake*, [*to*] *excite*, [*to*] *alarm*, [*to*] *confuse*, [*to*] *unnerve*, [*to*] *flex*, [*to*] *fluster*, [*to*] *perturb*, [*to*] *derange*, [*to*] *unsettle*, [*to*] *agitate* and others. As we can see, the synonyms which are shared by all the dictionaries include *to plague*, *to worry*, *to upset*,

*to confuse* and *to trouble*. Such near-synonyms are described by the dictionaries as verbs which are used to indicate distress and the feeling of anxiety about actual or potential problems. Therefore, their meanings appears to overlap with those of the verb *to disturb*.

As we have seen, the verb *to disturb* is used to convey various meanings which are related to the change of mental state, position, shape or activity caused by a movement or an interruption. It is also possible to notice a correspondence between the definitions given by the monolingual dictionaries and the Italian translation equivalents retrieved in the bilingual dictionaries. Finally, we have seen that the synonym dictionaries give a variety of synonyms of *to disturb*; among these, the most common are *to harass*, *to plague* and *to trouble*.

#### **4.2.2.2 To bother**

*The American Heritage Dictionary of Indo-European Roots* (2000) does not give any entry for the root of the verb *to bother* and, according to the *OED*, the etymology of the verb has to be considered unknown, even if the dictionary suggests an Anglo-Irish origin probably because the earliest attested instances of the term are found in the writings of Irishmen (T. Sheridan, Swift, Sterne) and in the Irish comic vocabulary of fiction and the stage.

Besides giving the etymology, the *OED* also gives a meaning of the verb *to bother* which is regarded as no longer in use in the modern English language, namely, the action of “bewilder[ing] with noise, confus[ing], muddl[ing or] put[ing] into a fluster or flutter”.

With regard to the current meaning of the term, the monolingual dictionaries and the *OED* distinguish between its transitive and its intransitive function. In particular, the verb is said to be used transitively when conveying the following meanings: (1) ‘to annoy someone by interrupting them when they are busy or want to be left alone’ (e.g. “Danny, don’t bother Ellen while she’s reading”; *Longman*); (2) ‘to make someone feel worried, upset or concerned’ (e.g. “secrecy is an issue which bothers journalists”; *Oxford Dict.*); (3) ‘to stalk, to frighten someone by following them around or trying to talk to them when they do not want to’ (e.g. “If he keeps bothering you, you should call the police”; *MacMillan*); (4) ‘to cause someone physical pain’ (e.g. “my stomach is *bothering* me”; *Merriam-Webster*). When used intransitively, *to bother* is said to be used in negative sentences or questions to indicate that someone does not want to make



the effort to do something because there is no good reason, it is unnecessary to do it or because they are too lazy (e.g. “Nothing I do makes any difference anyway, so why bother?”; *Cobuild*, p.155). Another intransitive use of the verb is also attested in sentences in which *to bother* means “to take the time or trouble; concern oneself” (e.g. “don't bother to come with me”; *Collins*). The *Longman* and the *OED* add another meaning which is used to express an exclamation indicating a sudden feeling of annoyance about something (e.g. “Oh bother it! The thread’s broken again!”; *Longman*). The *OED* also adds that *to bother* can be used alone “[i]n the imperative (logically 3rd pers. sing. with implied subject after analogy of verbs of cursing) as a mild imprecation”; the use of *to bother* as an exclamation or a imprecation is not shared by the verb *to disturb*. Finally, the *OED* gives a definition in which the verb can be used both transitively and intransitively, namely “to blarney, to ‘humbug’”.

The dictionaries of synonyms, too, give some explanations on the use of *to bother*; in particular, according to the *Cassell’s Dict.*, the verb *to bother* is used when a person is actively annoying another or in a situation in which a person is upset by something or for “a minor complaint that may come and go” (e.g. “frequently bothered by a slight stiffness in his joints”; p. 58). In addition, the verb is said to be used to signal an action that may be done intentionally to disturb (e.g. “Just pay the bill and I’ll stop bothering you”; p. 58). To define *to disturb*, the *Webster’s Dict.* makes reference to the verb *to annoy*; in particular, in the entry of *to annoy*, *to bother* is given as one of its near-synonyms implying “minor disturbance of one’s peace of mind and (...) suggest[ing] mild perplexity or anxiety”.

The bilingual dictionaries also give different entries for the transitive and the intransitive meaning of the verb *to bother*. In particular, when the verb is used transitively, it is translated as *dare fastidio*, *incomodare*, *importunare*, *seccare* (e.g. “Don’t bother him when he’s working: non disturbarlo quando lavora”; *Hoepli*; p.70), whereas, when the verb is used intransitively and mainly in negative sentences, it is translated as *preoccuparsi*, *prendersi il disturbo di* (e.g. “Don’t bother fixing it: non disturbarti ad aggiustarlo”; *Il Ragazzini*). Both the *Hoepli* and the *Il Ragazzini* give *al diavolo!*, *maledetto!*, *uffa!* and *accidenti!* as the translation equivalents of the imperative of the verb used as an exclamation to express irritation for something.

According to *Thesaurus.com*, *to bother* can be replaced by the following constructions: “[to] be concerned about, [to] concern oneself, [to] exert oneself, [to] fuss over, [to] go out of one’s way, [to] make a fuss about, [to] make an effort,[to] put

*oneself out, [to] take pains,[to] try, [to] worry about*” and the *ADESSE* lists “[to] perplex, [to] worry, [to] harass, [to] trouble, [to] annoy, [to] tease, [to] vex, [to] plague, [to] molest, [to] incommode, [to] disturb,[to] pester, [to] bother” as near-synonyms of the verb. In addition to the above mentioned synonyms, the *Merriam-Webster* lists some synonyms which are not found in the synonym dictionaries, such as “[to] bug, [to] chivy and [to] intrude (upon)”. The *Collins* gives some synonyms which are also common to the verb *to disturb* such as “[to] trouble, [to] concern, [to] dismay, [to] upset, [to] alarm, [to] vex, [to] pester, [to] plague”, but it also adds “[to] irritate, [to] annoy, [to] harass, [to] molest, [to] faze and [to] put or get someone’s back up”, all of which are said to mean ‘interfering with the activity of someone and irritating them’.

In conclusion, *to bother* has various meanings such as ‘to interrupt someone, to make someone feel worried or upset, to frighten someone, to suffer for a physical pain and to concern oneself’. Moreover, the verb can also be used as an exclamation or an imprecation. The bilingual dictionaries give two main translation equivalents of the verb: *preoccuparsi* and *dare fastidio*. Finally, there are apparently many synonyms of the verb *to bother*: for example, *to harass, to trouble, to plague, to concern, to annoy, to molest, to worry* and *to make an effort*.

#### **4.2.2.3 Conclusion**

As we have seen, the verbs *to bother* and *to disturb* have some common characteristics: both are used to describe the act of making a person feel worried or upset, or interrupting someone engaged in some activity they were doing. Moreover, both verbs share synonyms such as *to harass, to plague* and *to trouble*. But the two verbs have also some different meanings: for example, while *to disturb* is always used as part of a sentence, *to bother* can stand alone as an exclamation or imprecation; while *to disturb* can indicate a movement that causes a change in the position of something, *to bother* does not convey this meaning; while *to disturb* can describe the act of someone who wants to frighten animals, *to bother* cannot; and, while *to disturb* is said to be used when someone annoys someone else by accident or unintentionally, *to bother* is said to be used when such annoyance is made on purpose. Finally, the bilingual dictionaries do not give exactly the same translation equivalents for the two verbs; in particular, *to disturb* is mainly translated as *disturbare*, while the main translation equivalents of *to bother* are *dare fastidio* and *preoccupare*.

In conclusion, as summarised in Table 4.2, the data retrieved in the dictionaries shows that *to disturb* and *to bother* are not freely interchangeable because, although both are used to indicate the act of interrupting someone or making feel worried or upset, each also conveys meanings not available to the other term, that is: “to move something or make it change its position”, “to frighten animals” and “to annoy unintentionally” are only activated for *to disturb*, while “to annoy deliberately” and “to stalk someone” are only activated for *to bother*.

Table 4.2 *To disturb, to bother*

Word class: verbs			
		<i>TO DISTURB</i>	<i>TO BOTHER</i>
Meanings	To make someone feel worried or upset	+	+
	To interrupt someone	+	+
	To move something or make it change its position	+	-
	To frighten animals	+	-
	To stalk someone	-	+
	To annoy unintentionally	+	-
	To annoy deliberately	-	+
Italian equivalents	<i>Disturbare</i>	+	+
	<i>Preoccupare</i>	+	+

(The plus and minus signs indicate presence vs absence, respectively, of a given feature.)

### 4.2.3 The adjectives

In the following sub-sections, I will introduce and give dictionary definitions of the adjectives *compulsory*, *obligatory* and *mandatory*.

#### 4.2.3.1 *Compulsory*

According to the *OED*, the adjective *compulsory* derives from Medieval Latin *compulsorius*, whose root *compuls-* means ‘driven, forced’ (from the verb *compellere*

meaning 'to compel'). In particular, in the early 16<sup>th</sup> century, the adjective began to be used as a noun "denoting a legal mandate which had to be obeyed" (*OED*).

The *OED* specifies that the adjective can indicate something "depending on or produced by compulsion; compelled, forced, enforced, obligatory" and also "involving or exercising compulsion; compelling, coercive"; on the other hand, when it refers to an agent, the adjective is said to indicate "acting under compulsion; compelled; involuntary". The dictionary also gives possible collocates of the adjective such as "education" and "games". An obsolete collocation is found in the *OED*, namely "compulsory letters", which were "letters issued to compel the production of documents or appearance of witnesses". Although indirectly, this example too evokes the role of an authority, whose role is explicitly mentioned in the definitions found in the other monolingual dictionaries.

According to the monolingual dictionaries, the word *compulsory* is used with something that must be done because of a law, a rule or someone in a position of authority orders to do it. Only the *Longman-Activator* indicates that *compulsory* is used to indicate something that is intended to be done to keep people safe, to improve their education and so on (e.g. "[i]n France seat belts are compulsory"; *Longman-Activator*, p. 874). More examples are given in the other monolingual dictionaries: "[s]chool uniform is no longer compulsory in many British schools" (*MacMillan*); "[i]n East Germany learning Russian was compulsory" (*Cobuild*, p. 283). The *Cassell's Dict.* underlines the role of authorities and points out that *compulsory* "suggests that someone in authority has imposed a course of action that may not be departed from" (e.g. "compulsory attendance at lectures"; p. 106). The synonym dictionary also indicates that the adjective may imply a punishment for those who do not respect the compulsory ruling, which can be enforced also with the use of coercion (e.g. "compulsory blacklisting"; p. 106).

The bilingual dictionaries give lists of translation equivalents for the adjective *compulsory* like *obbligatorio*, *coattivo*, *coatto*, *forzato*, *forzoso*, *coercitivo*, *cogente*, and the dictionaries also give relevant collocations with their corresponding Italian translation: "compulsory education: educazione obbligatoria"; "compulsory winding up: liquidazione coatta" (*Hoepli*; p. 132); "compulsory sale: vendita forzosa"; "compulsory saving: risparmio forzato" (*Il Ragazzini*); "compulsory loan: prestito forzoso" (*Garzanti*; p. 243).

Finally, among the most common synonyms of the word *compulsory*, we find *binding, de rigueur, imperative, mandatory, obligatory* (*Cassell's Dict.*), *forced, imperative, imperious, necessary, required, requisite* (*Thesaurus.com*), and also *unavoidable, enforced, not to be evaded, compelling, constraining* (*ADESSE*). The *MacMillan* also adds *essential, vital, basic, indispensable, without fail, at all costs, of necessity* and the *Merriam-Webster* also suggests *incumbent, peremptory* and *non-elective*. Overall, the synonyms which commonly appear to be related to the term *compulsory* are *binding, mandatory* and *obligatory*, all of which represent something that people must absolutely do, in opposition to words like *necessary* and *indispensable*, which indicate a milder type of obligation, or at least an obligation not necessarily coming from an authority.

As we have seen in the above definitions, the adjective *compulsory* indicates something which must be done because of a law or an order of an authority; its Italian equivalents are words like *obbligatorio, forzato* and also *coatto*. The synonym dictionaries consulted list a lot of synonyms for the term, the most common being *binding, obligatory, mandatory, imperative* and *de rigueur*.

#### 4.2.3.2 *Obligatory*

*The American Heritage Dictionary of Indo-European Roots* states that the root of the term *obligatory* is the lengthened-grade form *\*lēg* of short-form *leg-*, meaning ‘to collect, which shows up in the Latin word *lēs* meaning ‘law’. According to the etymological dictionary, the root *\*lēg* is also present in the Latin words *lēgāre*, which means ‘to depute, commission, charge’ and *obligāre*, which means ‘to oblige’. *The OED* also adds that the adjective was present in post-classical Latin (with the meaning of ‘constituting an obligation’, according to 13<sup>th</sup> century British sources) and in Anglo-Norman and Middle French *obligatoire* (13<sup>th</sup> century).

The adjective *obligatory* has one main meaning which is shared by all the monolingual dictionaries and the *OED*, namely indicating something which must be done or practiced in order to obey a law or rule (e.g. “[i]t is obligatory for companies to provide details of their industrial processes”; *Longman*). Furthermore, the dictionaries indicate that when the term goes before a noun, it is used humorously as a cliché and describes something which is always done or included on a particular occasion (e.g. “[t]here are always those girls who know that they look fantastic in everything... doing model poses in the mirror saying, ‘Does it make me look fat?’ to their obligatory obese

friend”). Moreover, the explanation of the term *obligatory* given by the *Cassell’s Dict.* highlights that it is “like *binding* in its legal sense”; an additional, general definition is given by the *Cassell’s Dict.*, and that is “suggest[ing] something expected or made necessary by morality or conscience, rather than by a higher authority”. In particular, *obligatory* is said to focus more on what should be done than what is or must be done (e.g. “notions of obligatory timidity she could not subscribe to”; p. 106).

Although the monolingual dictionaries give more than one definition of *obligatory*, the bilingual dictionaries give only one Italian translation equivalent for the adjective, that is *obbligatorio* (e.g. “[t]o make it obligatory (up) on s.o. to do sth: obbligare qualcuno a fare qualcosa”; *Garzanti*, p. 841); “[o]bligatory attendance: frequenza obbligatoria”; *Hoepli*, p. 469). In these sentences the term is used to indicate something that must be done even though a law or an authority is not mentioned.

The synonyms of *obligatory* found in the *ADESSE* include *binding* and *coercive*, to which *Thesaurus.com* adds *compulsory* and the same synonyms given for that adjective, namely *compulsatory*, *de rigueur*, *enforced*, *essential*, *imperative*, *imperious*, *mandatory*, *necessary*, *requisite*, *required* and *unavoidable*. The *Cassell’s Dict.* gives the same synonyms as the above-mentioned synonym dictionaries. The *MacMillan* lists synonyms specific to the two definitions of *obligatory* given therein: *necessary*, *essential*, *vital*, *basic*, *compulsory*, *indispensable*, *without fail*, *at all costs* and *of necessity*, which recall the obligation imposed by a law or rule, and *typical*, *traditional*, *usual*, *classic*, *characteristic*, *archetypal*, *prototypical*, *customary* and *essential*, which are related to the humorous meaning outlined above.

From the dictionaries consulted, it appears that the term *obligatory* has two main meanings: 1) something that must be done to obey a law or rule, whose synonyms are *binding*, *necessary*, *compulsory*, *imperative*, *mandatory*, *de rigueur* and others; 2) something that is done out of habit or as a customary practice, whose corresponding synonyms are *typical*, *traditional*, *classic*, *customary* and others. The bilingual dictionaries give only one possible Italian translation equivalent, that is *obbligatorio*.

#### **4.2.3.3 Mandatory**

*The American Heritage Dictionary of Indo-European Roots* reveals that the root of the word *mandatory* is *man-*, which means “hand”. In line with this, the *OED* indicates that the origin of the adjective *mandatory* goes back to the post-classical Latin

*mandatorius* meaning ‘one who orders a commission or initiates something’, which derives from Latin compound *mandāre* ‘to put into someone’s hand, entrust, order’.

Both the *OED* and the other monolingual dictionaries give a definition of the term *mandatory* which makes reference to a command, a commission or something which is ordered by law or rule (e.g. “A new accounting system will soon become mandatory for all departments”; *MacMillan*). An additional meaning is instead only given by the *OED*, the *Merriam-Webster* and the *Collins*, which indicate that *mandatory* can designate a power or a State holding a mandate from the League of Nations, or the system of rule by mandate and also a territory subject to rule by mandate (e.g. “When the rump of mandatory Palestine was conquered by Israel in 1967, [...]”; *OED*). The *Cassell’s Dict.* describes *mandatory* as “compulsory in suggesting an imposed role” (e.g. “mandatory silence in the library”; p. 106); it further specifies that *mandatory* is milder than *compulsory* in “that it stops short of suggesting coercion and punishment as methods of enforcing the ruling”.

The bilingual dictionaries propose *obbligatorio*, *vincolante*, *imperative* and *ingiuntivo* as the translation equivalents of *mandatory* without taking into consideration the differences highlighted in the monolingual dictionaries, which, as we have already seen, give two main definitions of *mandatory*. Both the *Garzanti* and the *Il Ragazzini* give the word *mandatario* as the historical translation of *mandatory* (e.g. “mandatory state: potenza mandataria”; *Il Ragazzini*).

The synonyms of *mandatory* given in the dictionaries are more or less the same as for *compulsory* and *obligatory*; in particular, besides *compulsory* and *obligatory* themselves, the *Thesaurus.com* gives *binding*, *commanding*, *compelling*, *compulsatory*, *de rigueur*, *essential*, *forced*, *imperative*, *imperious*, *indispensable*, *involuntary*, *irremissible*, *needful*, *requisite*, to which the *ADESSE* adds *perceptive* and *directory*.

In conclusion, as in the case of *compulsory* and *obligatory*, the meaning of *mandatory* too appears to be directly related to something that is to be done because of a law, order or rule. The bilingual dictionaries mainly give *obbligatorio* as the Italian translation of the adjective *mandatory*, and also add *vincolante* and *imperative* as other possible translation equivalents. Finally, some of the synonyms given for the term are *binding*, *obligatory*, *compulsory* and *imperative*.

#### 4.2.3.4 Conclusion

From the analysis of the definitions of the adjectives, it emerges that the three adjectives *compulsory*, *obligatory* and *mandatory* convey the notion of something that people must do because of a rule or law. The bilingual dictionaries give only one translation equivalent for the three near-synonyms, namely *obbligatorio*. According to the *Thesaurus.com* and the *MacMillan*, the three adjectives have the same corresponding near-synonyms: *binding*, *required*, *imperative* and *de rigueur*.

In conclusion, as Table 4.3 shows, the data retrieved in the dictionaries seem to show that *compulsory*, *obligatory* and *mandatory* are not freely interchangeable because, although the three of them are used to indicate something that must be done because of a rule or an authority, the definitions also differ in some details, that is only *compulsory* implies a punishment for those who do not respect the obligation, while *obligatory* and *mandatory* do not, and only *compulsory* appears to be used when something is done to keep people safe.

Table 4.3 *Compulsory, obligatory, mandatory*

Word class: adjectives				
		<i>COMPULSORY</i>	<i>OBLIGATORY</i>	<i>MANDATORY</i>
Meanings	Something that must be done because a rule or law	+	+	+
	Something done to keep people safe	+	-	-
	It involves a punishment for those who do not respect the obligation	+	-	-
Italian equivalents	<i>Obbligatorio</i>	+	+	+
	<i>Coatto, forzato</i>	+	-	-
	<i>Vincolante</i>	-	-	+
Synonyms	<i>Binding, required, de rigueur, imperative</i>	+	+	+

(The plus and minus signs indicate presence vs absence, respectively, of a given feature.)



#### 4.2.4 The adverbs

In the following sub-sections, I will introduce and give dictionaries definitions of the adverbs *maybe*, *perhaps* and *possibly*.

##### 4.2.4.1 *Maybe*

The adverb *maybe* is the shortened version of the phrase *it may be*. *The American Heritage Dictionary of Indo-European Roots* gives *magh-* as the original root, which means ‘to be able, to have power’. With particular reference to its first usage, the *OED* reports that, even if the word *maybe* can be found in texts of major writers of the 17<sup>th</sup> century, it is not frequently used in standard literary English before the mid-19<sup>th</sup> century. In the 19<sup>th</sup> century, the word occurs in poetic sources and also in novels, in which it is used as a marker of dialectal or colloquial speech.

Before giving the different meanings of the term, the *MacMillan* specifies that *maybe* can be used as a sentence adverb (in a comment referring to a whole sentence or clause) or as an ordinary adverb<sup>15</sup> (before a number). The *Collins*, too, differentiates between the adverbial and the sentential use of the adverb and then provides a list of synonyms that can be used in specific contexts, namely *perhaps* as a substitute of the adverb *maybe* and *possibly* and *neither yes nor no* as a substitute of *maybe* used as a sentence adverb.

Moreover, the definitions given by the monolingual dictionaries indicate that *maybe* is mainly used to express uncertainty when someone does not know if something is true or may happen in the future (e.g. “I do think about having children, maybe when I’m 40”; *Cobuild*, p. 888); the dictionaries also add that *maybe* can be used in the spoken language to make suggestions or give advice when someone is not quite sure about what to do or does not want to agree or disagree (e.g. “Maybe we should call a doctor”; *MacMillan*). *Maybe* is also said to mean ‘sometimes’ (e.g. “At weekends she would drive into Oxford, do the shopping, and maybe visit a few friends”; *MacMillan*) and finally, to indicate that someone is guessing a number, a quantity or a value, rather than stating it exactly (e.g. “The problems really started maybe two or three years ago”; *Longman*). To these definitions, the *OED* adds that the adverb *maybe* also has a rare use in the American colloquial language “before a negative, as an emphatic assertion of the corresponding positive statement” and cites an example from *In American’ Poems 18*

---

<sup>15</sup> I used the term *ordinary adverb* following the wordings given in the dictionaries that I consulted.

published in 1926: “[w]ell, boys, how would you like to go to the circus?’ Say, maybe Ern and me didn’t jump up!”. The *OED* also gives another colloquial and chiefly American use of the term, stating that the adverb can be “used after a statement to indicate that the facts are indisputable, or following a stipulated condition, indicating that it is not open to negotiation” (an example is given from the *Las Vegas Journal* of 1999: “Danny Crawford played tough guy when he robbed a rural Illinois bank last summer, holding out a nylon bag to a teller and demanding that she ‘fill it up, and I don’t mean maybe’”).

The bilingual dictionaries give *forse*, *probabilmente* and *può darsi che* as the Italian translation equivalents of *maybe* used as a sentence adverb, and only the *Hoepli* indicates the translation equivalents *più o meno* and *su per giù*, which correspond to *maybe* when it is used as adverb marker of vagueness or imprecision before a number or an amount (e.g. “We’ll stay for maybe a week: ci tratterremo più o meno una settimana”; p. 425).

With reference to the synonyms of the adverb *maybe*, the *Thesaurus.com* gives the following phrasal constructions: *it may be*, *can be*, *might be*, *it could be* and single words such as *conceivably*, *credible*, *feasible*, *imaginably*, *obtainable*, *perchance*, *perhaps*, *possibly* and *weather permitting*. Besides *perhaps*, *possibly* and *it may be*, the *ADESSE* adds *haply* and *as luck may have it*. In addition, the *Merriam-Webster*, the *Collins* and the *Oxford Dict.* monolingual dictionaries also, and only, suggest a few synonyms for the term, namely *perhaps*, *possibly*, *neither yes nor no*, *it could be*, *conceivably*, *perchance*, *mayhap*, *peradventure*, which correspond to the synonyms given in the synonym dictionaries.

From the survey on the dictionaries, it emerges that *maybe* can be used as a modifier of the whole sentence to signal that the speaker/writer is not sure that something can happen or may be true, and with the meaning of ‘sometimes’; in this cases the bilingual dictionaries suggest the Italian equivalents *forse* and *probabilmente*. On the other hand, when *maybe* is used before numbers, the Italian translation equivalents given by the bilingual dictionary *Hoepli* are *più o meno* and *su per giù*. Finally, the synonyms of *maybe* found in the dictionaries are both single words like *possibly* and *perhaps*, and phrases like *it may be* (the phrase which the adverb derives from) and *it could be*.

#### 4.2.4.2 *Perhaps*

According to the *OED*, the adverb *perhaps* comes from the “Anglo-Norman *per* or its etymon classical Latin *per* + the plural of *hap*”. The dictionary adds a religious comment on the term, in which it is specified that the term “occurs only three times in the Bible of 1611, all in the New Testament, and all originally present in the Rheims New Testament (1582)”.

As in the case of *maybe*, before giving any definition, the *MacMillan* specifies that *perhaps* can be used both as a sentence adverb (when making a comment on the whole sentence or clause) and as an ordinary adverb (before a number); the same distinction is given in the *Collins*, which also lists *possibly* and *maybe* as near-synonyms of *perhaps*, while adding that *it may happen*, *be so*, etc. can be used as sentence substitute of *perhaps*. In addition, the *OED* indicates that *perhaps* can express “a hypothetical, contingent, conjectural, or uncertain possibility” by “modifying a statement or question” or by “modifying a word or phrase independently”.

The monolingual dictionaries indicate that *perhaps* is used to express uncertainty about something or to signal that someone does not know if something is true or not (e.g. “If you want a new summer dress, perhaps have a look in Marks and Spencer”; *Longman-Activator*, p. 838). More specifically, they point out that it occurs when expressing polite opinions without being too certain or definite (e.g. “This is perhaps her finest novel yet”; *Longman*), that it is also used when giving suggestions or advice (e.g. “[W]ould you perhaps consent to act as our guide?”; *Oxford Dict.*), that it also occurs when replying politely to a question without agreeing or disagreeing with it (e.g. “‘I think he must have made a mistake.’ ‘Yes, perhaps’”; *MacMillan*) and that, especially in formal English, it appears in sentences when making a polite request (e.g. “Well, perhaps you’ll come and see us at our place?”; *Cobuild*, p. 1066). Finally, *perhaps* is said to be used when guessing a number, time or amount without being certain that it is correct (e.g. “The shed is 20, perhaps 25, feet long”; *Merriam-Webster*).

The bilingual dictionary *Hoepli* and *Garzanti* give *circa*, *all’intorno* and *si e no* as the Italian equivalents of *perhaps* when used before a number, as in the sentences “[w]e waited for perhaps an hour: aspettammo per circa un’ora” (*Hoepli*; p. 509) and “[t]here were perhaps 50 in the audience: c’era un pubblico si e no di 50 persone” (*Garzanti*; p. 906). The dictionaries also indicate *forse*, *probabilmente*, *può darsi* and *magari* as possible alternatives (e.g. “Why not invite one of our cousins, Doreen or Irene perhaps? Perché non invitare una delle nostre cugine, magari Doreen o Irene?”; *Garzanti*, p. 906).

Once more, the *Thesaurus.com* and the *ADESSE* list the same near-synonyms for *perhaps* as for the adverb *maybe*, namely *as it may be*, *as the case may be*, *conceivably*, *feasibly*, *for all one knows*, *imaginably*, *it may be*, *maybe*, *perchance*, *possibly*, *reasonably*, *haply*, *belike*, *by chance*, *as luck may have it* and *peradventure*. To these, the *Merriam-Webster* adds *mayhap*, which derives from the construction *it may hap* (meaning ‘it may happen’), while the *MacMillan* adds *about*, *around*, *conservative*, *fair*, *some*, *speculative* and *what* as synonyms to be used when the adverb precedes a number.

In conclusion, according to the dictionaries, the adverb *perhaps* is used both as an adverb modifying a number and as an adverb modifying a whole sentence). The adverb appears to occur in contexts in which the word is used when a certain degree of politeness is required: for example, when making polite suggestions or requests, giving polite advice or opinions. The Italian translation equivalents given by the bilingual dictionaries can be divided into two groups: the ones referring to *perhaps* as an adverb expressing vagueness (*circa*, *all’intorno*, *sì e no*) and the ones referring to it as a sentence adverb (*forse*, *probabilmente*, *magari*, *può darsi*). Finally, the synonyms which are commonly associated with *perhaps* are *perchance*, *possibly*, *maybe* and *for all one knows*.

#### 4.2.4.3 Possibly

*The American Heritage Dictionary of Indo-European Roots* gives *poti-* as the root of the word *possible* with the meaning of ‘powerful, lord’. In particular, *possible* is said to derive from the Latin compound *posse* meaning ‘to be able’. The *OED* does not provide the etymology of the adverb *possibly* and only suggests a comparison with the Middle French word *possiblement*, which is said to have largely been used in 1337 and to have now become rare.

The *OED* lists some meanings of *possibly* which nowadays are no longer active: (1) “[u]sed for possible in adverbial phrases, as if possibly, soon as possibly, by all means possibly” (e.g. “[h]e then declared an Intention of going ... to Fitzpatrick’s Lodgings, in order to prevail with him, if possibly, to consent to a Separation from his Wife”; H. Fielding, 1749); (2) “[a]ccording to one’s ability; as much or as well as one can” (e.g. “[y]ou have provided for every one of them as a Free Man, as a man that doth (act) possibly, rationally, and conscientiously”; O. Cromwell, 1657). The *OED* also lists two definitions showing how nowadays the adverb is chiefly used as an intensifier of

*can/could*, when it indicates a possible manner within a range of possibility, and of *may/might*, when it is used to qualify a statement and express contingency or uncertainty. Such uses as an intensifier of the modal verbs *can/may* are also given in the *Oxford Dict.* and the *Longman*, which specify that *possibly* can be used with a modal verb when making a polite request (e.g. “Could you possibly close the window?”; *Longman*).

According to the monolingual dictionaries, *possibly* is used in some specific situations: when something is likely to happen but there is uncertainty (e.g. “Exercise will not only lower bloody [sic] pressure but possibly protect against heart attacks”; *Cobuild*, p. 1113); when someone wants to emphasise that they have tried their hardest to do something (e.g. “She’s done everything she could possibly think of to help”; *MacMillan*); when someone feels surprised, shocked or puzzled by something they have seen or heard (e.g. “It was the most unexpected piece of news one could possibly imagine”; *Cobuild*, p. 1113). The *Longman-Activator* adds that the adverb can also be used when a number, amount, time etc. may be correct but there is no certainty (e.g. “It may possibly be ten or eleven o’clock before I get home tonight”; p. 838).

The bilingual dictionaries provide different Italian equivalents of *possibly* depending on its contextual meaning. In particular, they give *forse*, *può darsi*, *probabilmente* as the most generic translation equivalents; (e.g. “This may possibly be the reason: questa è probabilmente la ragione”; *Hoepli*, p. 533) and list *affatto*, *assolutamente*, *proprio* as synonyms to be used in negative sentences; (e.g. “I can’t possibly remember all those details: non posso assolutamente ricordare tutti quei particolari”; *Garzanti*, p. 948). On the other hand, the adverb is translated as *mai* in idiomatic questions (e.g. “What can she possibly mean? Che cosa vorrà mai dire?”; *Il Ragazzini*) and as *in qualche modo*, *in un modo o nell’altro* and other emphatic expressions to convey emphasis (e.g. “He read everything he possibly could about the subject: ha letto tutto ciò che gli era umanamente possibile sull’argomento”; *Garzanti*, p. 948). Finally, only the *Hoepli* indicates that *possibly* can be used as an intensifier of *can/could*; (e.g. “How could they possibly expect me to do it? Come potevano mai sperare che lo facessi?”; p. 533).<sup>16</sup>

The synonyms shared by the *Thesaurus.com* and the *ADESSE* are *perhaps*, *perchance*, *peradventure* and *maybe*; the former also lists *probably*, *conceivably*,

---

<sup>16</sup> *Il Ragazzini* 2013 warns against the Italian false friend of *possibly*, specifying that the adverb does not mean ‘possibilmente’.

*maybe, likely, God willing, at all, by any chance, by any means, could be, if possible, in any way, not impossibly, within realm of possibility*, and the latter includes *mayhap, haply, it may be* and *as luck may have it*. Both the *Webster's Dict.* and the *Cassell's Dict.* do not give any entry for the adverb *possibly*.

In conclusion, the adverb *possibly* can be used combined with the modal verbs *can/could* and *may/might* in the expression of uncertainty, emphasis or surprise. Besides *forse, può darsi* and *probabilmente*, the translation equivalents given in the bilingual dictionaries are related to the different meanings given in the monolingual dictionaries (i.e. in negative sentences: *affatto, assolutamente, proprio*; in idiomatic questions: *mai*). Among the synonyms given, we find *maybe* and *perhaps*, which are the other adverbs here under study.

#### **4.2.4.4 Conclusion**

The analysis of the definitions given in the dictionaries of the three adverbs *maybe, perhaps* and *possibly* shows that the three of them have very similar meanings, especially when they are used to indicate a state of uncertainty. They appear to share Italian translation equivalents, which are *forse, probabilmente* and *può darsi*, while the synonyms that they are said to have in common are *as luck may have it, perchance, peradventure, haply, conceivably* and *as it may be*.

In conclusion, as is summarized in Table 4.4, the data retrieved in the dictionary survey show that *maybe, perhaps* and *possibly* are not freely interchangeable because, although the three of them are used to express uncertainty about something, such as a situation or value, the meanings also differ in other respects, that is, both *maybe* and *perhaps* can be used as sentence adverbs and ordinary adverbs and to suggest something or give advice, while only *possibly* can be used to give emphasis to a sentence and express surprise about something.

Table 4.4. *Maybe, perhaps, possibly*

Word class: adverbs				
		<i>MAYBE</i>	<i>PERHAPS</i>	<i>POSSIBLY</i>
Meanings	To express uncertainty	+	+	+
	Sentence adverb vs ordinary adverb	+	+	-
	To guess a number or a value	+	+	+
	To express polite requests	-	+	-
	To suggest something or give advice	+	+	-
	To emphasize or express surprise about something	-	-	+
Italian equivalents	<i>Forse, probabilmente, può darsi</i>	+	+	+
	<i>Affatto assolutamente, proprio</i> (in negative sentences)	-	-	+
	<i>Circa, all'intorno, più o meno</i>	+	+	-
Synonyms	<i>As luck may have it, perchance, peradventure, haply, conceivably, as it may be</i>	+	+	+

(The plus and minus signs indicate presence vs absence, respectively, of a given feature.)

### 4.3 Conclusion

To sum up, the survey of the dictionaries shows that the sets of words considered are likely to be near-synonyms; on the one hand, they have some meanings and patterns of use in common; on the other, they have distinctive specific characteristics. In particular, the members of each set mainly share one meaning (the act of killing for the nouns; the act of making someone feel worried or upset or interrupting for the verbs; the reference to something ordered by a law or an authority for the adjectives, and the notion of uncertainty for the adverbs), but at the same time, each of them has peculiar meanings not shared by the others: for example, only *assassin* involves important political figures as victims, *killer* comprises the death of animals; *to disturb* conveys the meaning of annoying unintentionally, while *to bother* denotes the act of annoying deliberately; *compulsory* is used when an obligation is intended to keep people safe, whereas *mandatory* and *obligatory* are not; and finally, *maybe* and *perhaps* appears to

occur when suggesting something, whereas *possibly* seems to be used to give emphasis to the sentence.

However, despite the wealth of definitions and examples retrieved, the dictionaries do not appear to provide all the information necessary to a foreign language learner to accurately use those terms. For example, none of the dictionaries gives details on the frequency of occurrence of the words and only in few cases do they indicate the register or the field in which the terms are more likely to occur. Moreover, the dictionaries do not give information on the possible syntactic restrictions and collocational preferences of the terms in question. This kind of information, however, may be more easily accessible through corpus data.

In the next chapter I will report on the findings obtained from the consultation of large corpora. Concordances of the terms under study will reveal their comparative frequencies of occurrence across genres and some of their phraseological patterns.



## CHAPTER 5

### FINDINGS FROM CORPUS SOURCES

#### 5.1 Introduction

In this chapter I will present corpus data regarding the terms under study (i.e. the nouns *murderer*, *killer* and *assassin*; the verbs *to disturb* and *to bother*; the adjectives *compulsory*, *obligatory* and *mandatory*; and the adverbs *maybe*, *perhaps* and *possibly*). Concordances from the *BNC*, the *COCA* and the *CWo* allowed me to compare the frequency of occurrence of the above sets of near-synonyms (both as lexemes and in their various word-forms), both in general and in different genres or text types, and also to identify some of their phraseological patterns, by considering their left and right collocates.

In the following sections and sub-sections I will outline the findings concerning the above mentioned sets of near-synonyms. In particular, when describing the different patterns of occurrence, I will first give the data referring to the whole corpora and then I will refer specifically to my datasets, namely the 200 occurrences retrieved from the *BNC* and the 200 from the *CWo* per each term (see section 3.3.2). The aim of my corpus analysis is to check if, to what extent and in what ways such terms can be considered completely interchangeable.

#### 5.2 The nouns

In the following sub-sections I will report on corpus findings about the nouns *assassin*, *murderer* and *killer*. In each sub-section I will first present the findings referring to the overall raw and normalised frequency of occurrence of the terms in question, in their various word-forms, and their distribution across different genres both in the *BNC* and the *COCA*. Then, I will look at both my datasets from the *BNC* and *CWo* and at the whole corpora to identify the left and right collocates of the terms analysed, and to identify the meanings these terms appear to convey and their colligational patterns.

##### 5.2.1 *Assassin*

The noun *assassin* occurs 454 times in the *BNC* and 2,262 times in the *COCA*, which corresponds to almost 5 times pmw in both corpora. As Table 5.1 shows, the

singular<sup>17</sup> form of *assassin* is almost twice as frequent as its plural counterpart<sup>18</sup> both in the *BNC* and the *COCA*. In addition, in both corpora, the highest frequency of the term *assassin* occurs in fiction, and in particular, the term occurs 322 times in the *BNC* and 968 in the *COCA*, that is more than 20 times pmw in the former and 10 times pmw in the latter. The reason may be that the plot of popular literary genres (e.g. detective stories, thrillers), but also plays and movie scripts, often pivots on a criminal character such as an assassin (e.g. “The Assassin (18) in Orchard Cinema Two at 9pm”, *BNC*: K2Y).

Table 5.1 Frequency of occurrence of the term *assassin* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens <sup>19</sup>	Frequency pmw <sup>20</sup>	Tokens	Frequency pmw
<b>Word-forms</b>				
<i>Assassin, assassin's</i>	302 (66.52%)	3.14	1,400 (61.89%)	3.02
<i>Assassins, assassins'</i>	152 (33.48%)	1.58	862 (38.11%)	1.86
<b>Genres</b>				
Spoken	2 (0.44%)	0.20	295 (13.94%)	3.09
Fiction	<b>322 (70.92%)</b>	<b>20.54</b>	<b>968 (42.79%)</b>	<b>10.70</b>
Magazine	26 (5.73%)	3.58	476 (21.04%)	4.98
Newspaper	28 (6.17%)	2.68	389 (17.20%)	4.24
Non-academic	49 (10.79%)	2.68	-	-
Academic	5 (1.10%)	0.33	134 (5.93%)	1.54
Other	22 (4.85%)	1.06	-	-
TOTAL	454 (100%)	4.72	2,262 (100%)	4.87

Both in my datasets and in the whole *BNC* and *CWo* corpora, the left collocates of the noun *assassin* are attributive adjectives and past participles used as pre-modifiers

<sup>17</sup> In my frequency counts, after having checked the instances manually, I considered both *assassin* and *assassin's* (the genitive word form meaning ‘of assassin’); none of the instances considered shows *assassin's* with the meaning of ‘assassin is’.

<sup>18</sup> I counted both the plural word forms *assassins* and *assassins'* (meaning ‘of assassins’).

<sup>19</sup> Here and in the following tables, the column *Tokens* lists the raw number of occurrences of the term in the whole corpus.

<sup>20</sup> Here and in the following tables, the column *Frequency pmw* includes the normalized frequency of occurrence pmw.

(within my dataset: *BNC* 23.5%; *CWo* 25%). Table 5.2 shows the normalized frequency counts of some of the most frequent pre-modifiers within the whole corpora, which is also compared to their frequency of occurrence in my dataset, where, however, they occur infrequently. Figg. 5.1.1 and 5.1.2, instead, show two screenshots of concordances from my datasets in which the noun is modified by the adjectives, past participles, numbers and proper names listed in Table 5.2.

In addition to the adjectival pre-modifiers listed in Table 5.2, in my datasets are given other adjectival pre-modifiers, which occur only once; in particular in the *BNC* dataset there are adjectives such as *curious*, *cynical*, *expert*, *famous*, *fellow*, *well-bribed* and in the *CWo* dataset, there are other pre-modifiers such as *grey-haired*, *mystical*, *political*, *possible*, *presidential*, *shadowy*, *sultry*. These attributes denote personal characteristics or contextual circumstances relevant to the assassin being talked about. In addition to the adjectival and participial pre-modifiers, other types of pre-modifiers are numbers (*BNC*: 6 occurrences, 3%; *CWo*: 3 occurrences, 1.5%) and proper names, which identify either an organization the assassin is a member of, or the name of the assassin's victim (*BNC*: 4 occurrences, 2%; *CWo*: 3 occurrences, 1.5%).

Fig. 5.1.1 Screenshot of pre-modifiers of *assassin* in the *BNC* dataset

2	HKY	imprisonment for the murder of Benigno Aquino and of his alleged < assassin > Rolando Galman. Among 20 others acquitted was the former
26	EF1	stand to lose a great deal if Mobutu has his way. A team of four < assassins >, made up from the now disbanded Security Police, have vowed
28	HKS	heavy personal bodyguard. His killer, a 16-year-old hired < assassin > who was wounded and captured, later claimed that a stranger
39	CDY	something from him, this stranger whose job was to shoot the IRA < assassins > if they came. He didn't know what it was, but now it wasn't
106	HH3	their suppression of political dissent by utilizing the trained < assassin >: the sicario. It is not uncommon to hear about the assassination

Fig. 5.1.2 Screenshot of pre-modifiers of *assassin* in the *CWo* dataset

39	usnews	just the creator of "The Dating Game," but also was a CIA < assassin >. Did you want to say, "Come on, nice fantasy?" "A. I really
69	indnews	have been detailed." The report is a goldmine for any potential < assassin >," one official observed, referring to portions in Vol. IV
70	usbooks	was before. Life imitated art and I became the professional < assassin >. Today, sitting in this room, I can look at the way Magenta
105	usnews	Garcia Knight Ridder News Service KABUL, Afghanistan - Two < assassins > gunned down an Afghan vice president yesterday as he was being
110	carnews	members swept the Shiite city of Karbala looking for would-be < assassins > of Izzat Ibrahim, Saddam's deputy on the Revolution Command

Table 5.2 Attributive left collocates of the term *assassin*:  
pre-modifiers in the *BNC* and *CWo* datasets

Pre-modifiers of <i>assassin</i>	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<b>Adjectival and participial pre-modifiers</b>			<b>47(23.5%)</b>	<b>50 (25%)</b>
<i>alleged</i>	5.13	4.97	2 (1%)	1 (0.5%)
<i>baby-faced</i>	-	-	-	2 (1%)
<i>dead/deadly</i>	-	5.13	-	4 (2%)
<i>great</i>	-	-	2 (1%)	-
<i>hired</i>	8.96	6.12	3 (1.5%)	1 (0.5%)
<i>lone</i>	-	-	-	2 (1%)
<i>potential</i>	4.43	3.69	3 (1.5%)	4 (2%)
<i>professional</i>	4.54	3.81	1 (0.5%)	1 (0.5%)
<i>silent</i>	4.73	5.15	3 (1.5%)	1 (0.5%)
<i>suicidal/suicide</i>	-	2.09	-	2 (1%)
<i>suspected</i>	3.90	6.53	1 (0.5%)	-
<i>trained</i>	6.48	7.24	2 (1%)	1 (0.5%)
<i>unknown</i>	-	-	-	2 (1%)
<i>would-be</i>	8.31	9.40	6 (3%)	5 (2.5%)
<b>Numbers</b>			<b>6 (3%)</b>	<b>3 (1.5%)</b>
<i>two</i>	-	-	2 (1%)	2 (1%)
<i>three</i>	-	-	1 (0.5%)	1 (0.5%)
<i>four</i>	-	-	3 (1.5%)	-
<b>Proper names</b>			<b>4 (2%)</b>	<b>3 (1.5%)</b>
<i>CIA</i>	-	9.83	-	2 (1%)
<i>Dark Elf</i>	7.69	-	1 (0.5%)	-
<i>IRA</i>	4.81	3.70	2 (1%)	1 (0.5%)
<i>de Montfort</i>	8.66	-	1 (0.5%)	-

As Table 5.3 shows, the immediate left collocates of *assassin* that stand out are determiners (*BNC*: 98 occurrences, 49%; *CWo*: 80 occurrences, 40%) or possessive pronouns (*BNC*: 11 occurrences, 5.5%; *CWo*: 7 occurrences, 3.5%).

Table 5.3 Additional left collocates of *assassin(s)* occurring in the *BNC* and the *CWo* datasets: determiners, possessive pronouns, numbers and proper names

Frequency		
	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<b>Determiners</b>	<b>98 (49%)</b>	<b>80 (40%)</b>
<i>an</i>	16 (8%)	18 (9%)
<i>the</i>	82 (41%)	62 (31%)
<b>Possessive pronouns</b>	<b>11 (5.5%)</b>	<b>7 (3.5%)</b>
<i>his</i>	10 (5%)	6 (3%)
<i>her</i>	1 (0.5%)	1 (0.5%)

*Assassin* can also be used as a pre-modifier of another noun; in the *BNC*, the most frequent nouns used as right collocates are *bug* (5.15 pmw), *feed* (4.32 pmw), *lie* (3.51 pmw), *troops* (1.74 pmw). (In my dataset, however, of the above terms only *troops* occurs, and only once in the *BNC* dataset.)

In my datasets, as Fig. 5.2 displays, when the term *assassin* is used in the genitive form (*assassin's/assassins' + noun*), the following head noun can indicate objects normally found, or events taking place, at a crime scene (i.e. *arrest, blow, bullet, dagger, fuse, gun, sword, trigger*); alternatively, it can indicate parts of the body; these occurrences might be part of descriptions of given characters in fiction (e.g. *arm, face, grey eyes, mouth, voice*).

Fig. 5.2 Screenshot of *assassin* in the genitive form in the *CWo* dataset

```

80   CRE           The bundle of clothes dropped from the < assassin >'s arm leaving him naked. But he put up no defence; he simply
84   K95           And you, too, sir Brian. why? Do you also fear the < assassin >'s dagger? How well do you sleep at night? What secrets did
86   CRE           Estabrook watched the smoke drift up over the < assassin >'s grey eyes, and before he could prevent himself he was telling
90   GL7           his coat off what remained of the lantern pole. The cut of the < assassin >'s sword had split it from below the shoulder to the hem. The
91   BN1           you know?' Mouse shrugged. 'Made it up, hrh?' The < assassin >'s voice was lazy with contempt. The dog was snarling

```

As Figg. 5.3.1 and 5.3.2 show, in my datasets, the noun *assassin* can occur in sentences when reference is made to the victims of assassination. These can be political figures such as Margaret Thatcher, Benigno Aquino, the Afghan Vice-President Haji Abdul Qadir, the Pope and others (*BNC*: 13 occurrences, 6.5%; *CWo*: 34 occurrences, 17%) and, less frequently, common people (*BNC*: 10 occurrences, 5%; *CWo*: 21

occurrences, 10.5%). In such cases *assassin* may be the subject of a transitive verb or the head of a noun phrase including the post-modifying preposition *of*.

Fig. 5.3.1 Screenshot of *assassin* in the *BNC* dataset:  
important political figures and common people as the victims

```
5 K95 there is!' Fitzzormonde hissed. 'I am a murderer, Father. An < assassin >. I took my friend's life. No! No!' He shook his head as if
29 ABJ indignation at the killing. 'By murdering President Rene Muawad, his < assassins > chose to strike down a man of dialogue and reconciliation,'
32 ECT . A few nights later the job was done. Coetzee met his < assassins > in a bar around about ten o'clock. It seemed that they had
35 G04 rope, her other hand seeking the wallbars. The middle < assassin > fired even as she dropped. Wood splintered next to her. She
108 AJF Ford, who sings a duet with John Warnock Hinckley, the would-be < assassin > of President Ronald Reagan. And Squeaky, it may be remembered
```

Fig. 5.3.2 Screenshot of *assassin* in the *CWo* dataset:  
important political figures and common people as the victims

```
7 times born in Swindon, Wiltshire, 1931. DEATHS: Marcus Junius Brutus, < assassin > of Julius Caesar, committed suicide near Philippi, Macedonia
11 usspok today, on November 24th, 1963, Lee Harvey Oswald, the accused < assassin > of President John F. Kennedy, was murdered. Dallas, Texas,
47 times a threat. The incident raised concerns that a determined < assassin > could have attempted to kill the Pope, as happened in 1981,
50 oznews collateral. It stars Tom Cruise as Vincent, a suave grey-haired < assassin > on a mission to kill five people in one night, and Jamie Foxx
103 times shouted at his tormentors that he was not afraid to die. 'When the < assassin > was pointing the pistol at him, this boy tried to take off
```

In addition, the concordances instantiating the noun *assassin* also suggest that the referent of the noun is talked about in terms of his/her profession; this occurs when, for instance, the predicate (e.g. *to win*, *to fail*, *to miss*, *to make a mistake*) makes reference to the mission of an assassin, which can be unsuccessful because of a mistake or a failure (*BNC*: 4 occurrences, 2%; *CWo*: 3 occurrences, 1.5%; e.g. “[y]et, Sir Edmund, I am confident the *assassin* will eventually make a mistake”, *BNC*: H90; “[w]hen Morjin’s *assassins* failed to kill you, [...]”, *CWo*: usbooks).

In my datasets, the noun *assassin* can occur preceded or followed by the coordinating conjunctions *and/or* (*BNC*: 16 occurrences, 8%; *CWo*: 12 occurrences, 6%), which are immediately followed or preceded, respectively, by another noun. In this pattern, the noun *assassin* occurs only in its plural word-form and the nouns it is coordinated with all indicate people who are involved in criminal and illegal acts (e.g. *accomplice*, *bombers*, *burglars*, *criminals*, *crook*, *merchants*, *rebellions*, *rioters*, *rippers*, *robbers*, *spies*, *terrorist*, *thieves* and *trolls* in the *BNC* dataset; *kidnappers*, *spies*, *terrorists*, *thieves*, *traitors* and *turncoats*, 1 instance each in the *CWo* dataset + *thugs*, 4 instances in the *CWo* dataset). The pattern is more frequent in the *BNC* than the *CWo* dataset (see Fig. 5.4).

Fig. 5.4 Screenshot of *assassin* in the *BNC* dataset:  
nouns coordinated with *assassin* by means of *and/or*

```
2 HA3 said a voice by his ear. Down below, thieves, < assassins >, trolls and merchants all realised at about the same moment
8 HU0 blackguards within the liberties of Paris; rioters, burglars and < assassins >, hand in glove with some of the most desperate characters of
9 HLR parliamentary debate as a "charter for crooks, criminals and < assassins >". Faced with parliamentary defeat, de Klerk had on Oct. 21
11 AT7 number on the market now about Secret Service and spies and < assassins > and rebellions.' There was silence at the table for a
88 FBL nowhere else to go. If sending in the bombers or the trained < assassins > is out, what else does that leave? Probably the best yet least
```

The colligational patterns of *assassin* instantiated in the whole corpora and my specific datasets show that the term can frequently occur as the subject, the object of predicates or in a prepositional phrase. When the noun is used as the subject of a predicate (*BNC*: 64 occurrences, 32%; *CWo*: 51 occurrences, 25.5%, see Table 5.4), the verbs describe the action in which the assassin is involved (e.g. *to fire* and *to shoot*, which indicate the use of weapons to attack someone; *to kill* and *to murder*, which convey the meaning of making an attempt upon someone's life; *to assassinate* is instantiated only once in the *CWo* dataset). Other predicates referring to the act of killing which occur with low frequency scores in my datasets are *to gun down*, *to attack*, *to attempt to kill*, *to pour fire*, *to slaughter*, *to infect*, *to stalk*, *to scare*; we also find verbs that show the physical movements or position of the assassin (*to stand*, *to walk*, *to run off*, *to rush out*, *to escape*, *to follow*, *to point*, *to disappear*); there are also verbs of communication (*to say*, *to nod*, *to speak*, *to stare back*, *to scream*, *to phone*); this is exemplified in Fig. 5.5.2.

Table 5.4 *Assassin* as the subject of predicates in the two corpora and in the *BNC* and *CWo* datasets

<i>Assassin</i> as the subject	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>to kill</i>	2.64	3.2	6 (3%)	19 (9.5%)
<i>to fire</i>	3.74	2.02	2 (1%)	2 (1%)
<i>to murder</i>	-	4.84	9 (4.5%)	13 (6.5%)
<i>to shoot</i>	2.64	3.63	4 (2%)	7 (3.5%)

Fig. 5.5.1 Screenshot of *assassin* in the *CWo* dataset as subject of the verbs *to kill*, *to murder* and *to assassinate*

1 safrica father killed so as to lure the couple out of hiding. The < assassin > was then to murder the young couple at the funeral.  
 2 indnews became a bit of a Luddite though. Seventeen bullets and two < assassins > killed him eventually; his ideals had been mangled far earlier  
 3 usspok the Lincoln conspiracy trials. The four people who helped < assassin > John Wilkes Booth assassinate President Lincoln were all tried

Fig. 5.5.2 Screenshot of *assassin* in the *BNC* dataset: other verbs

50 G04 rope, her other hand seeking the wallbars. The middle < assassin > fired even as she dropped. wood splintered next to her. She  
 58 CRE stay still, she started to stumble towards the place where her < assassin > had fallen. The driver was being helped from his smashed vehicle  
 97 CJP Musa Anter was hit by four bullets, Orhan Miroglu by three. The < assassin > then ran off. Musa Anter died on the way to hospital.  
 71 CEM country music. The right equipment for Quayle-hunting. < Assassins > don't scare my boy Bill BILL Clinton has thought about the  
 92 AD7 the window of a car as he returned from exercise. The would-be < assassins > missed their prey, and Phar Lap, unperturbed by the incident

Instead, when the noun *assassin* is used as the object of a predicate (*BNC*: 29 occurrences, 14.5%; *CWo*: 25 occurrences, 12.5%; the most frequent predicates are shown in Table 5.5); the term can occur with verbs such as *to hire* and *to send* (*BNC*: 8 occurrences, 4%; *CWo*: 8 occurrences, 4%), that is in contexts that make reference to jobs and errands, thus implying that the term identifies a profession, however illegal. Other predicates, which *assassin* is the object of, identify actions carried out by police or authorities (e.g. *to put to death*, *to catch up*, *to flush out*, *to look for*, *to find*, *to search for*, *to despatch*, *to execute*, *to forgive*, *to identify*, *to bring to justice*, *to kill*, *to shoot*).

Table 5.5 *Assassin* as the object of predicates  
in the two corpora and in the *BNC* and *CWo* datasets

<i>Assassin</i> as the object	Frequency		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>to hire</i>	5.03	6.12	3 (1.5%)	3 (1.5%)
<i>to send</i>	2.27	2.81	5 (2.5%)	5 (2.5%)

Fig. 5.6 Screenshot of *assassin* in the *CWo* dataset:  
object of predicates

```

1 brbooks way. And so the first courtesans pooled their money to hire < assassins >. The assassins exterminated the procurers, of course, and then
48 usbooks please the Lord of Lies --- he who has schemed and sent out < assassins > so that this war might take place. Will you do the bidding
54 usbooks There was always the possibility, though, that Jagang had sent < assassins > to Ayyindrill; some of those assassins could be sisters of the
56 usbooks he had his own household troops installed there." He sent < assassins > to intercept me at the border. I wonder if he thought I would
57 usbooks atching through this man's eyes. Jagang had in the past sent < assassins > after he had slipped into their thoughts. with Jagang in a

```

The noun *assassin* is also frequently instantiated in a prepositional phrase, that is as the object of the preposition that precedes it (*BNC*: 56 occurrences, 28%; *CWo*: 66 occurrences, 33%; see Figg. 5.7.1 and 5.7.2). In such cases, the PP can occur as the post modifier of a head noun or of a verb; the most frequent preposition occurring in this construction is *of* (*BNC*: 34 occurrences, 17%; *CWo*: 35 occurrences, 17.5%). Other prepositions which occur only few times are *for*, which identifies *assassin* as the beneficiary, *to*, which shows *assassin* as the indirect object and *by*, which indicates *assassin* as the agent in a passive clause.

Fig. 5.7.1 Screenshot of *assassin* embedded in a prepositional phrase  
in the *BNC* dataset

```

18 CM4 said disconsolately. 'I could atone.' 'The role of an < Assassin >,' Jaq reminded her, 'is not to feel remorse in any respect.
94 EF1 tailor's heart.' 'At least you're alright. Any news on the < assassin > and his accomplice?' 'Nothing yet,' Kolchinsky replied
43 ECU draws up her invincible battle plans, foils the conspiracies of < assassins >, learns of the infidelities and corruptions for which she
49 CM4 felt profoundly glad that he had never studied in the school of < Assassins >. 'Maybe later I'll learn how,' he said. 'Meanwhile, let's
94 CE5 the hotel who could have been of any possible interest to the < assassins >.' Cowley acknowledged that. 'Always assuming, Colonel

```



Fig. 5.7.2 Screenshot of *assassin* embedded in a prepositional phrase in the *CWo* dataset

```

4    usbooks    Pualani's eyes made me want to relate as well the story of the < assassin >'s arrow and all that had occurred since that dark afternoon
14   indnews    have been detailed." The report is a goldmine for any potential < assassin >," one official observed, referring to portions in Vol. IV
62   brbooks    murdered, a CIA case officer handed a poison pen to a Cuban < assassin > for use against Castro. Many of these schemes later came to
91   brbooks    local knowledge and some generally accepted notions about the < Assassins > that Marco Polo may have heard before his journey, or even
110  brbooks    century later that historians began to approach the history of the < Assassins > without bias and polemical preconceptions. Four historians

```

Finally, as Figg. 5.8.1 and 5.8.2 show, the noun *assassin* also appears to be instantiated (with lower frequency scores) in another colligational pattern, that is as the subject complement (*BNC*: 10 occurrences, 5%; *CWo*: 19 occurrences, 9.5%).

Fig. 5.8.1 Screenshot of *assassin* as subject complement in the *BNC* dataset

```

16   CRE        screamer, wavered, then returned to Jude. It wasn't an < assassin > any longer. Nor was it gentle. If it had a self, perhaps this
19   BNL        That thought nearly shocked Fox fully awake again. She was an < assassin > - a psychopathic killer. She had nothing to do with peace,
41   KRT        the army in Northern Ireland early today, was a top Republican < assassin >; he was named as Desi Grew, a leading member of the I R A.
73   HUO        I answered, 'but Moodie gave me the cloth, so he must be the < assassin >.' [Ah, there goes my chaplain again, jumping up and down
80   H90        'The servants of this house could, one or all, be either the < assassins > or in their pay.' 'we shall deal with them in the morning

```

Fig. 5.8.2 Screenshot of *assassin* as subject complement in the *CWo* dataset

```

12   usbooks    murderous look in her husband's eyes. "He's a child!" "He's an < assassin >," Damin corrected. Brak climbed to his feet, offering R'shieh
17   usbooks    Martin's grin turned wry. "Jules is an ex-Corsican bandit, an < assassin >, among other things. He was once sent to kill me." Luc beside
29   brbooks    planned assassinations for the Farseers. Had I said I was not an < assassin >, would never be one again? I wondered if I had been a liar
76   usbooks    secrets. 'Tulann was an assassin,' Kane said to me. 'And I'm an < assassin > of assassins. Some day I may kill the Great Beast himself
109  usnews     loves to mix things up inside. Laid-back Duncan is a silent < assassin > with more inside moves than anyone in the game. Advantage:

```

In conclusion, the data retrieved from the corpora show that the term *assassin* occurs with higher frequency scores in fiction. An examination of its immediate co-text shows the following patterns: *assassin* can be preceded by attributive adjectives and past participles which tend to portray a professional, hired and paid person; *assassin* also appears collocated with nouns which identify objects commonly found at a crime scene and is coordinated with other nouns denoting other kinds criminals; therefore these contexts suggest that *assassin* has a negative connotation. The noun *assassin* co-occurs both with proper names of political persons and nouns denoting common people as the victim of the killing action.

The term is instantiated in some colligational patterns; in particular, when the noun is used as the subject of transitive verbs (*BNC*: 32%; *CWo*: 25.5%), it is followed by verbs that describe the criminal actions carried out by the assassin such as *to kill* and *to murder*. When the noun is used as the object of a predicate (*BNC*: 14.5%; *CWo*: 12.5%), it occurs with verbs that both can make reference to illegal jobs and errands and to actions carried out by authorities. When the noun is used in an embedded construction (*BNC*: 28%; *CWo*: 33%), it frequently occurs as the object in a prepositional phrase

with the prepositions *of*, *for*, *to* and *by*. Finally, the term can also occur as a subject complement (*BNC*: 5%; *CWo*: 9.5%) identifying the subject as an assassin.

### 5.2.2 Murderer

The analysis of the frequency of occurrence of the term *murderer* reveals that the term occurs 454 times in the *BNC* and 2,262 times in the *COCA* (i.e. 5 times pmw in both corpora). As Table 5.6 displays, the singular word-form *murderer* is almost three times as frequent, and twice as frequent, as the plural *murderers* in the *BNC* and the *COCA*, respectively.<sup>21</sup> Furthermore, Table 5.6 also shows that in the *BNC* the word *murderer* has a high frequency of occurrence in fictional genres (53.87%) and a relatively high frequency of occurrence in the newspaper genre (133 instances, i.e.12.89%). In the *COCA* the term is highly used in the spoken language (33.84% of the times), and, as in the *BNC*, it is also relatively frequent in fiction (1,423 instances, i.e. 29.68%).

Table 5.6 Frequency of occurrence of the term *murderer(s)* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Word-forms</b>				
<i>Murderer, murderer's</i>	<b>760 (73.64%)</b>	7.90	<b>3,242 (67.63%)</b>	7.00
<i>Murderers, murderers'</i>	272 (26.36%)	2.83	1,552 (32.37%)	3.54
<b>Genres</b>				
Spoken	24 (2.33%)	2.41	<b>1,622 (33.84%)</b>	<b>16.97</b>
Fiction	<b>556 (53.87%)</b>	<b>34.95</b>	<b>1,423 (29.68%)</b>	<b>15.74</b>
Magazine	51 (4.94%)	7.02	528 (11.01%)	5.53
Newspaper	<b>133 (12.89%)</b>	<b>12.71</b>	802 (16.73%)	8.74
Non-academic	111 (10.76%)	6.73	-	
Academic	82 (7.94%)	5.35	419 (8.74%)	4.60
Other	75 (7.27%)	3.60	-	
<b>TOTAL</b>	<b>1,032 (100%)</b>	<b>10.72</b>	<b>4,794 (100%)</b>	<b>10.32</b>

<sup>21</sup> As in the case of *assassin*, the word-forms counted included the singular and the plural genitives, i.e. *murderer's* and *murderers'*.

As is shown in Table 5.7, among the left collocates of *murderer* in the whole corpora, we frequently find the terms *convicted*, *mass* and *serial* (*BNC*: 11 occurrences, 5.5%; *CWo*: 16 occurrences, 8%) used as pre-modifiers of the term. The last two of the above pre-modifiers suggest that the term *murderer* is used to indicate people who commit more than one murderous act (e.g. “When they think of mass *murderers*, they normally think of one person killing unlawfully a handful of other people”, *BNC*: CHL).

Table 5.7 Attributive left collocates of the term *murderer*: pre-modifiers in the two corpora and in the *BNC* and *CWo* datasets

Pre-modifiers of <i>murderer</i>	Frequency		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>alleged</i>	5.46	5.73	1 (0.5%)	1 (0.5%)
<i>convinte</i>	10.79	9.31	5 (2.5%)	7 (3.5%)
<i>mass</i>	8.21	8.00	5 (2.5%)	7 (3.5%)
<i>multiple</i>	6.36	5.97	1 (0.5%)	-
<i>notorious</i>	6.37	6.77	1 (0.5%)	2 (1%)
<i>serial</i>	7.87	8.33	1 (0.5%)	2 (1%)

Fig. 5.9.1 Screenshot of the pre-modifiers of *murderer* in the *BNC* dataset

26 HLS opposition demonstrations. Mass executions Eleven convicted < murderers > were hanged on Nov. 15 in what was claimed by human rights  
52 E9U Hospital between February and April 1991. The families of mass < murderer > Allitt's victims are also demanding that any inquiry be held  
53 HR9 look sullen. Especially when I am watching puerto Rican mass < murderers >. This one was now sobbing over his attorney's shoes. Quigley  
66 K3C lingers over the grisly reminiscences of a psychotic serial < murderer >? Again, bombarding the young day by day with fictitious  
76 ECU the polyhomicide acts as an aggrandized suicidal act. Serial < murderers > have long identified women in their sights, but the

Fig. 5.9.2 Screenshot of the pre-modifiers of *murderer* in the *CWo* dataset

37 usbooks example, are expected to be lower in societies where convicted < murderers > are executed than in societies where the murderers are given  
49 brbooks murdered in the next street, now it seemed that the convicted < murderer > had spent time in Eleanor's company. I'm sure it's the same  
60 times difficulties inherent in dealing with actual and potential mass < murderers >. The obsessive fear of creating new martyrs should be  
61 sunnow the fact that 54-year-old Shipman is Britain's biggest mass < murderer >. He is suspected of killing at least 146 people. Asbestos  
73 times serious testing of the patient. To understand what makes a serial < murderer > kill, it's important to know how all of us develop.

As Table 5.8 shows, other immediate left collocates of the term *murderer* which frequently occur in my dataset are determiners (*BNC*: 84 occurrences, 42%; *CWo*: 64 occurrences, 32%), possessive pronouns (*BNC*: 9 occurrences, 4.5%; *CWo*: 5 occurrences, 2.5%) and, as exemplified in Figg. 5.10.1 and 5.10.2, proper names or

nouns identifying victims in the genitive form<sup>22</sup> (*BNC*: 8 occurrences, 4%; *CWo*: 8 occurrences, 4%).

Table 5.8 Additional left collocates of *murderer(s)* in the *BNC* and the *CWo* datasets: determiners and possessive pronouns

Frequency		
	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<b>Determiners</b>	<b>84 (42%)</b>	<b>64 (32%)</b>
<i>an</i>	30 (15%)	20 (10%)
<i>the</i>	54 (27%)	44 (22%)
<b>Possessive pronouns</b>	<b>9 (4.5%)</b>	<b>5 (2.5%)</b>
<i>his</i>	8 (4%)	3 (1.5%)
<i>her</i>	1 (0.5%)	2 (1%)

Fig.5.10.1 Screenshot of left collocates of *murderer* in the *BNC* dataset: proper names

```

62 BP9 had been killed. Also, Pierre asked, wouldn't Angelica's < murderer > have been covered with blood? The whole toilet compartment
63 K95 that?' Cranston mumbled, his mouth full of food. 'Horne's < murderer > means the assassin knows us or why should he send such
65 H8A firmly. 'we want to hasten the discovery of Sir Thomas's < murderer >. Mr Didier, just in case the police think it's me,' Angelina
72 HL8 Republican Army]" and Libyan co-operation in bringing Fletcher's < murderers > to justice. In a speech to the General People's Congress
73 F58 overwhelming; and alongside them was a fierce desire to see Jay's < murderers > punished. Not so much the man who had killed him, though it

```

Fig.5.10.2 Screenshot of left collocates of *murderer* in the *CWo* dataset: proper names

```

4 usmags an called Amsterdam (Leonardo DiCaprio) against his father's < murderer >, an American protestant named William Cutting, also known as
5 usbooks washing over her. Convinced her she was walking beside Kitty's < murderer >, that it wasn't Henry, or James ... "Aren't you?" She halted
6 usbooks Black Trinity. But to find it, I---we ---will have to find Len's < murderer >. Whoever killed him took the pearls. If you help me find what
7 brbooks attacked by an armed raider, who shot and killed him. Pull's < murderer > was one of the last men to be hanged in England. Like Langdal
8 times progress. I wonder if the will is really there to see Stephen's < murderers > brought to justice. " US officials claim that the Greek

```

The noun *murderer* is also used as a head noun followed by the prepositional phrase *of* + proper names or nouns denoting people (*BNC*: 4 occurrences, 2%; *CWo*: 10 occurrences, 5%); these post-modifying noun phrases identify ordinary people as the victims of criminal acts (see Fig. 5.11), unlike what happens in the case of *assassin*, where they also denote victims who are important political figures (see section 5.2.1).

<sup>22</sup> On reference to the victim, see also about *murderer+ of* below.

Fig. 5.11 Screenshot of *murderer+ of* in the *CWo* dataset

85	brspok	TURKISH DETECTIVE KILLED Turkish police are hunting for the < murderers > of an ex-policeman who may have been killed by a left wing
86	usspok	States will work -- in Rwanda and elsewhere -- to ensure that the < murderers > of innocent people are arrested and tried for their
87	usspok	Hudson Institute, say the government would like to bring the < murderers > of journalists to justice. But most attacks are carried out
88	usbooks	terrorists for years, decades. They have supported, directly, the < murderers > of many hundreds of people. They are now protecting and
90	brnews	parallels in Hamlet with their own situation. Claudius, the < murderer > of Hamlet's father, is seen as Netanyahu, whose virulent

With regard to the immediate right collocates of the term, the data show that *murderer* can be followed by a noun only if it shows up in the genitive form (*BNC*: 4 occurrences, 2%; *CWo*: 3 occurrences, 1.5%), and not as a mere nominal pre-modifier. The nouns which follow *murderer* in the genitive construction make reference to the murderer's belongings or way of behaving (see Figg. 5.12.1 and 5.12.2).

Fig. 5.12.1 Screenshot of the right collocates of *murderer* in a genitive form in the *BNC* dataset: *murderer's* + noun

94	FBM	nauseating first chapter contains a graphic description of an axe < murderer >'s modus operandi ('You have to spell it out clearly, simply
95	A18	and Punishment , his switch from first-person narration - the < murderer >'s story - to what is formally third-person but proves so
96	FF0	was convenient to do so through a story of a murder and the < murderer >'s subsequent feelings and actions had nothing essential to
97	A18	sociological. The thrust of their loose end, as of details like the < murderer >'s yellow cubby-hole - coffinlike, his mother calls it - and

Fig. 5.12.2 Screenshot of the right collocates of *murderer* in a genitive form in the *CWo* dataset: *murderer's* + noun

8	usnews	the beginning of the song--in order--repeated by the women on < murderer >'s row? 20. what song, featured in "Catch Me If You Can," is
10	usbooks	daughter was murdered in a brutal assault at her college. At the < murderer >'s trial, a bird appeared in the window, making an awful
11	brbooks	tie anything we found to any killer. Even if we'd found the < murderer >'s wallet a sharp lawyer would make the case that his client

The word *murderer* can occur preceded or followed by the coordinating conjunction *and/or* (*BNC*: 16 occurrences, 8%; *CWo*: 31 occurrences, 15.5%), which is immediately followed or preceded, respectively, by another noun. Table 5.9 shows the nouns which most frequently occur in this kind of construction both with the singular word-form *murderer* and the plural word-form *murderers* within the whole *BNC* and *CWo* corpora as well as in my datasets. In the *BNC* dataset, the terms *criminals*, *enemy*, *mercenaries*, *outlaws*, *robbers*, *thieves* are instantiated in such a construction, but each one occurs only once. All the nouns that co-occur with *murderer(s)* in this patterns are negatively connoted because they represent people who are known to commit violent act against people or private properties.

Table 5.9 Nominal collocates of the term *murderer* in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
Nouns coordinated with <i>murderer</i> , by means of <i>and/or</i> , either preceding or following	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>drug dealer/traffickers</i>	-	-	3 (1.5%)	3 (1.5%)
<i>mugger</i>	8.00	6.45	2 (1%)	-
<i>rapist(s)</i>	10.09	10.11	8 (4%)	12 (6%)
<i>robber(s)</i>	7.84	7.34	1 (0.5%)	3 (1.5%)
<i>spies</i>	-	-	-	1 (0.5%)
<i>terrorist(s)</i>	-	5.48	-	9 (4.5%)
<i>thug(s)</i>	-	6.46	-	4 (2%)

Fig. 5.13.1 Screenshot of *murderer* in the *BNC* dataset: *and/or*

2 HHV Should not the money be used instead to deal with rapists, < murderers > and muggers in my constituency and in other parts of Greater  
5 K1M national prison housing amongst others robbers, rapists and < murderers > serving lengthy sentences.. Three-times monthly visits from  
10 C9B fleas and snakes are destroyed, and robbers and < murderers > are killed, in order to protect the lives of others. It might  
28 G3E 's shoe, the trace which might prove a man was a rapist or a < murderer >. And yet, she had thought, he doesn't really mind. All he cares  
54 FAN Japanese; and Aung San, previously portrayed by Dorman-Smith as a < murderer > and an enemy. Confusion over the Murder Charge Now he

Fig. 5.13.2 Screenshot of *murderer* in the *CWo* dataset: *and/or*

1 usnews a rapist, a gay basher or just a run-of-the-mill mugger or < murderer>." The lawsuit, filed Monday in U. S. District Court, targets  
3 brbooks there had been child molesters, there had been < murderers > and rapists and abductors. Fewer, perhaps. Less remorselessly  
4 brbooks fascinating effect. I suppose I can understand now why < murderers > and rapists are able to behave in the way they do: if I, a  
6 brnews the flat, but refused to speculate as to whether the < murderer > or murderers were known to their victims. He said: 'we are  
9 brbooks the grim walls of Bedford Jail without thinking of all the < murderers > and robbers locked up inside -- creatures who constituted a

Among the terms denoting criminals that collocate with *assassin* are those identifying perpetrators of sexual violence such as child molesters or rapists are (*BNC*: 14 occurrences, 7%; *CWo*: 3 occurrences, 1.5%; e.g. “Recent feminist work has focused not only on the continuum of male sexual violence, but also on the serial rapists and sex *murderers*”, *BNC*: CSI; “[...] in future I intend to exercise my discretion so that murderers of police or prison officers, terrorist *murderers*, sexual or sadistic *murderers* of children and *murderers* by firearm in the course of robbery can normally expect to serve at least 20 years in custody; [...]”, *BNC*: FDV). Moreover, the term *murderer* co-occurs 3 times in the *BNC* (1.5%) and 9 times in the *CWo* (4.5%) with reference to terrorist groups, and, when these are the terrorists who attacked the twin towers on September 11<sup>th</sup> 2001, they are similarly qualified as *murderers* (e.g. “But the victims of September 11th did die in vain, which is why we pursue their *murderers*”, *CWo*: usmags; “[...] suicide bombers they are martyrs. They’re not martyrs. They’re

*murderers*”, *CWo*: times; “[t]errorists responsible for atrocities, mass *murderers* and child killers will serve their entire lives in jail”, *CWo*: times; “[u]nder Islamic law, rapists, *murderers* and drug traffickers are beheaded in public”, *BNC*: CEM; “Asked if the death penalty should be brought back for terrorist *murderers*, a majority supported its return”, *BNC*: CFH).

As in the case of *assassin*, my datasets show that term *murderer*, too, can occur as the subject or as the object of a predicate. Table 5.10 shows the verbs that occur in the former pattern in the *BNC* and the *CWo* (*BNC*: 50 occurrences, 25%; *CWo*: 37 occurrences, 18.5%); whereas *to kill* is used to refer to the action a murderer commits, the other verbs describe the movements he or she may make (*to walk*, *to leave*, *to hide*) or the way in which he/she uses his/her weapons (*BNC*: 1 occurrence on the use of guns and 1 on shooting; *CWo*: 1 occurrence on the use of weapons and 3 occurrences referring on shooting; e.g. “Had a moment’s warning the *murderer* was behind him and grabbed the only weapon to hand”, *CWo*: usbooks). Other verbs which occur only once in my datasets can be classified under the same category, in the sense that they denote the actions that the referent of *murderer* carries out (e.g. *to beg*, *to commit a crime*, *to declare their innocence*, *to disappear*, *to end up*, *to escape*, *to get*, *to go away*, *to go down*, *to keep*, *to identify*, *to make a mistake*, *to reach*, *to say*, *to serve in prison*, *to shoot*, *to spend*, *to stand*, *to strike*, *to take [care]*, *to turn out*, *to watch*) or which he/she is affected by (e.g. *to be allowed*, *to be arrested*, *to be beaten up*, *to be caught*, *to be discovered*, *to be executed*, *to be released*, *to be seen*, *to be sent*).

Table 5.10 Right collocates of *murderer*:  
*murderer* as the subject of predicates  
in the two corpora and in the *BNC* and *CWo* datasets

<b><i>Murderer as the subject</i></b>	<b>Frequency</b>			
	<b><i>BNC</i> (pmw)</b>	<b><i>CWo</i> (pmw)</b>	<b><i>BNC dataset</i> (tokens and percentages)</b>	<b><i>CWo dataset</i> (tokens and percentages)</b>
<i>to hide</i>	3.61	3.10	2 (1%)	1 (0.5%)
<i>to kill</i>	2.24	2.16	6 (3%)	4 (2%)
<i>to leave</i>	1.41	0.91	3 (1.5%)	2 (1%)
<i>to walk</i>	2.55	2.13	2 (1%)	1 (0.5%)

When *murderer* occurs as the object of a predicate (*BNC*: 40 occurrences, 20%; *CWo*: 52 occurrences, 26%), this describes the murderers as prey, convicts or people in hiding.

Other predicates, which are instantiated only once or twice in the sentences in my datasets, indicate actions carried out by courts of justice or authorities (e.g. *to allow, to arrest, to bring to justice, to deter, to execute, to find, to free, to hang, to put/sentence to death, to persecute, to release, to seek, to shelter, to support, to trap*).

The noun *murderer* also occurs as the object of the preposition, which can occur in an embedded structure (*BNC*: 54 occurrences, 27%; *CWo*: 48 occurrences, 24%). In such cases, the most frequent prepositions occurring in my datasets are *of, by, to, about, with, on for, from* (see Figg. 5.14.1 and 5.14.2).

Table 5.11 Left collocates of *murderer*:  
*murderer* as the object of predicates  
in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
<b>Murderer as the object</b>	<b>BNC (pmw)</b>	<b>CWo (pmw)</b>	<b>BNC dataset (tokens and percentages)</b>	<b>CWo dataset (tokens and percentages)</b>
<i>to catch</i>	4.39	4.35	3 (1.5%)	1 (0.5%)
<i>to hunt</i>	5.64	5.16	2 (1%)	3 (1.5%)
<i>to punish</i>	6.47	3.79	2 (1%)	1 (0.5%)
<i>to release</i>	3.64	2.61	1 (0.5%)	1 (0.5%)
<i>to unmask</i>	8.25	6.36	-	-

Fig. 5.14.1 Screenshot of *murderer* in a prepositional phrase in the *BNC* dataset

```

38 CBC interviewed my fair share of criminals and oddballs, ranging from < murderers > to men with silver foil round their heads claiming protection
67 H9C you broke her neck. Quite simple, I understand, for a skilled < murderer >. A matter of touch, of knowing where to hold and quickly turn
96 CKB The game was over. Ginny never learned the identity of the < murderer >. In the real world, the safe world of the present, with
99 A1B Meditations in Time of Civil War: 'vengeance upon the < murderers >', the cry goes up, 'vengeance for Jacques Molay'. In cloud-pale
106 H9C laughed abruptly, 'did they just offer their throats to the < murderer >?' He turned. 'do you make any sense of it, Ranulf?'

```

Fig. 5.14.2 Screenshot of *murderer* in a prepositional phrase in the *CWo* dataset

```

19 oznews Australian. The Advertiser 2 - Metro Don's life was torn apart by a < murderer > but happy memories help him cope JUST days before she was
36 usspok as authorities investigate possible links between the alleged < murderer > and international terrorist groups. There have been
41 brregnews considering extending minimum starting point guidelines for < murderers > from England and Wales into Northern Ireland. Safety fear
64 times money from murdered children by staging a drama about the Moors < murderer > Myra Hindley. Winnie Johnson, whose son Keith Bennett,
95 usmags substantiate clues and provides them with a better description of the < murderer > and how he felt at the time, according to Montz.

```

In my datasets, the term *murderer* is frequently instantiated as the subject complement (*BNC*: 18 occurrences, 9%; *CWo*: 25 occurrences, 12.5%), that is in sentences where people are identified and marked them as murderers (e.g. “[...] the trace which might prove a man was a rapist or a *murderer*”, *BNC*: G3E; “I was not her *murderer*”, *CWo*: usbooks).



To sum up, the data analysed show that the term *murderer* is frequently instantiated in fiction in the *BNC*, and both in fiction and the spoken language in the *CWo*. The noun *murderer* is preceded by attributive adjectives such as *convicted* and *serial*, which indicate that the term is used to portray people who have committed multiple murders. According to the data analysed, the noun *murderer* is also preceded by proper names and common nouns in their genitive forms or followed by the preposition *of* + NPs to identify the victims of the criminal acts; these tend to be ordinary people rather than important political figures. *Murderer* can also be followed by the 's genitive form and a noun referring to the murderer's belongings or behaviour. Finally, *murderer* is also coordinated (*and/or*) with nouns such as *rapist* and *terrorist*, which identify other kinds of serious crime offenders.

With reference to the colligational patterns of *murderer*, my datasets show that the term is likely to occur as the subject or the object of a predicate, in a prepositional phrase and as subject complement. When occurring as the subject of a predicate (*BNC*: 25%; *CWo*: 18.5%), it can be followed by verbs which indicate the use of weapons for the purpose of killing people or which indicate the movements and actions the murderer carries out. When the noun occurs as the object of a predicate (*BNC*: 20%; *CWo*: 26%), it can be preceded by verbs which signal that the criminals are hunted down by police officers. When the term occurs in a PP (*BNC*: 27%; *CWo*: 24%), *assassin* is preceded by prepositions such as *of*, *by*, *to*, *about*, *with*, *on for*, *from*. Finally, when the noun occurs as a subject complement (*BNC*: 9%; *CWo*: 12.5%), it is used to identify people and mark them as murderers.

### 5.2.3 Killer

The examination of the term *killer* reveals that it occurs more frequently in American English than in British English; in particular, as Table 5.12 shows, the term occurs less than 2,000 times (17.68 pmw) in the *BNC* but more than 14,000 (31.35 pmw) in the *COCA*, and in both corpora the singular word-form is almost 4 times as frequent as its plural counterpart. In addition, the *BNC* shows that the term *killer(s)* has a high frequency score in fictional genres (about 26%) and an even higher frequency score in newspapers (about 38%); the *COCA* shows higher frequencies of occurrence of the term in the spoken register (about 33%), and to a lesser extent, in fiction (about 22%).

Table 5.12 Frequency of occurrence of the term *killer* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Word-forms</b>				
<i>Killer</i>	<b>1,341 (78.79%)</b>	13.93	<b>11,360 (78.03 %)</b>	<b>24.48</b>
<i>Killers</i>	361 (21.21%)	3.93	3,198 (21.97 %)	6.93
<b>Genres</b>				
Spoken	57 (3.35%)	5.72	<b>4,806 (33.01%)</b>	<b>50.28</b>
Fiction	<b>441 (25.91%)</b>	<b>27.72</b>	<b>3,298 (22.66%)</b>	<b>36.47</b>
Magazine	151 (8.87%)	20.79	2929 (20.12%)	30.65
Newspaper	<b>652 (38.31%)</b>	<b>62.29</b>	2,768 (19.01%)	30.18
Non-academic	160 (9.40%)	9.70	-	-
Academic	74 (4.35%)	4.83	757 (5.20%)	8.31
Other	167 (9.81%)	8.02	-	-
<b>TOTAL</b>	<b>1,702 (100%)</b>	<b>17.68</b>	<b>14,558 (100%)</b>	<b>31.35</b>

Attributive adjectives and nouns are often used as pre-modifiers of the noun *killer*, as Table 5.13 displays. One may notice that, apart from *serial*, these pre-modifiers are present with low percentages in my datasets; however, they all share a common feature, that is they describe the killer. In addition, the concordances show that, when *killer* is modified by nouns, the pre-modifier may refer to a human agent or to an inanimate instrument (see Fig. 5.15.1 and 5.15.2). In the former case, the pre-modifier more technically identifies a sub-category of killers, while in the latter it identifies the entity affected by the action of the killing agent.

Table 5.13 Attributive left collocates of *killer*:  
pre-modifiers in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
Adjectival pre-modifier of <i>killer</i>	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>alleged</i>	6.23	7.25	1 (0.5%)	1 (0.5%)
<i>convicted</i>	8.09	-	3 (1.5%)	-
<i>loyalist</i>	7.37	5.68	1 (0.5%)	1 (0.5%)
<i>notorious</i>	6.38	7.50	1 (0.5%)	2 (1%)
<i>psychopathic</i>	8.68	7.40	1 (0.5%)	1 (0.5%)
<i>serial</i>	10.62	10.35	8 (4%)	12 (6%)
<i>silent</i>	-	6.64	-	3 (1.5%)
<i>would-be</i>	7.72	7.22	1 (0.5%)	-
<b>Nominal pre-modifiers of <i>killer</i></b>				
<i>contract</i>	-	4.88	3 (1.5%)	1 (0.5%)
<i>cop</i>	7.92	6.84	-	-
<i>Ira</i>	6.51	6.69	3 (1.5%)	-
<i>moss</i>	7.29	5.94	-	1 (0.5%)
<i>pain</i>	-	6.14	4 (2%)	4 (2%)
<i>weed</i>	7.61	8.29	2 (1%)	1 (0.5%)

Fig. 5.15.1 Screenshot of *killer* in the *BNC* dataset: pre-modifiers

50 HJ4 round-the-clock protection for fear he could be in the sights of an IRA < killer > gang. Thieves stole a car from a house last month in  
64 CBT chance of catching the whistler, his theories about psychopathic < killers >, his own experience of serial murder.  
72 CD5 and portraits of the most infamous mass murderers and serial < killers >. These feature informative biogs detailing the number  
105 CLT as I have done for tubifex. However, following use of a weed < killer > on the lawn, earthworms should not be collected for  
110 BLX direction and dash rapidly forward, leaving the frustrated would-be < killers > snapping at empty water. On this butterfly fish

Fig. 5.15.2 Screenshot of *killer* in the *BNC* dataset: pre-modifiers

16 bregnews Italian authorities have been investigating whether the alleged < killers > were part of a wider network of Satanists, according  
45 sunnow involving 189 claims in the last two years, include Loyalist < killer > Trevor McKeown and murderer John Murdock, left. The family  
48 sunnow lawn. If moss is already widespread, rake it out first. Moss < killers > just burn off the top of thick moss. On all except the  
51 times Frances Gibb reports on a law lords case that could give notorious < killers > their freedom The Home Secretary should be stripped of  
63 usnews suspect in the 1970s, adds: "This man is like no other serial < killer > that ever existed. He doesn't fit a profile --- the

As Table 5.14 shows, other left collocates of the term *killer* in my datasets are determiners (*BNC*: 56 occurrences, 28%, *CWo*: 68 occurrences, 34%) and possessive pronouns (*BNC*: 14 occurrences, 7%, *CWo*: 11 occurrences, 5.5%); therefore, most of the time, the noun phrase with *killer* as its head is definite and thus presented as known to the audience. In addition, as Figg. 5.16.1 and 5.16.2 show, *killer* is also immediately

preceded by proper names or nouns in the genitive form (*BNC*: 13 occurrences, 6.5%; *CWo*: 12 occurrences, 6%), which identify people who are victims of killers (the same pattern characterises the term *assassin*; see section 5.2.1).

Table 5.14 Additional left collocates of *killer(s)* in the *BNC* and *CWo* datasets: determiners and possessive pronouns

	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<b>Determiners</b>	<b>56 (28%)</b>	<b>68 (34%)</b>
<i>an</i>	10 (5%)	12 (6%)
<i>the</i>	46 (23%)	56 (28%)
<b>Possessive pronouns</b>	<b>14 (7%)</b>	<b>11 (5.5%)</b>
<i>his</i>	12 (6%)	7 (3.5%)
<i>her</i>	2 (1%)	4 (2%)

Fig. 5.16.1 Screenshot of *killer* in the *BNC* dataset: 's-genitive pre-modifier

```

62 CBF      for murder victim's clothes POLICE hunting a businessman's < killer > appealed yesterday for help to find his clothes. Company
63 CHZ      curse on strangler A SOBBING mother cursed her daughter's < killer > yesterday as she collapsed at the graveside of pop fan Helen
64 KLE      death has made an appeal for help in the hunt for her father's < killer >. She said he didn't have an enemy in the world and that his
65 K4M      October, 1988, when police were still searching for Mr Galvin's < killer >. He was also wanted in connection with another matter
70 K5M      anybody's child next time.' Detectives searching for Paul's < killers > are sifting through more than 2,000 different leads they have

```

Fig. 5.16.2 Screenshot of *killer* in the *CWo* dataset: 's-genitive pre-modifier

```

2 sunnow   hunted; Exclusive 09 January 2000 POLICE hunting Jill Dando's < killer > want to question an obsessed fan who vanished soon after her
3 sunnow   Chief Supt Andy Murphy, leading the hunt for the French girl's < killer >, yesterday stressed the man was being sought purely as a 'very
4 brmags   unheard John Lennon songs to feature in new musical John Lennon's < killer > wanted to 'steal' fame John Lennon art exhibition launched
7 usbooks   out? (was the phone ripped out?) Or one of them, his mother's < killers >, waiting in the darkened kitchen by the phone? waiting to kill
8 brregnew to Gerry Adams and Martin McGuinness suggesting that Robert's < killers > should do the patriotic thing and be answerable for their

```

The immediate right collocates of the term *killer* can occur either in a genitive construction (*killer's/killers' + noun*) or in a nominal compound (*killer + noun*). The former construction is instantiated only once in my datasets: “Crandy’s death was the first from a tornado in Utah. *Killer’s* gold Chris Pharo 13 August 1999 Mandy murderer left necklace clue THE ruthless woman suspected of killing bisexual mum Mandy Power and her family left a gold necklace at the murder scene, police revealed yesterday” (*CWo*: sunnow). The latter construction is more frequent (*BNC*: 29 occurrences, 14.5%; *CWo*: 31 occurrences, 15.5%), and is exemplified by such phrases as *killer bee*, *killer disease*, *killer whale* (see the examples in Figg. 5.17.1 and 5.17.2 and Table 5.15). In such nominal compounds *killer* means ‘that kill(s)’ and identifies:

- dangerous animals/wildlife, which kill their prey, and products which negatively affect plants (*BNC/CWo*: 2.5%; e.g. “Moss *killers* just burn off the top of thick moss”, *CWo*: sunnow; “It’s not, that’s a lawn food and weed *killer*”, *BNC*: KCN);
- diseases such as heart disease, TB, cancer, tetanus, measles and viruses like SARS and meningitis, or, in one case, also a symptom of severe discomfort, namely stress (*BNC*: 11 occurrences, 5.5%; *CWo*: 17 occurrences, 8.5%). In such cases, *killer* can be either a pre-modifier or the noun denoting the disease or the subject complement (e.g. “Cancer-related diseases are the next biggest *killers*. Lung cancer kills 13 per cent of men [...]”, *BNC*: K5M; “Dr Hugh-Jones is extremely concerned that *killer* diseases like TB now seem to be on the increase again amongst homeless people”, *CWo*: brephem; “Well, stress kills people. It’s a big *killer*. Stress kills”, *BNC*: KM5);
- natural disasters (*BNC*: 1 occurrence, 0.5%; *CWo*: 5 occurrences, 2.5%; e.g. “[...] to raise money for victims of the *killer* tsunami”, *CWo*: oznews; “Bad nasty storm coming.’ The *killer* hurricane hit Barbados at 145 miles per hour [...]”, *BNC*: FRS).

Fig. 5.17.1 Screenshot of *killer* followed by nouns in the *BNC* dataset

36	K35	said sweeping reform of the constitution was needed. MEXICO < killer > bees on attack A SWARM of killer African bees attacked some
40	GLX	three hundred years? Or disappear forever wiped out by some < killer > bug and all that will be left of them is a film in which they
42	HWS	that do not possess mature T lymphocytes but do have natural < killer > cell activity. As cimetidine has been shown to enhance natural
88	CEM	bug ARNOLD Schwarzenegger's wife is being treated for the < killer > virus meningitis. The multi-millionaire actor has maintained
98	ABC	rounded up a mixed herd of 50 to 60 bottlenose dolphins and false < killer > whales ( Pseudorca ), and drove them through the narrow

Fig. 5.17.2 Screenshot of *killer* followed by nouns in the *CWo* dataset

41	usmags	body more quickly unleash its front-line defense -- natural < killer > cells. Bad gas: Peru's Camisea natural gas reserves promise
43	brephem	risen alarmingly. Dr Hugh-Jones is extremely concerned that < killer > diseases like TB now seem to be on the increase again amongst
54	oznews	whose usual solidity in defence combined with an unexpected < killer > instinct in attack. As impressively as Steve McClaren's team
100	safrica	workers have said they feel inadequately protected from the < killer > virus, whose symptoms include headaches, internal bleeding,
107	sunnow	the town Barry Roche 18 June 2001 (Photograph) - THE three < killer > whales cavorting in Cork Harbour fancied seeing some-fin new

Table 5.15 Right collocates of *killer*:  
*killer* as a pre-modifier of nouns  
in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
<b>Killer as pre-modifier of nouns</b>	<b><i>BNC</i></b>	<b><i>CWo</i></b>	<b><i>BNC dataset</i></b>	<b><i>CWo dataset</i></b>
	<b>(pmw)</b>	<b>(pmw)</b>	<b>(tokens and percentages)</b>	<b>(tokens and percentages)</b>
<b>Animals/wildlife</b>				
<i>bee(s)</i>	6.80	6.99	1 (0.5%)	-
<i>bug</i>	6.93	7.28	2 (1%)	-
<i>whale(s)</i>	9.52	9.60	7 (3.5%)	1 (0.5%)
<b>Diseases</b>				
<i>cell</i>	4.20	5.80	2 (1%)	1 (0.5%)
<i>disease(s)</i>	5.28	5.95	1 (0.5%)	2 (1%)
<i>virus</i>	5.99	6.10	1 (0.5%)	2 (1%)
<i>heart disease, TB, cancer, tetanus, measles, SARS, meningitis, stress</i>	-	-	7 (3.5%)	10 (5%)
<b>Natural disasters</b>				
<i>instinct</i>	6.73	8.87	-	3 (1.5%)
<i>tsunami</i>	-	6.77	-	2 (1%)
<i>wave</i>	-	6.86	-	2 (1%)

When the term *killer* is used as the subject of predicates (*BNC*: 42 occurrences, 21%; *CWo*: 40 occurrences, 20%), these encode the action of killing people (e.g. “Perhaps Darnley’s *killers* wrote these stories about Bothwell, before they killed Darnley [...]”, *BNC*: FRD; “*Killers* prey on people at random”, *CWo*: usbooks) or other deliberate material actions (e.g. “So it pays the nest-owners to sit tight and hope that the *killers* will pass them by”, *BNC*: BLX; “Her *killer* was whisked to Libya under diplomatic immunity”, *CWo*: sunnow). In addition to these verbs, in my datasets the predicates used in this pattern – some of them instantiated only once – include, for instance, *to shout*, *to commit murder*, *to open fire*, *to finger*, *to want to steal*, *to walk*, *to pounce*, *to shoot* but also verbs of communication (e.g. *to write*, *to hear*, *to speak*) and a few experiential verbs (e.g. *to go to sleep*, *to think*, *to know*); finally, the verb *to be* is also used, and it introduces a characteristic describing the referent of the killer (e.g. *to be tall*, *to be shy*).

Table 5.16 Right collocates of *killer*:  
*killer* as the subject of predicates  
in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
<i>Killer as the subject</i>	<i>BNC (pmw)</i>	<i>CWo (pmw)</i>	<i>BNC dataset (tokens and percentages)</i>	<i>CWo dataset (tokens and percentages)</i>
<i>to attack</i>	-	-	4 (2%)	1 (0.5%)
<i>to escape</i>	4.06	4.09	1 (0.5%)	1 (0.5%)
<i>to lurk</i>	6.55	5.43	-	-
<i>to murder</i>	4.37	5.64	6 (3%)	4 (2%)
<i>to prey</i>	6.29	5.37	-	2 (1%)
<i>to stab</i>	5.16	5.16	2 (1%)	1 (0.5%)
<i>to strike</i>	4.44	4.42	1 (0.5%)	1 (0.5%)
<i>to target</i>	3.73	4.79	-	3 (1.5%)

The noun *killer* occurs as the object of predicates such as those given in Table 5.17 (*BNC*: 53 occurrences, 26.5%; *CWo*: 41 occurrences, 20.5%), which are used to define actions that are normally carried out by a law-enforcing authority (such as the police) against crimes. Other verbs which occur in this pattern – and which are instantiated only once in my datasets – are *to capture*, *to look/search for*, *to miss*, *to identify*, *to protect*, *to release*, *to nail*, *to search for*, *to free*, *to accuse*, *to trap*, *to curse*, *to approach*, *to condemn*, *to fire*, *to trip up*, *to pledge*, which indicate actions carried out by police or justice courts.

Table 5.17 Left collocates of *killer*:  
*killer* as the object of predicates  
in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
<i>Killer as the object</i>	<i>BNC (pmw)</i>	<i>CWo (pmw)</i>	<i>BNC dataset (tokens and percentage)</i>	<i>CWo dataset (tokens and percentage)</i>
<i>to catch</i>	5.44	5.92	2 (1%)	1 (0.5%)
<i>to find</i>	-	-	6 (3%)	4 (2%)
<i>to hunt</i>	9.19	7.88	5 (2.5%)	4 (2%)
<i>to jail</i>	7.55	7.07	4 (2%)	1 (0.5%)
<i>to bring to justice</i>	-	-	2 (1%)	4 (2%)

The noun is also instantiated as in prepositional phrases, that is as the object of a preposition, which are part of embedded structures (*BNC*: 36 occurrences, 18%; *CWo*: 44 occurrences, 22%). The most frequent prepositions occurring in this structure in my datasets are *of, by, in, to, against, with, on, for, from* (see Figg. 5.18.1 and 5.18.2).

Fig. 5.18.1 Screenshot of *killer* in a prepositional phrase in the *BNC* dataset

```

33 ABK      him sacked (for 'human-rights violations' against his hired < killers >), and a cast-iron pardon from the assembly. That might give
48 CSI     widow continued to deny that her husband had had sex with his < killer >'s 15-year-old daughter, and the dead man's brother spoke of
87 CBF     the new year. Olive has received 450 letters from the < killer >. 'He has talked about it before and nothing came of it
95 CBJ     the densely-populated riverside parishes; the victims of the < killer > - or rather the remains - were laid to rest in a corner of
107 GOL    Nothing on this scale. This was an organization of trained < killers >. Hundreds, maybe even thousands of them. And he knew the

```

Fig. 5.18.2 Screenshot of *killer* in a prepositional phrase in the *CWo* dataset

```

41 usnews   including Missouri, have a death penalty for juvenile < killers >, only Texas, Oklahoma and Virginia have carried out such executions
47 sunnow  sight, there is something repulsive in our pictures of mass < killer > Michael Stone as he lives his new life in Spain. That's
62 canews  (AP) - Connecticut again set an execution date for serial < killer > Michael Ross but legal manoeuvring could once more derail the
85 usnews  Perry to explain his disappearance, were actually sent by the < killers >. DiTommaso's fingerprints and DNA were also found on the letters
88 safrica workers have said they feel inadequately protected from the < killer > virus, whose symptoms include headaches, internal bleeding,

```

In my datasets, the term *killer* is also frequently instantiated as the subject complement (*BNC*: 19 occurrences, 9.5%; *CWo*: 13 occurrences, 6.5%), that is in sentences where given people are identified as killers.

Fig. 5.19.1 Screenshot of *killer* as subject complement in the *BNC* dataset

```

15 CN3     'You should be armed,' Tweed decided. 'Manescu is a < killer >. The file records he's strangled people with his bare hands
34 FRJ     what?' 'They're thugs, Eric. These Tans are just hired < killers >. That can't be news to you, man. You've had plenty of incidents
52 BLX     shock tactic is to create the impression that you are a lethal < killer > yourself, and not to be trifled with. Many a harmless animal
82 FBN     stressed. Incorrect diagnosis, not the virus itself, can be the < killer >, since Fish Pox can be mistaken for bacterial, fungal or protozoan
101 H9Y     stand up when they'd just sat down. 'The boxes were the < killers >. The Fortune is a tiny gem of a theatre where the boxes sit

```

Fig. 5.19.2 Screenshot of *killer* as subject complement in the *CWo* dataset

```

20 usbooks  as he commented on Cooper and Crèvecoeur: The hunter is a < killer >. The husbandman, on the other hand, brings about birth and
33 usnews   reserves for terrorists, Bush added, "He's a cold-blooded < killer >. This is a man who I believe strongly thinks he can use terrorist
44 brmags  another lot who haven't had it. Heart disease is an important < killer >, linked to a string of avoidable risks from cigarettes to
91 usbooks His ears are gorged so that he cannot speak clearly. He is the < killer > so that he can be the recorder and use language effectively
100 brbooks was unhelpful, and the police were almost certain he was the < killer >. He was known to have a terrible temper, especially when anyone

```

To sum up, the instances of the term *killer* collected show that this term occurs with a quite frequently in newspapers in the *BNC* and in the spoken language in the *CWo*, although their frequency of occurrence scores in fiction is rather high in both corpora. The noun *killer* appears to display a negative connotation, that is when reference is being made to someone/something that has killed or kills; none of the instances taken into consideration seems to show *killer* as having a positive connotation. With reference to the pre-modifiers of the noun, this can be preceded by attributive



adjectives or by common nouns in their genitive form: the former are used to describe the behavioral patterns of the killer him/herself; the latter identify particular *kinds* of killer; that is, the former are descriptive, the latter classifying in function. In addition, the term can also be preceded by proper names and common nouns with the 's genitive marking the victims of a killer. Finally, the term appears to be followed by common nouns which identify animals, diseases or natural phenomena which are dangerous and can kill (e.g. *killer whale, killer disease, killer tsunami*).

With regard to the colligational patterns of the term, *killer* appears to occur with high frequency scores as the subject or the object of a predicate, in a prepositional phrase and as subject complement. In particular, when the noun occurs as the subject of a sentence (*BNC*: 21%; *CWo*: 20%), it is instantiated with communication or experiential verbs and also with verbs that describe the action of killing someone. When the term occurs as the object of predicates (*BNC*: 26.5%; *CWo*: 20.5%), the verbs that are instantiated mainly indicate actions carried out by authorities to capture of the criminal. When the noun occurs in a prepositional phrase (*BNC*: 18%; *CWo*: 22%), it is preceded by prepositions such as *of, by, in, to, against, with, on for, from*. Finally, *killer* can occur as a subject complement (*BNC*: 9.5%; *CWo*: 6.5%), which is used to identify people and mark them as killers.

#### **5.2.4 Conclusion**

In conclusion, the analysis of the occurrences of the three nouns *assassin, murderer* and *killer* reveals similarities and differences in the frequencies and patterns of occurrence of these terms. First of all, the terms have different frequency scores, that is *assassin* occurs almost 5 times, *murderer* 10 times *killer* between 17 and 31 times pmw, respectively; and for all of them, the highest frequency of occurrence is in fiction.

Moreover, as Table 5.18 displays, although the three nouns appear to have a similar denotation (i.e. 'someone who kills someone else'), they occur in different patterns: with regard to their pre-modifiers, the three near-synonyms have a high frequency of occurrence with the adjective *alleged*; instead, *assassin* and *killer* are more likely to be modified by *would-be*; and finally, *murderer* and *killer* are modified by the adjectives *serial* and *notorious*. With regard to verbs collocating with the near-synonyms, *to kill* is used as a predicate with both subjects: *assassin* and *murderer*; and the verbs *to catch* and *to hunt* are used as predicates of *murderer* and *killer* occurring as objects. Finally, only *killer* often appears to be used in the medical field to identify

diseases so serious that can kill, and only *murderer* seems to occur to identify someone that kills people after having done sexual violence to them. We may also notice that the three nouns can occur in the same colligational patterns, that is as the subject or the object of a predicate, in a prepositional phrase and as subject complements.

If we compare the findings from the entire corpora with those from my datasets, we will notice that the former show recurring patterns which are not instantiated in the latter. More specifically, whereas the whole corpora highlight a high recurrence of a pattern, the same pattern does not necessarily also occur in my datasets.

Table 5.18 Colligational and semantic patterns of *assassin*, *murderer* and *killer*

		<i>Assassin</i>	<i>Murderer</i>	<i>Killer</i>
<b>Pre-modifiers</b>	<i>alleged</i>	x	x	x
	<i>notorious</i>	-	x	x
	<i>serial</i>	-	x	x
	<i>would-be</i>	x	-	x
<b>Predicates</b>	<i>to kill</i>	x	x	-
	<i>to catch</i>	x	x	x
	<i>to hunt</i>	-	x	x
<b>Victims</b>	important political figures	x	-	-
	common people	x	x	x
<b>Meanings/Use</b>	To kill someone	x	x	x
	Reference to diseases that cause death	-	-	x
	To kill after sexual violence	-	x	-

Therefore, as summarized in Table 5.18, the data analysed seem to show that the three nouns are not completely interchangeable, since there are some contexts in which one noun is preferred rather than the others; in particular, although the three of them are used to indicate someone who kills, only *killer* can be used with reference to animals and diseases, while only *murderer* occurs when referring to sexual violence that kills. Finally, the three terms are used when the victims are common people, but only *assassin* can be used when the victims are important political figures.

### 5.3 The verbs

In the following sub-sections I will report on the findings relevant to the verbs *to disturb* and *to bother*.<sup>23</sup> Every verb will be discussed with regard to its raw and normalised frequency of occurrence both as a lexeme and in its various word-forms, and its distribution across genres both in the *BNC* and the *COCA*. Then I will examine my datasets from the *BNC* and *CWo* to identify the terms' main colligational and semantic patterns.

#### 5.3.1 *To disturb*

The verb *DISTURB* appears to occur more frequently in British English than in American English; more specifically, as Table 5.19 displays, although the number of tokens in the *BNC* is lower than that in the *COCA* (1,992 times in the former and 6,130 times in the latter), the normalised frequency of occurrence is 7 times higher in the *BNC* than in the *COCA* (20.69 times pmw in the *BNC* vs 13.20 times pmw in the *COCA*).

Of all the word-forms of the verb *DISTURB* occurs, the simple past/past participle form *disturbed* is the most frequent one (see Table 5.19). The frequency hierarchy of the various forms of *DISTURB* in the *BNC* dataset is as follows: *disturbed* (56.5%) > *disturb* (15.5%) > *to disturb* (14%) > *disturbing* (10.5%) > *disturbs* (3.5%); the frequency hierarchy in the *CWo* dataset is as follows: *disturbed* (58.5%) > *disturb* (18.5%) > *disturbing* (12%) > *to disturb* (11%) > *disturbs* (5%).

In addition, Table 5.19 also shows that both in the *BNC* and in the *COCA* the term *DISTURB* occurs most frequently in fiction (773 and 2,329 times, respectively). In this case as well, the normalised frequency score is higher in the *BNC* than in the *COCA*, although there are fewer tokens of the term in the former than in the latter.

In my datasets, the verb *DISTURB* mostly occurs in declarative clauses, especially affirmative (*BNC*: 74%; *CWo*: 80.5%), but also negative (*BNC*: 23.5%; *CWo*: 17.5%); interrogative clauses are a minority, whether positive or negative (*BNC*: 2.5%; *CWo*: 2%). In addition, the verb is more frequently used in the active (*BNC*: 70%; *CWo*: 55%) than the passive voice (*BNC*: 30%; *CWo*: 45%).

---

<sup>23</sup> Unless otherwise specified, the forms *DISTURB* and *BOTHER* in capital letters will refer to the lexemes, independently of their variant inflectional realizations, while the forms *to disturb* and *to bother* will refer only to the *to*- infinitive forms of the verbs.

Table 5.19 Frequency of occurrence of the term *DISTURB* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Word-forms</b>				
<i>Disturb</i>	593 (29.76%)	6.16	1,800 (29,37%)	3.88
<i>Disturbs</i>	77 (3.86%)	0.80	490 (7,99%)	1.06
<i>Disturbed</i> <sup>24</sup>	<b>1,116 (56.02%)</b>	10.88	<b>3,151 (51,40%)</b>	6.79
<i>Disturbing</i> <sup>25</sup>	207 (10.36%)	2.15	689 (11.24%)	1.48
<b>Genres</b>				
Spoken	90 (4.52%)	9.03	830 (13.54%)	8.68
Fiction	<b>773 (38.80%)</b>	<b>48.59</b>	<b>2,329 (37.99%)</b>	<b>25.75</b>
Magazine	133 (6.68%)	18.31	1,107 (18.06%)	11.58
Newspaper	133 (6.68%)	12.71	806 (13.15%)	8.79
Non-academic	250 (12.55%)	15.16	-	-
Academic	256 (12.85%)	16.70	1,058 (17.26%)	11.62
Other	357 (17.92%)	17.13	-	-
TOTAL	1,992 (100%)	20.69	6,130 (100%)	13.20

Among the left collocates of *DISTURB*, we find adverbs which are used as pre-modifiers of the past participle (*disturbed*); the most frequent ones in the *BNC* and the *CWo* are displayed in Table 5.20. As we may notice, these adverbs mainly express the notion of intensity (e.g. *barely*, *very*), only one that of frequency (e.g. *constantly*), while a few identify the realm of experience where the concept of ‘disturbing’ applies (e.g. *emotionally*, *mentally*, *psychologically*; e.g. “[t]hat, the court heard, gave him the chance to study the behaviour of people who really were mentally *disturbed*”, *BNC*: CEN; “Assuming that the verbally abusive patterns are not the product of a severely *disturbed* psyche, [...]”, *CWo*: usbooks).

<sup>24</sup> *Disturbed* represents the simple past, the past-participle form or a deverbal adjective.

<sup>25</sup> *Disturbing* appears to be used both as part of a progressive form and as an adjective.

Table 5.20 Left collocates of *DISTURB*:  
adverbial pre-modifiers in the two corpora and in the *BNC* and *CWo* datasets

Pre-modifiers of <i>DISTURB</i>	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>barely</i>	6.03	3.20	1 (0.5%)	-
<i>constantly</i>	5.20	2.62	1 (0.5%)	-
<i>deeply</i>	7.86	6.72	2 (1%)	3 (1.5%)
<i>emotionally</i>	8.10	7.52	-	2 (1%)
<i>little</i>	4.91	3.18	-	3 (1.5%)
<i>mentally</i>	7.60	7.46	2 (1%)	5 (2.5%)
<i>obviously</i>	4.55	2.70	-	1 (0.5%)
<i>particularly</i>	3.90	3.56	-	1 (0.5%)
<i>profoundly</i>	7.98	7.34	1 (0.5%)	1 (0.5%)
<i>psychologically</i>	6.95	6.92	1 (0.5%)	-
<i>seriously</i>	6.74	4.73	3 (1.5%)	1 (0.5%)
<i>severely</i>	7.53	6.60	-	3 (1.5%)
<i>very</i>	-	-	-	9 (4.5%)

Fig. 5.20.1 Screenshot of the left collocates of *DISTURB* in the *BNC* dataset:  
adverbs

```

28 JXX pervaded by a dreamy languor, a torpid drowsiness, which was barely < disturbed > as his hand closed over her wrist. And then, so gently
60 BMW standard of honesty. Ambiguities of this kind constantly < disturb > our attempts to describe the social organization of Europe
61 HLL claimed that" most courts in the civilized world will be deeply < disturbed >" by it. In a sharply worded communiqué released on Aug.
130 CLM private Reason'. These conciliatory figures were profoundly < disturbed > by the bitter armed conflict of the 1640s, and the Royalist
138 FVW state of distress and have less time for the more seriously < disturbed > and those with long-term disorders who are less readily

```

Fig. 5.20.2 Screenshot of the left collocates of *DISTURB* in the *CWo* dataset:  
adverbs

```

71 brbooks with personal crisis, to poltergeist cases, where emotionally < disturbed > teenagers seem to be an important component. In voodoo
98 brbooks rumours, I wouldn't be troubled; but as it is, I am a little < disturbed >. However all any of us can do is to pray to the Lord
100 usbooks son and daughter were shot to death, apparently by a mentally < disturbed > younger son who then drowned himself in the San Juan.
127 usbooks books: a book by a 'clever, gentlemanly fellow' particularly < disturbed > Oldmixon, for the gentleman 'took a run into the interior
135 usbooks verbally abusive patterns' are not the product of a severely < disturbed > psyche, changing these behaviors should prove no more

```

The past participle form *disturbed* and the word-form *disturbing* are preceded by the auxiliary verbs *to be* and *to have* in their various word-forms, none of which being particularly prominent in the corpora: *is/are, was, has/have, had, have been, had been, will be, being* (*BNC*: 57 occurrences, 28.5%; *CWo*: 78 occurrences, 39%); modal auxiliaries are also instantiated, but less frequently: *might be, must have been, may have been, could be, should be* (*BNC*: 12 occurrences, 6%; *CWo*: 9 occurrences, 4.5%).

As Table 5.21 displays, in my datasets the subjects that occur with the predicate *DISTURB*, both in the active and passive voice, are mainly personal pronouns (*I, you, he/she/it, we, they*: *BNC*: 72 occurrences, 36%; *CWo*: 65 occurrences, 32.5%), proper names and common nouns identifying people or animals (e.g. *John, Elizabeth, Roland, parents, neighbours, camels, the spider, birds, people, police, the family*). Other subject noun phrases, used in active sentences, denote sources of disturbance such as inanimate entities like news, noisy events (e.g. *noise, sound*) or concrete entities (e.g. *hat, injuries, wind, cysts, earthquake, walls, protective layer*). In passive sentences, we find subject noun phrases that indicate the condition or entity affected by the source of disturbance (e.g. *thought, silence, harmony, balance, equilibrium, ideas, question, sleep*).

Table 5.21 Left collocates of *DISTURB*:  
subjects of *DISTURB* in the *BNC* and *CWo* datasets

Frequency					
Subjects of <i>DISTURB</i>		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)		
<b>Personal pronouns</b>		<b>72 (36%)</b>	<b>65 (32.5%)</b>		
<i>I</i>		16 (8%)	15 (7.5%)		
<i>you</i>		6 (3%)	6 (3%)		
<i>he</i>		10 (5%)	10 (5%)		
<i>she</i>		8 (4%)	9 (4.5%)		
<i>it</i>		15 (7.5%)	10 (5%)		
<i>we</i>		2 (1%)	6 (3%)		
<i>they</i>		15 (7.5%)	9 (4.5%)		
<b>Nouns</b>	<b><i>BNC</i> (pmw)</b>	<b><i>CWo</i> (pmw)</b>			
<i>news</i>	1.59	0.89	1 (0.5%)	2 (1%)	
<i>noise, sound</i>	5.06	7.66	3 (1.5%)	2 (1%)	
<i>sleep</i>	-	-	3 (1.5%)	2 (1%)	

The direct objects of the verb *DISTURB* are personal pronouns: *me, you, him, her, us, them* (*BNC*: 47 occurrences, 23.5%; *CWo*: 30 occurrences, 15%) and other terms such as those listed in Table 5.22, which mostly refer to situations and concepts characterized by, or denoting, calmness and tranquillity and, occasionally, the quiet behaviour of a criminal such as a burglar (*BNC*: 69 occurrences, 34.5%; *CWo*: 58

occurrences, 29%; e.g. “Er if it’s at night time it annoys you, cos if all ring your neighbours they’re not gonna be so friendly if your alarm’s ringing forever and a day and you know, it’s *disturbing* your sleep”, *BNC*: KNF; “A knock at the door *disturbed* him, and he looked up in annoyance”, *CWo*: usbooks). Finally, the direct object of *DISTURB* may identify animals, whose peaceful or busy state may be disturbed by some external agent (e.g. “Do not *disturb* livestock or damage crops or trees”, *BNC*: ECG; “The ants, which get their name from their chaotic behaviour when *disturbed*, squirt formic acid over any animals that venture in their path, [...]”, *CWo*: cannews).

Table 5.22 Right collocates of *DISTURB*:  
direct objects of *DISTURB*  
in the two corpora and in the *BNC* and *CWo* datasets

Objects of <i>DISTURB</i>	Frequency		<i>BNC dataset</i> (tokens and percentage)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>balance</i>	5.06	4.96	3 (1.5%)	2 (1%)
<i>equilibrium</i>	6.92	6.91	3 (1.5%)	1 (0.5%)
<i>intruder/burglar</i>	6.65	11.72	2 (1%)	-
<i>neighbour/neighbourhood</i>	4.80	4.70	1 (0.5%)	-
<i>peace</i>	6.42	6.14	3 (1.5%)	5 (2.5%)
<i>sleep</i>	7.11	6.13	4 (2%)	2 (1%)
<i>tranquillity/calm</i>	7.22	10.12	1 (0.5%)	-
<b>Animals</b>				
<i>ants, bees, birds, cats, curlew, gulls, livestock, rabbits, sharks, sheep, spiders</i>	-	-	8 (4%)	4 (2%)
<i>wildlife</i>	5.33	4.11	1 (0.5%)	-

Fig. 5.21 Screenshot of the right collocates of *DISTURB* in the *BNC* dataset:  
direct objects

```

32  K1H      possibility they haven't ruled out is that Mr Miles could have < disturbed > a burglar. But they say as yet they've found no signs of
145 AKE      to gallop straight through a fence without falling, or even < disturbing > the equilibrium of his limpet-like companion. But for
157 JY9      and a fortnightly van.' 'Your commercial activities will < disturb > the neighbourhood,' he repeated. Ashley's eyes returned
102 KBR      everyone, he had found affinity. He hadn't wanted them to come and < disturb > his peace, that was obvious. There was no ease
170 CAM      at Upper Heyford right on schedule at 1.55 a.m. local time, < disturbing > the sleep of the villagers beneath him as he made his

```

In my datasets, the term *DISTURB* co-occurs with words indicating psychological turmoil and in expressions describing someone who feels upset and/or worried (*BNC*:

71 occurrences; 35.5%; *CWo*: 76 occurrences, 38%; e.g. “The boy was not then aware, of course, of how frequently an idling or unanchored mind can be *disturbed* by the accidental experience of receiving luminous images from out of clear, sharp light when it is itself cast in any sort of shadow”, *BNC*: ADA; “This is what’s *disturbing* about these pictures of performance grief”, *CWo*: oznews). The term also appears to be used when referring to the change of positions or shapes that something undergoes as a result of movement (*BNC*: 15 occurrences, 7.5%; *CWo*: 26 occurrences, 13%; e.g. “Her outstretched hand *disturbed* a fragment of loose rock, sending it tumbling over the precipice”, *BNC*: GVP; “This earthquake *disturbed* a 620-mile section along the boundary of huge geological plates”, *CWo*: usnews).

In my datasets, the verb *DISTURB* frequently occurs within the following colligational pattern: “*DISTURB* + preposition + noun”. In particular, the word-form *disturbed* is often instantiated in “*disturbed* + *by* + noun” (*BNC*: 24 occurrence, 12%; *CWo*: 22 occurrences, 11%) as part of passive clauses. Table 5.23 shows the nouns which most frequently occur in my datasets in such a pattern. Other terms which only occur once in my datasets are shown in Figg. 5.22.1 and 5.22.2. The referents of the agentive complement appear to be both animate and concrete (e.g. *birds*) and inanimate and abstract entities (e.g. *statements*).

Table 5.23 Right collocates of *disturbed* in passive clauses: *disturbed* + *by* + noun in the two corpora and in the *BNC* and *CWo* datasets

Nouns following the passive construction <i>disturbed by</i>	Frequency		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>noise/sound</i>	4.92	6.78	4 (2%)	1 (0.5%)
<i>report</i>	1.20	0.36	3 (1.5%)	5 (2.5%)

Fig. 5.22.1 Screenshot of ‘*disturbed* + *by*’ in the *BNC* dataset

53 GUA flocks make remarkable roaring sound on surface of water when < disturbed > by birds of prey. Much given to quarrelsome chases,  
 64 AR8 were beginning to run short of food and water, but had not been < disturbed > by the enemy. Their only foe had been the usual plague  
 65 HTM eat, to indicate their good intentions. The Green Jacks were < disturbed > by the fire, but Wynne-Jones placed pieces of flint on  
 66 EVA Freudian theory should therefore predict that sleep might be < disturbed > by the prevention of dreaming, but that any untoward effects  
 67 CAS playing softly in the background. A few minutes later, I was < disturbed > by the sound of 'boogie-woogie' dance music - coming from my

Fig. 5.22.2 Screenshot of ‘*disturbed* + *by*’ in the *CWo* dataset

59 sunnow lonely hearts ads after splitting with Debbie-and many were < disturbed > by his sexual appetities. Kinky One woman, referred to  
 60 usbooks of it. we circle solemnly above the Titanic, attracted and < disturbed > by its mystery, not as a technical problem to be solved  
 68 brbooks be waiting for him. On the fourth morning he had woken early, < disturbed > by the insistent rhythms of drills and hammers,  
 74 indnews democratic norms, justice and human rights in India, we are very < disturbed > by the statements emanating from the Delhi Police,  
 75 brbooks tapping on the window, but it was no more than a hawthorn bush, < disturbed > by the wind. 'disk error.' Damn. But delight. There



Another preposition that is used in the above mentioned colligational pattern is *about*, which however is infrequent (it occurs only 3 times in the *BNC* and in the *CWO* datasets in “*disturbed/disturbing + about + topic/event*”). As Fig. 5.23.1 and 5.23.2 show, another, albeit infrequent, pattern (*BNC/CWO*: 3 occurrences) is the form *disturbed* followed by the *to*-infinitive denoting the event which creates disturbance.

Fig. 5.23.1 Screenshot of “*disturbed/disturbing + about*” and “*disturbed + to+ verb*” in the *BNC* dataset

36	A4L	couple (Randy Quaid and Mary Beth Hurt) has a great deal to be < disturbed > about. To give away much more of first-time director
37	GTD	between the Mall and the backs of houses in Pall Mall, a garden < disturbed > about 1709 for the building of Marlborough House and
184	H9C	monk. Something stirred in his memory but he was too tired and < disturbed > to concentrate. Ranulf came up and expertly went
185	H9H	unreadable thoughts. 'Oh, him!' Belinda blurted dismissively, < disturbed > to find that Tom had been aware of at least some of
186	C98	inexorable forces of Nature. Both Annie and Elizabeth were < disturbed > to find that each morning brought fresh fox tracks

Fig. 5.23.2 Screenshot of “*disturbed/disturbing + about*” and “*disturbed + to+ verb*” in the *CWO* dataset

39	safrica	Nelwamondo, Chief Calvin Nelwamondo told The Star that he was < disturbed > about the news as it involved someone he had regarded
40	oznews	of a loved one? And what would you be feeling? This is what's < disturbing > about these pictures of performance grief.
41	brspok	evening everyone. <M01/> Yes Pauline. <F04/> Yes er I'm somewhat < disturbed > about what the proposals are about the destruction
198	brbooks	fundamentalist Christian while in his waking state, Cayce was profoundly < disturbed > to discover that while in trance he delivered life
199	usspok	for the increased number of cases. Still, he says Muslims are < disturbed > to see the trend increasing, not decreasing, over

My datasets also reveal the pattern “*sorry + DISTURB*” (see Fig. 5.24), in which ‘*disturbing*’ is the damage that people apologise for (*BNC*: 8 occurrences, 4%; *CWO*: 2 occurrences, 1%).

Fig. 5.24 Screenshot of ‘*sorry to disturb*’ in the *BNC* dataset

91	H8L	to have left my lighter in the dining-room. So sorry to keep < disturbing > you like this, I must go and get it.' She closed the
96	CA3	Ridley, do you remember? Laurence's daughter. I'm so sorry if I'm < disturbing > you. I've only popped in for a few minutes. It's Lee.'
166	G12	The doorbell rang. It was Mrs. Mounce. 'Sorry to < disturb >, darling,' she said to Bob, 'but I've brought you a bottle
167	CKC	and Montgomery at the breakfast table. 'Coffee? Sorry to < disturb > you but these are times that are sent to try men's souls.'
168	G0P	politely at her. 'Good morning,' he said. 'I'm sorry to < disturb > you. I was wondering if I could speak to Mrs Ward. Mrs Donna

To sum up, the data retrieved show that the term *DISTURB* in its various word-forms frequently occurs in fiction both in the *BNC* and the *COCA*, and that the most frequent word-form is *disturbed*. The verb appears to be preceded by adverbs that emphasise the way in which the action is carried out and to be followed by personal pronouns as direct objects. *DISTURB* also occurs followed by the preposition *by* + agent and rarely *about* and *to* + verb. The term *DISTURB* is instantiated in situations in which someone is upset, worried or interrupted by someone/something else and when there is the interruption of an activity or the disruption of the calmness of a situation or place. *DISTURB* also occurs when describing people that are mentally deranged and sometimes when the meaning conveyed is that of frightening animals.

### 5.3.2 To bother

The verb *BOTHER* occurs 3,956 times in the *BNC* and 18,148 times in the *COCA*, which corresponds to about 40 times pmw in both corpora and, as Table 5.24 displays, the most frequent word-form in both corpora is the base form *bother* (about 50% of all occurrences).

The frequency hierarchy of the various forms of *BOTHER* in the *BNC* dataset is as follows: *bother* (45%) > *bothered* (33.5%) > *bothering* (10%) > *to bother* (8%) > *bothers* (3.5%); the frequency hierarchy in the *CWo* dataset is as follows: *bother* (43.5%) > *bothered* (33.5%) > *bothers* (8.5%) > *bothering* (8%) > *to bother* (6.5%).

As we can see in Table 5.24, the verb *BOTHER* occurs the most frequently in fiction in both corpora: 1,402 times in the British English corpus and 7,765 times in the American English corpus, which corresponds to 88.12 times pmw and 85.87 times pmw, respectively. In addition, the verb occurs 125.56 times pmw in spoken British English and only 38.17 times pmw in spoken American English.

The verb *BOTHER* appears to occur mostly in declarative clauses, both affirmative (*BNC*: 48%; *CWo*: 32%) and negative clauses (*BNC*: 42%; *CWo*: 57%), and less frequently in interrogative clauses, both positive and negative (*BNC*: 10%; *CWo*: 7%).

Table 5.24 Frequency of occurrence of the term *BOTHER* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Word-forms</b>				
<i>Bother</i> <sup>26</sup>	<b>2,086 (52.73%)</b>	<b>21.67</b>	<b>8,983 (49.50%)</b>	<b>19.35</b>
<i>Bothers</i>	116 (2.93%)	1.21	1,970 (10.85%)	4.24
<i>Bothered</i> <sup>27</sup>	1,398 (35.34%)	14.52	5,262 (28.99%)	11.34
<i>Bothering</i> <sup>28</sup>	356 (9.00%)	3.70	1,933 (10.66%)	4.17
<b>Genres</b>				
Spoken	<b>1,251 (31.62%)</b>	<b>125.56</b>	<b>3,648 (20.10%)</b>	<b>38.17</b>
Fiction	<b>1,402 (35.44%)</b>	<b>88.12</b>	<b>7,765 (42.79%)</b>	<b>85.87</b>
Magazine	258 (6.53%)	35.53	3,072 (16.93%)	32.15
Newspaper	235 (5.94%)	22.45	2,809 (15.48%)	30.63
Non-academic	253 (6.39%)	22.45	-	-
Academic	113 (2.86%)	7.37	854 (4.70%)	9.38
Other	444 (11.22%)	21.31	-	-
<b>TOTAL</b>	<b>3,956 (100%)</b>	<b>41.10</b>	<b>18,148 (100%)</b>	<b>39.08</b>

The immediate left collocates of *BOTHER* are adverbs that specify circumstances (e.g. of manner, frequency or temporal collocation) of the action denoted by the verb (see Table 5.25). In particular, in my dataset, the most frequent adverbs pre-modifying the verb are time adverbs such as *never*, *ever* and *rarely*, adverbs of degree or intensity such as *hardly*, *too*, *much*, and adverbs of precision such as *really* (see Fig. 5.25.1 and 5.25.2).

<sup>26</sup> The term *bother* refers only to the verb; the noun was excluded from the research.

<sup>27</sup> *Bothered* represents the simple past, the past-participle form and the deverbal adjective.

<sup>28</sup> *Bothering* appears to be used both as part of a progressive form and an adjective.

Table 5.25 Left collocates of *BOTHER*:  
adverbial pre-modifiers of *BOTHER*  
in the two corpora and in the *BNC* and *CWo* datasets

Frequency				
Pre-modifiers of <i>BOTHER</i>	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>anyway</i>	7.68	2.92	3 (1.5%)	1 (0.5%)
<i>even</i>	7.57	5.25	8 (4%)	11 (5.5%)
<i>ever</i>	6.01	3.52	2 (1%)	1 (0.5%)
<i>hardly</i>	7.10	3.76	1 (0.5%)	-
<i>longer</i>	5.34	4.00	-	2 (1%)
<i>much</i>	5.56	2.92	2 (1%)	4 (2%)
<i>never</i>	6.65	4.46	6 (3%)	5 (2.5%)
<i>no</i>	5.64	3.80	3 (1.5%)	-
<i>particularly</i>	5.45	3.58	2 (1%)	-
<i>rarely</i>	6.69	4.56	1 (0.5%)	1 (0.5%)
<i>really</i>	6.27	4.62	1 (0.5%)	5 (2.5%)
<i>that</i>	4.33	5.24	2 (1%)	4 (2%)
<i>too</i>	4.08	2.93	1 (0.5%)	1 (0.5%)

Fig. 5.25.1 Screenshot of the left collocates of *BOTHER* in the *BNC* dataset:  
adverbs

```

30 H46      to put on its annual Carol Service last month and didn't even < bother > to tell anyone), at which the whole book business could get
43 A7A      about next week. You are crowded at home?' Erika hardly < bothered > to say 'yes'. For a family in the position of hers
156 CK6     also knows it might 'peak at 114'. He doesn't seem particularly < bothered > either way. He talks about being shipped around London
157 CEX     where she had been a quick, intense, thundery girl who rarely < bothered > to talk to other women. Now she was successful,
158 ED3     and buy a Jeff Banks, I want them to know that we have really < bothered >.' DAVID EMANUEL David Emanuel (above) is best known for

```

Fig. 5.25.2 Screenshot of the left collocates of *BOTHER* in the *CWo* dataset:  
adverbs

```

70 usbooks  toward the sounds instead of resenting them and they'll no longer < bother > you. Perhaps you live in a neighborhood where there is
73 brbooks  Not that Colonel Nasser or any of the Arab leaders were much < bothered > about Palestinian Arabs. But Nasser's credentials
132 usbooks  movie, and when he came to pick them up again at eleven he never < bothered > to ask what they had done. They must have been familiar
160 brregnews contents of my fridge." The television thing doesn't really < bother > me any more. I've been on GMTV a few times already and
191 brregnews doesn't and I'm not too worried about it." He may not be too < bothered > about what went on but Celtic and Balde will be anxiously

```

The past participle form *bothered* and the word-form *bothering* are preceded by the auxiliary verbs *to be* and *to have*, in their various word-forms (*BNC*: 48 occurrences, 24%; *CWo*: 41 occurrences, 20.5%) and occasionally by the modal verbs *can/could*, *may* and *should* (*BNC*: 12 occurrences, 6%; *CWo*: 14 occurrences, 7%).

In my datasets the subject noun phrases that occur with the predicate *BOTHER*, both in active and passive sentences, denote human and animal agents and are personal pronouns (*I, you, he/she/it, we, they*: *BNC*: 111 occurrences, 50.5%; *CWo*: 100 occurrences, 50%); proper names (e.g. *Hitler, Pamela, Jack, Erika, Hoffman, Thomas*) and common nouns (e.g. *the kids, people, the writer, lesbians, your sister, my husband, the author, fans, the girl, doctors, the novelist, squirrels*). When the verb is used in its active voice, these subject noun phrases are used to indicate entities that are responsible for the act of bothering; when the verb is used in its passive voice, these subject are used as the patients affected by the act of bothering. As Table 5.26 displays, other NPs, which are used in active instances, are words such as *noise* and *conscience*, which denote two inanimate causes of disturbance. In addition to these subjects, other NPs occurring with the verb in the active voice and instantiated only once in my datasets are inanimate and abstract nouns (e.g. *resemblance, incarceration, feedback, aspects of the reforms*), nouns denoting part of the body (*the leg, the knee*) and pronouns (*nothing, something, anything, nobody, someone, none*).

Table 5.26 Left collocates of *BOTHER*:  
in the whole corpora and in the *BNC* and *CWo* datasets: subjects

<b>Frequency</b>				
<b>Subjects of <i>BOTHER</i></b>			<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<b>Personal pronouns</b>			<b>111 (50.5%)</b>	<b>100 (50%)</b>
<i>I</i>			48 (24%)	24 (12%)
<i>you</i>			16 (8%)	8 (4%)
<i>he</i>			13 (6.5%)	22 (11%)
<i>she</i>			6 (3%)	7 (3.5%)
<i>it</i>			10 (5%)	24 (12%)
<i>we</i>			5 (2.5%)	2 (1%)
<i>they</i>			13 (6.5%)	13 (6.5%)
<b>Nouns</b>	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>conscience</i>	5.90	5.63	1 (0.5%)	-
<i>noise</i>	2.91	4.19	8 (4%)	11 (5.5%)

In my datasets, the verb *BOTHER* virtually always takes as its direct objects personal pronouns (see Figg. 5.26.1 and 5.26.2): *me, you, him, her, us, them* (*BNC*: 42 occurrences, 21%; *CWo*: 49 occurrences, 24.5%), while only rarely does it have other direct objects, which anyway always identify sentient beings or experiencers and never inanimate entities.

Fig. 5.26.1 Screenshot of the direct objects of *BOTHER* in the *BNC* dataset:  
personal pronouns

73	A74	's been a bit funny these last couple of days, like summat's < bothering > her or summat. I dunno what's up, but I reckon summat
74	CH3	that match, which showed he has a terrific attitude. Nothing < bothers > him.' Mick Quinn collected three crates of champagne
84	A74	am and shut my eyes. The noises swirl round me, but it don't < bother > me. I'm falling. Falling head over heels in black. I
112	G4S	There are fashions are there as well for the women not Yeah < bother > them quite so much. Erm recently this er Asian dress has
210	FBM	Am I right in assuming that you have no use for it? I might < bother > you with another one (or two!) that I have completed since

Fig. 5.26.2 Screenshot of the direct objects of *BOTHER* in the *CWo* dataset:  
personal pronouns

73	brbooks	small point, Mrs Dearden," Thorne said, after apologizing for < bothering > her at home. 'Could you tell me exactly what time of day
74	usnews	has grown in the pros, and his teammates know that any loss < bothers > him, including that 24-14 loss at Gillette Stadium on Jan.
99	sunnow	to me now and say 'love it, love it'. Truth "It doesn't < bother > me. I felt at the time what I said was right. I felt I was
102	brspok	my mother, having dropped off the van on the way. one thing < bothered > me: he always seemed to be in a rush to get on, and I
204	usbooks	toward the sounds instead of resenting them and they'll no longer < bother > you. Perhaps you live in a neighborhood where there is

The concordances in my datasets show that *BOTHER* instantiates the meaning of ‘to worry about something, to feel upset’, 102 times (51%) in the *BNC* and 71 times (35.5%) in the *CWo* (e.g. “Something’s *bothering* you, I know, she said quietly. Won’t you tell me, Beth?”, *BNC*: FPK; “this inconvenient fact evidently *bothered* Franklin, too, in later years, and so, as he so often did, he simply altered the past to suit the present”, *CWo*: usbooks). In addition, the verb *BOTHER* also appears to occur in the meaning of ‘giving annoyance/irritation’ (*BNC*: 24 occurrences, 12%; *CWo*: 16 occurrences, 8%; e.g. “What did *bother* him was that in his duel within McLaren with Prost, the Frenchman, six years his junior, always started with an advantage”, *BNC*: CD9; “I mean, it *bothers* me a LITTLE, but really, I’m fine with it”, *CWo*: usbooks). The term is also used to signal when someone or something interrupts the activity of a person (*BNC*: 16 occurrences, 8%; *CWo*: 22 occurrences, 11%; e.g. “the sound comes tumbling around the edge of Jacob’s dream like little wooden ark animals, flat-sided and barely familiar, but it doesn’t *bother* his sleep”, *CWo*: usbooks; “I’m sorry to *bother* you at work, Mrs Day”, *BNC*: GUM). The term *BOTHER* is used to refer to a person who suffers because of physical pain in 1 occurrence in the *BNC* and 7 in the *CWo* (e.g. “[...] Stephen Davis (992 yards), who is *bothered* by a sore ankle, [...]”, *CWo*: usnews). Finally, the term occurs as conveying the meaning of frightening someone by following them around in only 1% of the instances in the *CWo* (e.g. “[...] she had complained of

being *bothered* by “an Asian man””, *CWo*: safrica); and in only one of the occurrences retrieved in the *CWo* does *bother* convey the meaning of ‘being interested in something’ (e.g. “Oasis are still the best band in my mind, but I’m not *bothered* about that kind of stuff any more”, *CWo*: sunnow).

In my datasets, the verb *BOTHER* occurs in the colligational patterns: “verb + preposition + noun” and “verb + *to*-infinitive + verb” (see Fig. 5.27.1 and 5.27.2): “*to bother*+ *about* + topic/event” (*BNC*: 20 occurrences, 10%; *CWo*: 8 occurrences, 4%), “*bothered* + *by* + noun indicating an agent” (*BNC*: 4 occurrences, 2%; *CWo*: 12 occurrences, 6%), “*to bother*+ *with*+ noun indicating an issue to deal with” (*BNC*: 26 occurrence, 13%; *CWo*: 13 occurrences, 6.5%), “*to bother*+ *to*-infinitive” (*BNC*: 46 occurrences, 23%; *CWo*: 49 occurrences, 24.5%); this last pattern is frequently used to convey the meaning of ‘to take the trouble to do something’ and ‘to make the effort to do something’, typically in a negative context (e.g. “She didn’t *bother* to answer”, *BNC*: HH1; “This time she didn’t *bother* to reply”, *CWo*: usbooks; “I didn’t even *bother* to argue with him”, *CWo*: brbooks).

Fig. 5.27.1 Screenshot of “*BOTHER* + preposition + noun” and “*BOTHER* + *to*+ verb” in the *BNC* dataset

39	K88	there, then. Sunderland were there. Oh, I wouldn't have been < bothered > about any other cos you see it better on the, on the screen
60	G2T	the strap tended to lie on top of James' head - but he wasn't < bothered > by them.' Tomy Comfort Carrier, £33 Age: birth to 24 months
117	HH1	horse's back?' Isabel sighed and subsided. She didn't < bother > to answer. FitzAlan would only snap at her anyway, as he had
179	BP8	. At least there's no awkward questions asked, no needing to < bother > with insurance, does mean there's a bit to add to her measly
180	AC7	doubtfully, but I rejected the idea at once. why should God < bother > with someone who chose deliberately to do something foolish

Fig. 5.27.2 Screenshot of “*BOTHER* + preposition + noun” and “*BOTHER* + *to*+ verb” in the *CWo* dataset

31	sunnow	them with confidence when they play us. 'To be honest, I'm not < bothered > about how well they're doing. It's how we perform that
50	usnews	played without leading rusher Stephen Davis (992 yards), who is < bothered > by a sore ankle, and DeShaun Foster managed just 56 yards
136	indnews	deputed to the task. Like many first-time users, the family didn't < bother > to change the password-a simple-enough task-and India's first
166	brnews	shopkeeper ', Joyce called the old firm in Finnegans wake), seldom < bother > to read him. Auden tried hard to pique our interest, and the
198	sunnow	Toilet talk? MOST of the sailors were so drunk they didn't < bother > with toilets. There were trails of sick behind each of the

An additional pattern instantiated in my datasets and displayed in Fig. 5.28.1 and 5.28.2 is the semantic preference of *BOTHER* with the notion of ‘apologising’ as in ‘sorry *to bother*’(*BNC*: 5 occurrences, 2.5%; *CWo*: 2 occurrences, 1%); ‘apologizing for’ (*CWo*: 1 occurrence, 0.5%).

Fig. 5.28.1 Screenshot of “*sorry to bother*” in the *BNC* dataset

183	ASS	, cautiously. 'I know,' said Henry, 'and I am sorry to < bother > you at this time of night but my client has given me to understand
184	AMB	'Excuse me,' said the scientist politely. 'I'm sorry to < bother > you, but what were those directions again? I seem to have
185	GUM	, something. But I mean Ruth's quite all right. I'm sorry to < bother > you at work, Mrs Day. But I wanted a word with you when the
186	KB7	for showing us your home. Thank you very much. I'm sorry to < bother > you. If you can hear next door's music I can assure you it's
187	FAB	closed his notebook. 'That will be all. I am very sorry to < bother > you, but I'm sure you understand.' 'of course.' George

Fig. 5.28.2 Screenshot of “*sorry to bother*” in the *CWo* dataset

```
1 brbooks Taboured himself across the room. Peter felt sorry that he'd < notbothered > the poor old coot. Once the man managed to reach Thurzella,  
73 brbooks small point, Mrs Dearden," Thorne said, after apologizing for < bothering > her at home. 'Could you tell me exactly what time of day it  
188 brbooks morning. Finally, I rang Brian Chalmers of PharmAce. 'Sorry to < bother > you at home, Brian, but have any of your vans been in an accident
```

To sum up, corpus data shows that *BOTHER* occurs most frequently in fiction but quite frequently also in the spoken language both in the *BNC* and the *COCA*. The term is preceded by adverbs, which specify the manner, intensity and other circumstances of the action of disturbance; and it is followed by direct objects such as personal pronouns which identify human experiencers. It is also followed by prepositions such as *by+* noun denoting animated or inanimate entities, *about* + topic/event, *with* + noun indicating an issue to deal with, and *to* + verb indicating the cause of annoyance. The concordances also show that *BOTHER* is mostly used to indicate a feeling of worry or frustration, the interruption of an activity or an action carried out because it had to be done.

### 5.3.3 Conclusion

In conclusion, the findings related to the verbs *DISTURB* and *BOTHER* show that the verbs have different frequency scores but occur in the same genres. In particular, both verbs frequently occur in fiction, but *DISTURB* is less frequent than *BOTHER*. Both verbs are also frequently preceded by adverbs, which specify the way in which, or the extent to which, the action denoted by the verb is carried out; when accompanied by modal verbs, the meanings of the two verbs are coloured by the notions of ability, permission or obligation.

Both verbs show a high frequency of occurrence in instances in which they convey the meanings of (1) ‘making someone feel worried and/or upset’ or (2) ‘interrupting an activity’. Moreover, while *DISTURB* is used to indicate insanity or mental derangement, *BOTHER* is only used when physical – not mental – pain is involved. Finally, while *DISTURB* is more likely to be used when reference is made to the change of position/shape, the frightening of animals, or to the interruption of the quietness of a place or situation, *BOTHER* appears to be chosen when people want to underline the fact that someone/something is annoying them.

As in the case of the nouns, for the verbs too, the findings from the entire corpora and those from my datasets do not perfectly match, that is the *word sketches* retrieved from the whole corpora, which are tables showing the terms that frequently occur instantiated as immediate left or right collocates of the search term, included collocates



which were not listed at all in my datasets or showed frequency scores which did not correspond to those I identified in my datasets.

Table 5.27 Patterns in which *DISTURB* and *BOTHER* are likely to occur

		<i>To disturb</i>	<i>To bother</i>
<b>Pre-modifiers</b>	Adverbs	x	x
<b>Subjects</b>	Personal pronouns	x	x
	Animate and inanimate entities	x	x
<b>Objects</b>	Personal pronouns	x	x
	NPs denoting animals	x	-
	Animate entities	x	x
	Inanimate entities	x	-
<b>Prepositions</b>	<i>by</i> + noun denoting people	x	x
	<i>about</i> (+ noun denoting event/topic)	x	x
	<i>with</i> + noun denoting an issue to deal with	-	x
	<i>to</i> + verb denoting the action that causes disturbance	x	x
<b>Meanings</b>	To interrupt an activity	x	x
	To worry/ upset	x	x
	To interrupt quietness	x	-
	To change position	x	-
	To frighten animals	x	-
	To cause psychological turmoil	x	-
	To have physical pain	-	x
<b>Other</b>	<i>sorry to</i> + verb indicating the cause of disturbance	x	x

Therefore, as summarized in Table 5.27, corpus findings regarding *DISTURB* and *BOTHER* suggest that these verbs do not appear to be completely interchangeable in all situations; in particular, they are more likely to be replaced by one another only when they convey the meanings of interrupting an activity and making someone feel worried or upset. On the other hand, only *DISTURB* appears to be used when conveying the meanings of ‘interrupting the quietness’, ‘changing positions or shapes’, ‘frightening animals’ and ‘causing psychological turmoil’, whereas only *BOTHER* is used to convey the meaning of ‘having physical pain’.

## 5.4 The adjectives

In the following sub-sections I will report on the findings relevant to the adjectives *compulsory*, *obligatory* and *mandatory*. In particular, for every adjective I will present the raw and the normalised frequency of occurrence and its distribution across genres within the *BNC* and the *COCA*. Then, I will consider both my datasets retrieved from the *BNC* and *CWo* and the whole corpora to identify the left and right collocates of the term in question, the meanings it is used to convey and its colligational and semantic patterns.

### 5.4.1 *Compulsory*

The adjective *compulsory* occurs more frequently in British English than in American English and in particular, as Table 5.28 displays, it occurs 1,679 times in the former and 1,142 times in the latter (17.44 times pmw in British English vs 2.46 times pmw in American English). Table 5.28 also shows that the term *compulsory* has a high frequency of occurrence in the academic genre both in British English and in American English: as matter of fact, the term occurs 30.59 times pmw in the *BNC* and 7.79 times pmw in the *COCA*.

Table 5.28 Frequency of occurrence of the term *compulsory* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Genres</b>				
Spoken	80 (4.77%)	8.03	72 (6.31%)	0.75
Fiction	30 (1.79%)	1.89	56 (4.90%)	0.62
Magazine	57 (3.39%)	7.85	165 (14.45%)	1.73
Newspaper	185 (11.02%)	17.68	140 (12.26%)	1.53
Non-academic	373 (22.22%)	22.61	-	-
Academic	<b>469 (27.93%)</b>	<b>30.59</b>	<b>709 (62.08%)</b>	<b>7.79</b>
Other	485 (28.88%)	23.28	-	-
TOTAL	1,679 (100%)	17.44	1,142 (100%)	2.46

The immediate left collocates of the adjective *compulsory* can be determiners: *a*, *the* (*BNC*: 44 occurrences, 22%; *CWo*: 23 occurrences, 11.5%), which precede the

adjective when this is used as the pre-modifier of a noun (see the paragraph on attributive adjectives below); *compulsory* can also be immediately preceded by adverbs such as those shown in Table 5.29, or verbs such as *to become*, *to introduce*, *to make* (*BNC*: 11 occurrences, 5.5%; *CWo*: 13 occurrences, 6.5%); in the latter case, as Figg. 5.29.1 and 5.29.2 show, *compulsory* is used as a the subject complement or the object complement and thus as a predicative adjective (see the paragraph below).

Table 5.29 Left collocates of *compulsory*:  
adverbial per-modifiers in the two corpora and in the *BNC* and *CWo* datasets

Pre-modifiers of <i>compulsory</i>	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>almost</i>	2.32	1.61	1 (0.5%)	-
<i>legally</i>	6.07	2.26	1 (0.5%)	-
<i>no</i>	-	-	2 (1%)	4 (2%)
<i>no longer</i>	3.87	0.94	2 (1%)	-
<i>now</i>	2.90	-	4 (2%)	-
<i>practically</i>	5.38	4.17	1 (0.5%)	1 (0.5%)

Fig. 5.29.1 Screenshot of *compulsory* in the *BNC* dataset:  
*to become*, *to introduce* and *to make*

62 B76 Clunk-click-mumble Keith Dale on the myths of belting-up SEAT BELTS become < compulsory > dress for the well-turned out driver at the end  
95 BMB February. Hindustan Lever, Unilever's Indian subsidiary, introduced < compulsory > rural work for graduate trainees as part of  
110 EDK practice of vaccination spread fast at home and abroad. It was made < compulsory > in Denmark, Norway, and Sweden around 1810.

Fig. 5.29.2 Screenshot of *compulsory* in the *CWo* dataset:  
*to become*, *to introduce* and *to make*

52 oznews his chosen mode of transport. Les reckons crash helmets became < compulsory > about the time he started to lose his hair  
92 oznews to sack people. The coalition is also planning to introduce < compulsory > secret ballots of union members before strikes,  
122 brnews and called for tough new standards for lecturers to be made < compulsory >. The new national Institute for Learning and Teaching

The term *compulsory* can be used as an attributive adjective (60%) or as a predicative adjective (40%). In the former case, it modifies nouns such as those shown in Table 5.30 and in Figg. 5.30.1 and 5.30.2, which appear to be relevant to such fields as the administration, law and bureaucracy on the one hand (e.g. *retirement*, *insurance*, *levy*), and education, on the other (*education*, *schooling*, *attendance*, *registration/admission*).

Table 5.30 Right collocates of *compulsory*:  
*compulsory* as attributive adjective in the two corpora  
and in the *BNC* and *CWo* datasets

Frequency				
<i>Compulsory</i> as pre-modifiers	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>attendance</i>	6.55	5.54	1 (0.5%)	1 (0.5%)
<i>education</i>			7 (3.5%)	9 (4.5%)
<i>heterosexuality</i>	6.97	7.69	-	1 (0.5%)
<i>insurance</i>			1 (0.5%)	6 (3%)
<i>levy</i>	7.32	5.80	1 (0.5%)	1 (0.5%)
<i>liquidation</i>	6.14	6.15	-	1 (0.5%)
<i>purchase/ acquisition</i>	7.57	6.59	23 (11.5%)	7 (3.5%)
<i>redundancy</i>	9.56	9.58	11 (5.5%)	12 (6%)
<i>registration/ admission</i>	7.84	5.47	16 (8%)	1 (0.5%)
<i>retirement</i>			4 (2%)	3 (1.5%)
<i>schooling</i>	9.08	7.02	7 (3.5%)	-

Fig. 5.30.1 Screenshot of the right collocates of *compulsory*  
in the *BNC* dataset

24 BPS cent holding, whether or not there is a control contract. The < compulsory > acquisition procedures used in the UK after a takeover  
27 ALP need of assessment or treatment, these women did not require < compulsory > admission, with the attendant limits on civil liberties.  
39 FEP bulbous attention. This can be useful during the occasional < compulsory > attendance at official Indonesian functions, or when  
161 K4S by his agency Conrad Ritblatt he said: 'we must move towards < compulsory > registration of investment deals, lettings, mortgages  
178 A2C which characterises the true dialect speaker. Her 12 years of < compulsory > schooling had left her exactly where they had found her

Fig. 5.30.2 Screenshot of the right collocates of *compulsory*  
in the *CWo* dataset

94 usbooks heterosexual relationships in a context, no matter how disputed, of '< compulsory > heterosexuality.' The pervasive force of student  
115 brbooks transmitted the bulk of its programmes and has been financed from a < compulsory > levy, to a vastly different system in which vertical  
117 brregnews difficulties have proved impossible to rescue, resulting in < compulsory > liquidation." Under the new rules introduced in 2003,  
148 times structures. I have a couple of suggestions. First, the rule on < compulsory > purchase of annuities should be scrapped.  
160 brspok company is shaking up its electrical business and hasn't ruled out < compulsory > redundancies. Meanwhile at retail the supermarket price

In the latter case, when the term *compulsory* occurs as a predicative adjective (see Fig. 5.31), it follows a linking verb (e.g. *to be*, *to make*, *to be made*, *to become*) and qualifies noun phrases which, here as well as in the attributive function of the adjective, are relevant to the fields of education and bureaucracy (see Table 5.31.). In my datasets, the nouns qualified by *compulsory* make explicit reference to policies (e.g. *acts*, *reforms*, *regulations*, *legislations*, *laws*, *jurisdictions*; *BNC/CWo*: 14 occurrences, 7%) or official institutions (e.g. *government*, *commissions*, *inspectors*, *authorities*; *BNC*: 7 occurrences,

3.5%; *CWo*: 20 occurrences, 10%; e.g. “The Education Act (Scotland), making education *compulsory* for children from five to thirteen [...]”, *BNC*: EVJ; “Within two months the government had abolished *compulsory* religious education [...]”, *CWo*: usbooks).

Table 5.31 NP subjects in sentences in which *compulsory* is used as the predicative adjective

Frequency				
Subjects (X is <i>compulsory</i> )	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>attendance</i>	4.95	5.12	-	1 (0.5%)
<i>education</i>	5.36	-	2 (1%)	1 (0.5%)
<i>(seat)belt</i>	6.76	4.08	3 (1.5%)	1 (0.5%)
<i>insurance</i>	4.42	3.11	3 (1.5%)	2 (1%)
<i>membership/registration</i>	5.41	4.09	2 (1%)	-
<i>voting</i>	5.90	5.66	-	2 (1%)

Fig. 5.31 Screenshot of *compulsory* as predicative adjective with the verb *to be* in the *BNC* dataset

96 times and more than 50 women attend the jumma prayers. Attendance is < compulsory > for male Muslims, and the men number around 500  
 97 brbooks breakdowns (see Accidents & Breakdowns). The wearing of seat belts is < compulsory >. Motorway (autostrada) tolls are quite expensive  
 100 brephem also need your registration documents. Third party insurance is < compulsory >, and green cards are highly recommended. For regularly  
 209 safrica in the Hitler Youth during world war 2, when membership was < compulsory >, according to his autobiography. But he was never  
 141 oznews Australians abroad are the only adults for whom voting is not < compulsory >. "Being overseas is considered a valid reason for not

The meanings that the term conveys are strictly related to the terms which *compulsory* occurs with. In particular, the term *compulsory* often occurs when reference is made to the field of education in my datasets as well (44 times, 22%, in the *BNC* and 31 times, 15.5%, in the *CWo*; e.g. “In Britain *compulsory* schooling starts at age five and continues to age 16 [...]”, *BNC*: FP4; “Schools for the upper classes had existed in Britain since the mid-16th century but in 1880 *compulsory* education was introduced for the whole country - and the schoolchild emerged”, *CWo*: brbooks). In addition, the term is also used with reference to the workplace and employment issues (*BNC*: 27 occurrences, 13.5%; *CWo*: 37 occurrences, 18.5%), that is with words such as *redundancies*, *redeployment*, *work*, *workforce*, *employers*, *retirement*, *worker*, *wage control*, *contract* and others (e.g. “They are among 1,500 staff of the bank who are facing *compulsory* redundancy”, *BNC*: K5M; “Temporary residents will lose access to superannuation, with the employers’ *compulsory* 9 per cent superannuation contribution

going directly to government”, *CWo*: oznews). Finally, only in 7 (3.5%) occurrences in the *BNC* and 5 (2.5%) in the *CWo* datasets does the term *compulsory* refer to the military service (e.g. “At the 1991 annual military parade on July 14 Mitterrand proposed a reduction of *compulsory* national service from 12 to 10 months as from 1992”, *BNC*: HLB; “For boys, it arrived at 18, when *compulsory* military training began” *CWo*: usspok).

Although in the instances in my datasets the term appears to indicate ‘obligation’, it does not frequently co-occur with modal verbs such as *should* and *must*, which overtly convey the notion of ‘obligation’. As a matter of fact *should* only occurs 4 times (2%) in the *BNC* and 8 times (4%) in the *CWo*; and *must* occurs only once in both corpora (see Figg. 5.32.1 and 5.32.2).

Fig. 5.32.1 Screenshot of *compulsory* in the *BNC* dataset with the modal verbs *should* and *must*

23	K94	principal's manufacturing plant and such training should form a < compulsory > part of any agreement. Training may need to be a
56	EE9	contributions. Llewellyn-Smith proposed that insurance should be < compulsory > for all grades of workers in three trades; shipbuilding,
57	CNS	curriculum proposals, that in years four and five science should be < compulsory >; at least one arts course should be followed; and a
108	JNP	was that it was a laudable objective, but should not be made < compulsory >? Er can I answer that. We are fortunate enough in

Fig. 5.32.2 Screenshot of *compulsory* in the *CWo* dataset with the modal verbs *should* and *must*

48	brbooks	real long-term cost of what we are doing. Surely it should be < compulsory > to learn basic money management at school? For a simple
49	oznews	Iraq war was always a ridiculous idea. Her statement should be < compulsory > reading for Australians who either don't care enough about
50	safrica	a strong emphasis on good quality English teaching should be < compulsory >. Some say we have to look deep into the future before
35	brmags	Health Act order. Four conditions must be satisfied before any < compulsory > powers can be used: there must be a mental disorder (as

In the instances in my datasets, the adjective *compulsory* occurs in the following patterns (see Figg. 5.33.1 and 5.33.2): it can be followed by the preposition *for* + a noun which identifies people who are subjects to the obligation being talked about (*BNC*: 12 occurrences, 6%; *CWo*: 14 occurrences, 7%); it can also be followed *in* + a placename/year, a prepositional phrase which identifies the circumstances of the obligation (*BNC*: 5 occurrences, 2.5%; *CWo*: 3 occurrences, 1.5%); and it can also be followed by a *to*-infinitive which expresses the object of the obligation, that is, the action that is to be done (*BNC*: 6 occurrences, 3%; *CWo*: 2 occurrences, 1%).

Fig. 5.33.1 Screenshot of *compulsory* followed by prepositions in the *BNC* dataset

68	EVJ	experiment. The Education Act (Scotland), making education < compulsory > for children from five to thirteen, had only just been passed
69	HAB	to be infected with hepatitis B France: since a law making it < compulsory > for employers to offer hepatitis B vaccination to healthcare
81	EE9	to supervise school building and education. Education became < compulsory > in 1880 and free in 1891 - previously it had cost a small sum
84	EDK	of vaccination spread fast at home and abroad. It was made < compulsory > in Denmark, Norway, and Sweden around 1810. We now know that
198	BPB	can be covered with a silk to match your sweater. It is now < compulsory > to wear a back protector at most events, and if not, it is

Fig. 5.33.2 Screenshot of *compulsory* followed by prepositions  
in the *CWo* dataset

84	brspok	sort of thing in-built into the course so that it was maybe < compulsory > for everybody to have that kind of experience? <F02/> Erm
85	times	and more than 50 women attend the juma prayers. Attendance is < compulsory > for male Muslims, and the men number around 500 -so many that
104	sunnow	to swim or had life-saving skills. Swimming lessons are not < compulsory > in schools. The association will be based in Galway with five
194	brbooks	long-term cost of what we are doing, surely it should be < compulsory > to learn basic money management at school? For a simple yet
195	usbooks	hand out to students who need them. while it should not be < compulsory > to wear a name tag, every effort should be made to persuade

In conclusion, the data collected from the corpora show that the term *compulsory* appears to be used in academic writing in both the *BNC* and the *COCA* and that it is preceded by adverbs and introduced by verbs such as *to become*, *to introduce* and *to make*. In the instances retrieved in my datasets, *compulsory* is used as an attributive adjective 60% of the time, as a predicative adjective in the remaining 40%, with verbs such as *to be*, *to become*; in both cases, the terms which co-occur with *compulsory* refer to the fields of education and bureaucracy. The term also occurs in sentences related to the military field and the workplace, that is, those in which people have to respect rules and comply with policies. The adjective *compulsory* occurs in patterns in which it is followed by the preposition *for* + a noun identifying people, the preposition *in* + a placename/year identifying the circumstances of the obligation, a *to*-infinitive expressing the action that is *compulsory*. With reference to the meanings the term appears to convey, *compulsory* is used when reference is made to a law or an authority that imposes an obligation, which may be for the common good, as in the case of education.

#### 5.4.2 *Obligatory*

The adjective *obligatory* occurs 320 times in the *BNC* and 913 times in the *COCA*; despite what the raw data suggests, as Table 5.32 shows, it has a higher normalised frequency of occurrence in British English (3.32 times pmw in the *BNC*) than in American English (1.97 times pmw in the *COCA*).

Both in the *BNC* and in the *COCA* the highest frequency of occurrence of the term *obligatory* is attested in academic writing (5.74 times pmw in the *BNC* and 3.73 times pmw in the *COCA*), whereas the lowest frequency of occurrence of the term is attested in the spoken language, in which *obligatory* has less than one occurrence pmw. These frequency scores suggest that *obligatory* is more likely to occur in formal rather than in informal contexts.

Table 5.32 Frequency of occurrence of the term *obligatory* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Genres</b>				
Spoken	2 (0.62%)	0.20	53 (5.8%)	0.55
Fiction	25 (7.81%)	1.57	194 (21.25%)	2.15
Magazine	35 (10.94%)	4.82	169 (18.51%)	1.77
Newspaper	15 (4.69%)	1.43	157 (17.20%)	1.71
Non-academic	79 (24.69%)	4.79	-	-
Academic	<b>88 (27.50%)</b>	<b>5.74</b>	<b>340 (37.24%)</b>	<b>3.73</b>
Other	76 (23.75%)	3.65	-	-
TOTAL	320 (100%)	3.32	913 (100%)	1.97

The instances in my datasets show that among the immediate left collocates of the term *obligatory* there are the determiners *an* (*BNC*: 14 occurrences, 7%; *CWo*: 19 occurrences, 18.5%) and *the* (*BNC*: 48 occurrences, 24%; *CWo*: 71 occurrences, 15.5%), which are used when the adjective is used as pre-modifier of a noun (see this section below). The adjective appears to be preceded by adverbs (the most frequent ones both in the whole *BNC* and *CWo* and in my datasets are shown in Table 5.33), or verbs such as *to become* and *to make* (*BNC*: 13 occurrences, 6.5%; *CWo*: 4 occurrences, 2%).

Table 5.33 Left collocates of *obligatory*:  
adverbial pre-modifiers in the two corpora and in the *BNC* and *CWo* datasets

Pre-modifiers of <i>obligatory</i>	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentage)	<i>CWo dataset</i> (tokens and percentages)
<i>almost</i>	4.49	1.99	5 (2.5%)	4 (2%)
<i>legally</i>	6.67	-	2 (1%)	-
<i>morally</i>	6.05	5.21	-	1 (0.5%)
<i>practically</i>	6.40	-	2 (1%)	1 (0.5%)
<i>virtually</i>	3.69	-	1 (0.5%)	1 (0.5%)



Fig. 5.34.1 Screenshot of the adverbs which modifies *obligatory* in the *BNC* dataset

```

15  FB8      Hollywood executive, or agent, intent on a deal. The almost < obligatory > robot, Bishop (Lance Henriksen), is this time a hero and
105 C8R      Horizontal devolution Even if there is nothing legally < obligatory > about a particular scheme of vertical or horizontal devolution
106  CAM      mile later Quinn abandoned the motor-cycle; without the legally < obligatory > crash-helmet, he was likely to attract the attention of a
133  CK2      highly competitive. For visiting dignitaries it was practically < obligatory > to shoot a tiger. Subalterns posted to the Indian Army were
199  F9D      clarity, it should go without saying, is virtually < obligatory > in commercial sound tracks - something too many producers of

```

Fig. 5.34.2 Screenshot of the adverbs which modifies *obligatory* in the *CWo* datasets

```

21  brbooks   but only, eight-lane highway. It was almost < obligatory > to kiss the hand of Uncle this and Aunt that, or Cousin
22  oznews    "and he says that during the late '60s and '70s it was almost < obligatory > for everyone to venture into Third world culture, in search
96  usnews    to human beings in a so-called "vegetative state" is "morally < obligatory >." Cardinal McCarrick said the issue is "complex," but added
115 brbooks    When I first started playing for England it was practically < obligatory > to drink as much as you could the night before a game,
200 usbooks    curtain was vintage Caspar Neher. Mackie's costume (virtually < obligatory > since in all productions) was the invention of Harold Paulsen

```

The term *obligatory* can occur in sentences as an attributive adjective (56%) or as a predicative adjective (44%). In particular, in my datasets, when the term occurs in the former position, it is immediately followed by nouns such as those listed in Table 5.34 and exemplified in Figg. 5.35.1 and 5.35.2. As one may notice, some of the nouns listed identify requirements such as *lamps* for vehicles in traffic, others identify things that one regards as obvious or unavoidable such as the *stop* at McDonald's. In this latter case, then, the meaning conveyed is that of 'something expected' rather than 'something that must be done'.

Table 5.34 *Obligatory* as attributive adjective followed by nouns in the *BNC* and *CWo*

Modifies ( <i>obligatory</i> + X)	Frequency		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>appearance</i>	-	1.66	-	2 (1%)
<i>attendance</i>	3.38	3.50	1 (0.5%)	-
<i>character</i>	-	-	2 (1%)	-
<i>lamps/lights</i>	6.34		7 (3.5%)	-
<i>parasite</i>	5.60	4.75	-	1 (0.5%)
<i>question</i>	-	-	-	2 (1%)
<i>stop</i>	-	3.45	-	5 (2.5%)
<i>tower</i>	-	-	2 (1%)	-

Fig. 5.35.1 Screenshot of the right collocates of *obligatory* in the *BNC* dataset:

		nouns
71	A3T	revolutionaries' of Tiananmen Square. Now it has happened again. His < obligatory > attendance of the 40th birthday party of East Germany has
79	AN3	people imbue relationships of friendship with the lasting and < obligatory > character of kinship. Friendship has the advantage of
138	CE2	This point means that all the statutory provisions regarding < obligatory > lamps and reflectors, before the vehicle can lawfully be
141	CE2	constable who reports the offender for summons e.g. 'The vehicle's < obligatory > lights were not lit and the vehicle was in motion on the
198	AE6	proprietor at Haydock. The new building had an organ and the < obligatory > tower, this time 120 feet high. In addition there was a

Fig. 5.35.2 Screenshot of the right collocates of *obligatory* in the *CWo* dataset:

		nouns
44	times	that a well-adjusted person could not approach all of the < obligatory > appearances, grandiose pageantry and forced gallantry
45	usbooks	Curzon Street is big enough for a family, but other than the < obligatory > appearances during the height of the season, I can't
153	usnews	it all, unsure why his story was such a big deal. He answered < obligatory > questions about finding the cash, but was much more
183	usmags	the malls shopping for Levi's jeans, watch movies and make the < obligatory > stop at McDonald's. Though they both have big dreams
185	brbooks	China, the starting point for the grassland silk Road, and an < obligatory > stop for travellers. The mud walls that surrounded it,

In the *BNC* dataset the subjects of the instances in which *obligatory* is used as a predicative adjective reveal specific patterns; for example, as Fig. 5.36 shows, there are subject noun phrases which refer to the field of education (*education, foreign-language training, English*), others that are used in grammatical contexts (*modifiers, complement, use of passive*), still others that refer to the field of 'protective and special-equipment clothing' (*bandanas, sweaters, slim Jim ties, tight trousers, chisel toed shoes, protective clothing*) and others that regard the law and bureaucracy (*law, adherence to these rules, conscription, registration, voting*). In the *CWo* dataset, no particular patterns are evident, and the only terms which occur twice are *attendance* and *voting*, the former relevant to education and the latter to bureaucracy.

Fig. 5.36 Screenshot of *obligatory* as predicative adjective in the *BNC* dataset

78	CCV	be able to discuss the contexts in which Standard English is < obligatory > and those where its use is preferable for social
84	EAD	establishment of legislative, executive and judicial offices. Law is < obligatory > because it is authoritative. Oakeshott, however, also
89	FRL	only function. In other languages, the use of the passive is < obligatory > in certain contexts; for instance, the passive has
36	B14	strict contamination control measures and protective clothing are < obligatory >. Where workers make sharp equipment, or use machines
201	HKR	registered voters was 1,732,000 aged 16 and over. Registration was < obligatory > but voting was optional. The President was elected by

In my datasets, the term *obligatory* occurs in instances in which words such as *religions, Acts, laws, rules, Conventions, Council, a Moral authority, the Parliament* and others are used, which therefore see the law or some authority involved (*BNC*: 13 occurrences, 6.5%; *CWo*: 19 occurrences, 9.5%; e.g. "The Housing and Town Planning Act made it *obligatory* for local authorities to prepare surveys of their housing needs, to draw up plans to deal with them, and to carry out their schemes", *BNC*: G05; "Acts of Parliament went so far as to make *obligatory* the use of woollen cloth for mourning clothes", *CWo*: usbooks). In addition, when the term is used to indicate something done to obey a rule, the term occurs 13 times (6.5%) in the *BNC* and once in the *CWo* as referring to grammar rules that need to be respected in order to express grammatically

correct sentences (e.g. “In other languages, the use of the passive is *obligatory* in certain contexts [...]”, *BNC*: FRL). Finally, the term *obligatory* is also used in sentences expressing the idea that something is to be done because everyone expects you to do so or because it is considered the proper thing to do (*BNC*: 61 occurrences, 30.5%; *CWo*: 65 occurrences, 32.5%; e.g. “For visiting dignitaries it was practically *obligatory* to shoot a tiger”, *BNC*: CK2; “When I first started playing for England it was practically *obligatory* to drink as much as you could the night before a game [...]”, *CWo*: brbooks).

As in the case of *compulsory*, *obligatory* too is used to indicate ‘obligation’, and in this case too, in my datasets deontic modal verbs such as *should* and *must*, which encode ‘obligation’, are infrequent: *should* only occurs 5 times (2.5%) in the *BNC* and 4 times (2%) in the *CWo*; and *must* occurs 6 times (3%) in the *BNC* and only once in the *CWo* (see Figg. 5.37.1 and 5.37.2).

Fig. 5.37.1 Screenshot of *obligatory* in the *BNC* dataset with the modal verbs *should* and *must*

```

14  FF0      enough tell you what it is he should do next. It is almost < obligatory > that the private-eye should be a person who looks at society
41  J79      Society has directed that the following steps by a solicitor are < obligatory >: (1) A solicitor must, with his clients' authority, at once
48  A9W      way of reducing carbon dioxide emissions, should be < obligatory > in new buildings. Improved efficiency in electrical appliances
65  CS2      In order to be all-things-considered obligatory a prima facie < obligatory > act must not be incompatible with some more stringent prima
93  FF0      But the shabbiness does indicate one other quality that is < obligatory >; the private-eye must be a man of integrity. He may not exactly

```

Fig. 5.37.2 Screenshot of *obligatory* in the *CWo* dataset with the modal verbs *should* and *must*

```

8   oznews   girl off while she can still fetch a dollar. Alternatively, < obligatory > taxpayer funding of the ABC should be made optional.
57  sunnow   it that a chief Justice re- tires at 72 surely it should be < obligatory > to retire at 65. GORDON CUNNINGHAM Dublin 13. Piling on the
60  usbooks  apparatus. The nucleic acid creatures must have been < obligatory > parasites from the start, preying upon the protein creatures
115 usspok   the A-K-P might still have an Islamist agenda. Voting is < obligatory > in Turkey. unofficial results of the poll should be known by
191 oznews   girl off while she can still fetch a dollar. Alternatively, < obligatory > taxpayer funding of the ABC should be made optional.

```

In my datasets the adjective *obligatory* occurs in the following patterns (see Figg. 5.38.1 and 5.38.2): it can be followed by the preposition *for* + a noun identifying people who have to comply with the obligation (*BNC*: 9 occurrences, 4.5%; *CWo*: 12 occurrences, 6%); it can be followed by *in* + a placename identifying the place in which the obligation is effective (*BNC*: 11 occurrences, 5.5%; *CWo*: 5 occurrences, 2.5%) and it can be followed by a *to*-infinitive indicating the action that is obligatory (*BNC*: 6 occurrences, 3%; *CWo*: 8 occurrences, 4%).

Fig. 5.38.1 Screenshot of *obligatory* followed by prepositions in the *BNC* dataset

```

99  CAK      issue ('Hooked on IT'), and making foreign-language training < obligatory > for company executives (as well as MPs, especially left
104 HC5      giving due weighting to quality as well as price. It will become < obligatory > for members in practice to hold adequate professional
124 EX5      half-classes like Quantitative Methods or computing, which are < obligatory > in the Business School, may carry only a half
128 C89      Feb/March, '92) highlighted, health labelling of art materials is < obligatory > in the United States, though not, as yet, in this country.
196 CK2      highly competitive. For visiting dignitaries it was practically < obligatory > to shoot a tiger. Subalterns posted to the Indian Army

```

Fig. 5.38.2 Screenshot of *obligatory* followed by prepositions in the *CWo* dataset

```

97  usbooks staircase" leading from kindergarten to university. Attendance was < obligatory > for school-age children of both sexes, who were to be
111 brspok  abstention is by local standards likely to be high: voting is < obligatory > in Brazil but voters can register protest by leaving the
112 brboo somewhere in the canopy way above. Half-seen birds such as these are < obligatory > in Eden. when the waterfall is reached at last, some of
199 brbooks kingdom's beautiful, but only, eight-lane highway. It was almost < obligatory > to kiss the hand of Uncle this and Aunt that, or Cousin
202 times  might introduce pork instead of beef. I now know that it is < obligatory > to serve beef -because we wouldn't want anybody to think

```

In conclusion, as instantiated in my datasets, the term *obligatory* seems to be largely used in the field of academic writing both in the *BNC* and the *COCA*; it is preceded by adverbs and by verbs such as *to become* and *to make*. In 56% of the occurrences in my datasets *obligatory* is used as an attributive adjective and in 44% of the occurrences as a predicative adjective, and in both cases, it is attested in sentences referring to education, grammar, clothing and bureaucracy. The adjective is also used in a few phraseological patterns; followed by the preposition *for* + a noun denoting the people involved in the obligation; followed by *in* + a placename indicating the place in which the obligation has to be obeyed; and followed by a *to*-infinitive signalling the action that has or has not to be carried out. With reference to the meaning conveyed, the concordances show that the term *obligatory* is more likely to occur when related to the meaning of necessity, that is it indicates a moral duty rather than a legal one. We could also say that *obligatory* indicates, first, a personal sense of duty and second, the obligation to respect rules and laws.

### 5.4.3 Mandatory

The examination of the frequencies of occurrence of term *mandatory* shows that the term occurs 959 times in the *BNC* and 5,287 times in the *COCA*, with a frequency score that is around 10 times pmw in both corpora (9.96 times pmw in the *BNC* and 11.39 times in the *COCA*). As Table 5.35 shows, in the two corpora the term *mandatory* frequently occurs in academic genres (almost 16/18 times pmw).

Among the immediate left collocates of the adjective *mandatory*, in my datasets we find the determiners *a* (*BNC*: 27 occurrences, 13.5%; *CWo*: 29 occurrences, 14.5%) and *the* (*BNC*: 22 occurrences, 11%; *CWo*: 25 occurrences, 12.5%), which precede *mandatory* when this is used as an attributive adjective (see this section below). The adjective is also immediately preceded by verbs such as *to become* and *to make* (see Figg. 5.39.1 and 5.39.2), which may be followed by the syntactic direct object *it*, co-referential with a semantic direct object appearing after *mandatory* (*BNC*: 10 occurrences, 5%; *CWo*: 18 occurrence, 9%).

Table 5.35 Frequency of occurrence of the term *mandatory* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Genres</b>				
Spoken	30 (3.13%)	3.01	1,033 (19.54%)	10.81
Fiction	14 (1.46%)	0.88	204 (3.86%)	2.26
Magazine	31 (3.23%)	4.27	1,009 (19.08%)	10.56
Newspaper	66 (6.88%)	6.31	1,393 (26.35%)	15.19
Non-academic	109 (11.37%)	6.61	-	-
Academic	<b>244 (25.44%)</b>	<b>15.91</b>	<b>1,648 (31.17%)</b>	<b>18.10</b>
Other	465 (48.49%)	22.32	-	-
<b>TOTAL</b>	<b>959 (100%)</b>	<b>9.96</b>	<b>5,287 (100%)</b>	<b>11.39</b>

Fig. 5.39.1 Screenshot of *mandatory* in the *BNC* dataset:  
*to become and to make*

103 FBW practice were replaced by a caution, well before the caution became < mandatory > in 1848. Nevertheless, although the misuse of judicial  
104 FS6 resources for it are more widely available, the provision has become < mandatory >. For example, local authorities had power to provide a  
136 ACR chemicals of the UK. In Europe, pressure is growing to make it < mandatory > to have some form of closed transfer system for chemicals.  
140 CHT Any course you choose to undertake before updating is made < mandatory > may not later satisfy the Boards if you decide to delay  
143 EDL competent to issue and execute letters rogatory but it makes < mandatory > as between parties to the Protocol the use of Central

Fig. 5.39.2 Screenshot of *mandatory* in the *CWo* dataset:  
*to become and to make*

8 oznews girl off while she can still fetch a dollar. Alternatively, < obligatory > taxpayer funding of the ABC should be made optional.  
57 sunnow it that a chief Justice re- tires at 72 Surely it should be < obligatory > to retire at 65. GORDON CUNNINGHAM Dublin 13. Piling on the  
60 usbooks apparatus. The nucleic acid creatures must have been < obligatory > parasites from the start, preying upon the protein creatures  
115 usspok the A-K-P might still have an Islamist agenda. voting is < obligatory > in Turkey. unofficial results of the poll should be known by  
191 oznews girl off while she can still fetch a dollar. Alternatively, < obligatory > taxpayer funding of the ABC should be made optional.

With reference to the right collocates of *mandatory*, in the *BNC* we find that there are 35 occurrences (17%) in which the term is followed by *Input/Input Field/Display Field* (see Fig. 5.40). These instances may represent standard phrases used in the field of informatics to give instructions, as the instances below suggest.

Fig. 5.40 Screenshot of *mandatory* followed by *Input/Input field/Display Field* in the *BNC* dataset

16 HWF SPR title and originator. Do you accept this SPR (Y/N) (<Mandatory > Input Field) To accept the SPR enter Y and press the RETURN  
17 HWF corrected before the SSR can be updated. Are you sure (Y/N) (<Mandatory > Input Field) Finally, confirm that the details are correct  
20 HWF or deinstallations will be accepted. Are you sure (Y/N) (<Mandatory > Input Field) If the details are correct and you wish to proceed  
21 HWF characters long, including spaces. Originators address (<Mandatory > Input/Display Field) The address of the originator must be  
51 HWF Appendix A. You must provide the following details: User (<Mandatory > Input Field) The name of the LIFESPAN user whose packages

As in the cases of *compulsory* and *obligatory*, *mandatory* too can be used as an attributive adjective (68.5%) or as a predicative adjective (31.5%). In the former case, it

can be followed by nouns such as those shown in Table 5.36, which refer to the legal field.

Table 5.36 Right collocates of *mandatory*:  
*mandatory* as attributive adjective followed by nouns in the *BNC* and *CWo*

Modifies ( <i>mandatory</i> + X)	Frequency		<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)		
<i>challenger</i>	5.97	7.82	2 (1%)	2 (1%)
<i>detention</i>	-	7.36	4 (2%)	-
<i>injunction</i>	5.68	5.12	1 (0.5%)	1 (0.5%)
<i>life sentence</i>	-	-	4 (2%)	4 (2%)
<i>overtime</i>	4.77	6.42	-	-
<i>penalty(ies)</i>	3.98	-	4 (2%)	-
<i>privilege</i>	6.32	-	4 (2%)	-
<i>repatriation</i>	8.27	8.60	5 (2.5%)	2 (1%)
<i>requirement(s)</i>	-	5.13	6 (3%)	1 (0.5%)
<i>retirement</i>	4.81	5.43	1 (0.5%)	7 (3.5%)
<i>sanction(s)</i>	6.49	5.08	2 (1%)	1 (0.5%)
<i>sentence</i>	6.73	7.32	4 (2%)	7 (3.5%)

Fig. 5.41.1 Screenshot of *mandatory* used as attributive adjective in the *BNC* dataset

155 ACJ reason for these doctrines. It has been suggested here that the < mandatory > penalty is relevant but not critical. The key issues are,  
164 HWF tasks such as loading the offline media. Module owner < Mandatory > privileges None The module owner is generally the person  
175 A9N colony's administration to go ahead with the first batch of < mandatory > repatriations this week. Mr Peter Haliday, the district  
183 CKP asked. However, within the labour movement, interest in < mandatory > retirement was growing. Many trade unionists, acutely  
189 CRT prison service for some years now. Life imprisonment is the < mandatory > sentence for murder but it is also the maximum sentence for

Fig. 5.41.2 Screenshot of *mandatory* used as attributive adjective in the *CWo* dataset

51 times that. " If he must yield to his older brother, who is the < mandatory > challenger for the WBC title (the only belt 37-year-old  
65 oznews Human Rights Commission has found that Australia's policy of < mandatory > detention breaches international law, aligning us with some  
66 oznews between quarantine for cats and dogs entering Australia and the < mandatory > detention of people who come here without visas. But it's  
112 usbooks would obstruct any lot's view. The trial court granted a < mandatory > injunction requiring both defendants to trim their pine  
159 nznews training and refresher courses so that officers understand the < mandatory > requirements of the pursuit policy and the reasoning behind

In the latter case, when the term *mandatory* occurs as a predicative adjective, it follows linking verbs such as *to be*, *to become* and *to make*; in such instances, the subject NPs denote policies, injunctions, duties and legal consequences, as shown in Table 5.37. In addition to the terms listed in the table below, other terms, which occurs only once in

the datasets, identify equipment items, such as protective clothing (e.g. *helmets, protective headgear, head protection, flak jacket*).

Table 5.37 Subject in sentences in which *mandatory* is used as predicative adjective

Subjects (X is <i>mandatory</i> )	Frequency			
	<i>BNC</i> (pmw)	<i>CWo</i> (pmw)	<i>BNC dataset</i> (tokens and percentages)	<i>CWo dataset</i> (tokens and percentages)
<i>death penalty</i>	-	-	1 (0.5%)	1 (0.5%)
<i>insurance</i>	1.74	1.30	1 (0.5%)	1 (0.5%)
<i>penalty</i>	3.98	1.64	-	-
<i>provision</i>	2.92	-	2 (1%)	-

With reference to the meanings conveyed by *mandatory*, in 47 occurrences (23.5%) in the *BNC* and 48 (24%) in the *CWo* there is explicit reference to *Acts, reforms, laws, regulations, rules, councils, Courts, governments* and so on as the sources of the legal obligation conveyed by the term *mandatory*; in particular, an occurrence in the *BNC* clearly suggests that *mandatory* is used in opposition to something which is done illegally (e.g. “Bribery is a difficult area; what is illegal in one place may be almost *mandatory* in another”, *BNC*: A60). Moreover, the term *mandatory* appears to occur also in instances that make reference to life sentences which condemn people to life imprisonment (*BNC*: 13 occurrences, 6.5%; *CWo*: 24 occurrences, 12%; e.g. “In practice the only exceptions to judicial sentencing are fixed penalties and obligatory disqualification for certain Road Traffic offences, and the *mandatory* penalty of life imprisonment for murder”, *BNC*: EEC; “His sentencing is set for July 22, although the aggravated murder conviction carries a *mandatory* sentence of life in prison without possibility of parole”, *CWo*: usnews).

The findings appear to show that the term *mandatory* is used to signal something that necessarily has to be done; as in the case of the other two adjectives, *mandatory* too occurs with the deontic modal verbs *should* and *must*, but with higher frequency scores, that is *should* occurs 7 times (3.5%) in the *BNC* and 11 times (5.5%) in the *CWo* and *must* occurs 13 times (6.5%) in the *BNC* and twice (1%) in the *CWo* occurrences; see Figg. 5.42.1 and 5.42.2.

Fig. 5.42.1 Screenshot of *mandatory* in the *BNC* dataset with the modal verbs *should* and *must*

25	HWF	A. You must provide the following details: Charge code ( < Mandatory > Input Field) The new Charge Code to be created. The Charge
29	HWF	making entries in the following fields. Module filename ( < Mandatory > Input Field) You must enter the name of the file in to which
101	GOK	issue and wearing of protective headgear in the system should be < mandatory >, and 'non smoking' regulations rigidly enforced. The latter
196	FD9	injunction should be varied in order to give effect to the < mandatory > terms of section 39, so as to permit the defendants to comply
199	CLP	which might help or hinder, and he must also stay within the < mandatory > operating instructions. Unfortunately, because such situations

Fig. 5.42.2 Screenshot of *mandatory* in the *CWo* dataset with the modal verbs *should* and *must*

70	oznews	move opinions. Her analysis of ''Feminine Arts '' should be < mandatory > reading for all teenage girls. Although the book is one which
80	usmags	new vehicle in the European Union must comply with certain < mandatory > technical requirements. Once a so-called "type approval "has
114	nznews	violent crime." He said the law should be changed to make it < mandatory > for child murderers to spend the rest of their life in prison
126	indnews	Exchange Board of India, and forward pricing should be made < mandatory >. until then, long-term investors in open-ended equity funds
182	times	that. " If he must yield to his older brother, who is the < mandatory > challenger for the WBC title (the only belt 37-year-old Lewis

In the instances in my datasets, the adjective *mandatory* occurs in two main phraseological patterns (see Figg. 5.43.1 and 5.43.2): it can be followed by the prepositions *for* + a noun identifying people who have to obey to the obligation (*BNC*: 8 occurrences, 4%; *CWo*: 21 occurrences, 10.5%) and *in* + a placename/year denoting the circumstances in which the obligation takes place (*BNC*: 6 occurrences, 3%; *CWo*: 7 occurrences, 3.5%).

Fig. 5.43.1 Screenshot of *mandatory* followed by prepositions in the *BNC* dataset

61	HLA	penal code in line with Islamic law by making the death penalty < mandatory > for anyone convicted of defaming the Prophet Mohammed.
62	EEH	harmonization at the consumer level and their implementation is < mandatory > for member States. Given the objectives of the community
65	CBT	ascertain the procedure for the declaration of dividends. Is it < mandatory > for the company to determine the dividends and their date
73	FBW	practice were replaced by a caution, well before the caution became < mandatory > in 1848. Nevertheless, although the misuse of judicial
77	BNV	before filing our flight-plan to santiago. Flight-plans are < mandatory > in Spain for VFR as well as IFR flights; the landing fees

Fig. 5.43.2 Screenshot of *mandatory* followed by prepositions in the *CWo* dataset

85	nznews	violent crime." He said the law should be changed to make it < mandatory > for child murderers to spend the rest of their life
86	usbooks	has long been a subject of controversy. The interpretation < mandatory > for Communist historians and favored by western socialists
106	indnews	maximise the gains for farmers. While this will make subsidies less < mandatory > in future, the potential for returns, hopes chidambaram,
107	cannews	youth crime and effective learning and values. Sports should be < mandatory > in schools. But the Ottawa conference will hear some
109	oznews	in this type of work. Are flak jacket and head protection not < mandatory > in situations where danger lurks? My heart goes out to his

To sum up, the instances of the term *mandatory* indicate that it is used in academic writing both in the *BNC* and the *COCA*, and that it is often introduced by verbs such as *to become* and *to make*. The adjective *mandatory* is used in 68.5% of the occurrences as an attributive adjective in the judicial field and in 31.5% of the occurrences as a predicative adjective when reference is made to protective equipment, required by law, and the legal field (esp. imprisonment). *Mandatory* occurs in such patterns as: followed by prepositions such as *for* + a noun denoting people who have to obey the imposition, and *in* + a placename/year indicating the circumstances of the obligation. The adjective can occur in informatics to indicate a field in a form that has to



be filled in; it can also be used to express legal obligations determined by regulatory agencies and bodies such as governments and other powerful authorities, and the policies and rules issued by them.

#### 5.4.4 Conclusion

In conclusion, the findings related to the adjectives *compulsory*, *obligatory* and *mandatory* show that adjectives have different frequency scores and more specifically, *compulsory* is the term with the highest frequency scores in the British English, whereas *mandatory* is the one with the highest scores in the American English. In both the *BNC* and the *COCA* all the terms have their higher frequency of occurrence in the academic field.

The three terms appear to be used with the verbs *to become* and *to make* and as attributive or predicative adjectives. Both *compulsory* and *obligatory* are frequently preceded by adverbs such as *almost* and followed by the prepositional constructions with *for* followed by nouns identifying people that have to obey to the obligation, *in* followed by placenames denoting the circumstances in which the obligation has to be obeyed (only *compulsory* has the construction *in* + year), or a *to*-infinitive indicating the action that has to be obeyed, while *mandatory* is only followed by *for* + nouns denoting people and *in* + a placename/year in which the obligation takes place.

With reference to the meanings that these terms convey, the three adjectives appears to occur in sentences in which people are obliged by a law, a reform or an authority to behave or act in a certain way. The term *compulsory* is used when talking about something that must be done in the educational, work and military fields; *obligatory* is also used when we are talking of grammar rules to be obeyed; and *mandatory* is often associated with the legal field, in particular when reference is made to life sentences and imprisonment. In addition to this, only *obligatory* seems to convey the meaning of necessity rather than of obligation.

As in the case of the nouns and the verbs (see sections 5.2.4 and 5.3.3), the findings from the entire corpora do not fully match those from my datasets and in particular, the former show recurring patterns that so not appear to be instantiated in the latter.

Table 5.38 Collocational patterns of *compulsory*, *obligatory* and *mandatory*

		<i>Compulsory</i>	<i>Obligatory</i>	<i>Mandatory</i>
<b>Position</b>	Attributive adjective	60%	56%	68.5%
	Predicative adjective	40%	44%	31.5%
<b>Verbs</b>	to introduce	x	-	-
	to become	x	x	x
	to make	x	x	x
<b>Prepositions</b>	<i>For</i> + noun identifying people	x	x	x
	<i>In</i> + placename in which the obligation has to be obeyed	x	x	x
	<i>In</i> + year during which the obligation has to be obeyed	x	-	x
	<i>To</i> + verb signalling the action that has to be carried out	x	x	-
<b>Meanings/use</b>	To behave in a certain way because of a law, a reform or an authority	x	x	x
	Education	x	x	-
	Workplace	x	-	-
	Military	x	-	-
	Grammar	-	x	-
	Legal	-	-	x
	Bureaucracy	x	x	x

Therefore, the findings relevant to the three adjectives, which are summarized in Table 5.38, show that the adjectives are partially interchangeable, that is they can be replaced one with another when referring to an obligation imposed by a law or in bureaucracy. On the other hand, they also appear to have different semantic preferences: when referring to education, *compulsory* and *obligatory* tend to be used; when talking about workplaces and the military, *compulsory* is preferred; in grammar, *obligatory* is the default choice, while in the legal field, *mandatory* is a frequent choice.

### 5.5 The adverbs

In the following sub-sections I will report on the findings relevant to the adverbs *maybe*, *perhaps* and *possibly*. In every sub-section I will include the raw and normalised frequency scores of the term being analysed and its distribution across genres in the

*BNC* and the *COCA*. Then, by looking at both my datasets and the whole *BNC* and *CWo*, I will analyse the left and right collocates of the adverb in question and its role in the sentence, that is, if it occurs as a clause adverb, as a “normal” adverb or in both roles. Finally, I will describe the meanings the term appears to convey.

### 5.5.1 *Maybe*

The adverb *maybe*<sup>29</sup> occurs 9,883 times in the *BNC* and 126,011 times in the *COCA*, that is 102.67 time pmw in the former and 271.38 times pmw in the latter (see Table 5.39). Table 5.39 also shows that the term has its highest frequency score in fiction in both corpora and, in particular, that it occurs 3,815 times in the *BNC* (38.60%: more than a third out of the total amount of tokens) and 55,310 times in the *COCA* (43.89%: almost half of the total amount of tokens).

Table 5. 39 Frequency of occurrence of the term *maybe* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Genres</b>				
Spoken	3,107 (31.44%)	311.83	38,707 (30.72%)	404.98
Fiction	<b>3,815 (38.60%)</b>	<b>239.80</b>	<b>55,310 (43.89%)</b>	<b>611.64</b>
Magazine	713 (7.21%)	98.18	15,003 (11.90%)	157.00
Newspaper	565 (5.72%)	53.98	14,096 (11.19%)	153.70
Non-academic	462 (4.67%)	28.01	-	-
Academic	201 (2.03%)	13.11	2,895 (2.30%)	31.79
Other	1,020 (10.33%)	48.96	-	-
<b>TOTAL</b>	<b>9,883 (100%)</b>	<b>102.67</b>	<b>126,011 (100%)</b>	<b>271.38</b>

As Figg. 5.44.1 and 5.44.2 show, the term *maybe* often occurs as a sentence adverb, that is an adverb modifying the whole sentence, in sentence- or clause-initial position; it is therefore often preceded by punctuation marks such as the full stops, commas, inverted commas and others (*BNC*: 108 occurrences, 54%; *CWo*: 120 occurrences, 60%), the conjunctions *and/or/but* (*BNC*: 34 occurrences, 17 %; *CWo*: 36

<sup>29</sup> The form *maybes* was excluded from the research.

occurrences, 18%) and the adverbs *then* and *well* (*BNC*: 11 occurrences, 5.5%; *CWo*: 3 occurrences, 1.5%).

Fig. 5.44.1 Screenshot of the left collocates of *maybe* in the *BNC* dataset

3	CH2	we're asleep when it should be the other way round. < Maybe > he'll change when he starts school next year. 'I can't
120	KRT	though it is a pity what's happened to the R S C isn't it, and < maybe > out of these ashes will rise a new Royal de Vere Company, one
136	JY3	was no reply. She'd assumed he'd driven back to London, but < maybe > he hadn't. Maybe he was still in Oxfordshire? Another thought
193	HRA	under French civil law.' 'Good,' said the major. 'Then < maybe > you can give me a hand. We've got to get one set of charges
206	JYM	that you're gonna say yes no even though like you said, well < maybe > I do put it that way. But we know really you could say yes

Fig. 5.44.2 Screenshot of the left collocates of *maybe* in the *CWo* dataset

41	usbooks	have a little fun. You could stay over a couple of days too, < maybe > even bring your wives-at your own expense, of course-if you
44	usbooks	law, so no regrets." "None?" asked Cooper doubtfully. "well, < maybe > a few little ones," her father said. "The point is, I want
141	sunnow	burdens yourself. You need a little help with your mum and < maybe > the time has come to ask for it. I also feel that you just
150	oznews	radio sleuth's signature phrase. Just joking, of course, but < maybe >, after all the real and amateur psychologists, sociologists
174	usbooks	hadn't gotten the phone tap, not enough evidence or time, or < maybe > they were too dumb to figure out my favorite place in the world

*Maybe* is rarely followed by punctuation marks such as full stops, commas (*BNC*: 19 occurrences, 8.5%; *CWo*: 22 occurrences, 11%), that is, it is rarely found at the end of a clause or sentence. In addition, as Fig. 5.45 shows, in a few occurrences it is also used as a modifier of a number (8 occurrences in the *BNC*, 4% and 11 occurrences in the *CWo*, 5.5%).

Fig. 5.45 Screenshot of *maybe* in the *BNC* dataset: before a number

23	K2A	support on a Mary Black tour and from doing folk club gigs for < maybe > 100 customers I suddenly found myself doing 20 nights in a
55	BNT	tasted one or three of the strawberry, I tried a couple, or < maybe > half a dozen, of the raspberry. The jam, it turns out, is home
56	G02	nothing to keep them up, not even ankles, to be accurate, well < maybe > half an inch, no more... but the image changed. Snap! Just like
119	HET	needed, the only new thing wa was a very aged er er typewriter, < maybe > Mm. two. And there was the best one of course, your cousin
148	BPH	Much larger sums were involved, almost inevitably in five, or < maybe > six, figures. Some paid in cash, most borrowed the whole or

The right collocates of *maybe* (see Figg. 5.46.1 and 5.46.2) are often noun phrase subjects, which can be personal pronouns (*BNC*: 70 occurrences, 35%; *CWo*: 66 occurrences, 33%), proper names or NPs, which sometimes can be preceded by determiners or adjectives, with a common noun as their head (*BNC*: 48 occurrences, 24%; *CWo*: 51 occurrences, 25.5%). Other right collocates are verbs which indicate the action which cause uncertainty (*BNC*: 19 occurrences, 9.8%; *CWo*: 15 occurrences, 7.5%)

Fig. 5.46.1 Screenshot of the right collocates of *maybe* in the *BNC* dataset

36	KRH	merely subjective, or are delusory or are soft-headed or Though < maybe > Ayer would want to say that on another occasion. I mean Ayer
60	ASA	bound through. I think he was convinced the green was fast, but < maybe > he had forgotten we were coming out of the rough and the
118	F71	interpretation yes is different, and you need an original yes and you < maybe > miss something, from this you have your own, that's right and
120	HR9	close to Marr, wasn't I? And he was the only guy on to them. < Maybe > my dad was really and truly dead and what I had been looking
172	CDV	flowering only in 'another country': it sounds like virtue. < Maybe > the shepherd lad was the Good Shepherd himself. As for the

Fig. 5.46.2 Screenshot of the right collocates of *maybe* in the *CWo* dataset

36	usbooks	Drexel's Peter Ackerman, the trader stepping into Milken's shoes. < Maybe > Ackerman could come up with something Gutfreund could live
49	brbooks	shot. I asked whether we could get nearer the hospital and < maybe > drive round so that I could do a tracking shot from the van
81	usbooks	yeah, I'm good, I'm real good at this stuff. Good training, < maybe > I have a knack for it." kelly shrugged, not wanting to say
94	cannews	it happen. To our own children or many others who we know. Or < maybe > it has even happened to us. Well, Elizabeth French is one of
154	brnews	Dr Sheila Cassidy, and no doubt of other Britons as well. But < maybe > the spanish judge had not included it in his extradition request

After having looked at the instances in my datasets, I identified some meanings which are conveyed by the adverb *maybe*. The term appears to be used when someone is uncertain about something and does not want to fully commit to the truthfulness or validity or appropriacy of his/her statement or opinion; in such cases, a statement is expressed which conveys the speaker's/writer's tentativeness (*BNC*: 176 occurrences, 88%; *CWo*: 173 occurrences, 86.5%). This meaning may be stressed by the use of filler words such as *erm*, *mm*, *er* (e.g. "[...] the only new thing wa was a very aged er er typewriter, *maybe* Mm. two. And there was the best one of course, your cousin [...]", *BNC*: HET; "*Maybe* he won't call my supervisor if I let it pass", *CWo*: usbooks). *Maybe* also signals uncertainty and is frequently used to make suggestions, in association with *should* (*BNC*: 10 occurrences, 5%; *CWo*: 4 occurrences, 2%) or observations about something that will probably happen or that have already happened, in association with *can/could* (*BNC*: 12 occurrences, 6%; *CWo*: 18 occurrences, 9%; e.g. "*Maybe* Ross should never had got married in the first place", *BNC*: JXX; "*Maybe* one could open an attack with Kumble instead of the usual Srinath and Prasad", *CWo*: indnews). The term also occurs as an answer to questions when one interlocutor does not want to say explicitly if he/she agrees or not with the other interlocutor (*BNC*: 1 occurrence, 0.5%). Finally, my dataset shows some occurrences in which the term is used in association with the conjunction *or* when giving options, with the meaning of "you can do one or the other" (*BNC*: 16 occurrences, 8%; *CWo*: 14 occurrences, 7%; e.g. "Cut me 'ead off, he said grinning, Or *maybe* grow a fringe", *BNC*: ARP; "But he missed and shot Dr. Sayers instead. Or *maybe* Dimitri didn't miss", *CWo*: usbooks).

To sum up, the adverb *maybe* occurs with high frequency scores in fiction both in the *BNC* and the *COCA*, and is mainly placed at the beginning and only rarely at the end of sentences and clauses. According to the instances retrieved in my datasets, the term *maybe* can occur as a clause adverb (95.25%) or as an adverb before a value (4.75%). In most of the sentences, even in those in which the speaker is giving suggestions, the occurrence of the adverb *maybe* expresses a meaning of uncertainty and hesitation as if the interlocutor were not sure of what he/she is saying.

### 5.5.2 Perhaps

The adverb *perhaps* is instantiated 33,294 times in British English and 97,740 times in American English, that is 345.86 times pmw in the *BNC* and 210.49 times pmw in the *COCA*. As is shown in Table 5.40, *perhaps* occurs more frequently in fiction both in the *BNC* (27.23%) and in the *COCA* (24.23%), even if in the latter the term has also a high frequency score in academic writing (23.18%).

Table 5.40 Frequency of occurrence of the term *perhaps* in the *BNC* and the *COCA*

	<i>BNC</i>		<i>COCA</i>	
	Tokens	Frequency pmw	Tokens	Frequency pmw
<b>Genres</b>				
Spoken	4,586 (13.78%)	460.27	17,512 (17.92%)	183.22
Fiction	<b>9,067 (27.23%)</b>	<b>569.92</b>	<b>23,683 (24.23%)</b>	<b>261.89</b>
Magazine	1,912 (5.74%)	263.29	19,540 (19.99%)	204.48
Newspaper	1,655 (4.97%)	158.12	14,349 (14.68%)	156.45
Non-academic	4,745 (14.25%)	287.66	-	-
Academic	5,355 (16.08%)	349.28	<b>22,656 (23.18%)</b>	<b>248.79</b>
Other	5,974 (17.95%)	286.73	-	-
TOTAL	33,294 (100%)	345.86	97,740 (100%)	210.49

The adverb *perhaps* often occurs at the beginning of a clause or sentence as a sentence adverb (*BNC*: 98%; *CWo*: 97.5%; see Fig. 5.47); it is therefore frequently preceded by punctuation marks such as full stops and commas (*BNC*: 110 occurrences, 55%; *CWo*: 108 occurrences, 54%), the conjunctions *and/or/but* (*BNC*: 29 occurrences, 14.5%; *CWo*: 30 occurrences, 15%) and frequently followed by subjects of sentences/clauses such as personal pronouns (*BNC*: 50 occurrences, 25%; *CWo*: 44 occurrences, 22%). *Perhaps* can also precede and modify a number or value (*BNC*: 2%; *CWo*: 2.5%; see Fig. 5.48), which is thus qualified as not exact, but approximate (i.e. in the meaning of ‘more or less’).

Fig. 5.47 Screenshot of *perhaps* in the *BNC* dataset:  
sentence adverb

10	CB5	Ernest. But it's not the same for a woman, it can't be.' < Perhaps > it is in America. Who knows? And I'm not afraid.' But of
102	F9C	to his eyes and squeezed them together. He felt very tired. < Perhaps > it was the smoke. Perhaps it was William. Perhaps it was life
119	CRE	once. Men and women of great authority had walked in it, and < perhaps > left some ghost of themselves to calm him in extremis. No
142	CU1	contracting market. Peter Rochford deserved more scope. But < perhaps > he'd left it too late. Too many things, some self-induced,
188	H7F	books to keep warm; it offended her. Besides, she had said, < perhaps > they would take more time repairing the heating system if they

Fig. 5.48 Screenshot of *perhaps* in the *CWo* dataset:  
adverb before a number

25	brspok	successful. Over 1,800 websites exist in the gov. uk domain with < perhaps > 40 more every month. But adding 'directly relevant' or 'highly
69	usbooks	two of the men bore a heavy burden. Leaphorn followed them for < perhaps > fifty yards. The tracks were fading fast and he lost them when
190	brbooks	dark curved ridge that closed the top of the valley there was < perhaps > three miles of air, perfectly clear, but somehow evident. It
193	brspok	All you would need to do is dose this on perhaps a monthly, or < perhaps > two monthly basis and all your grease related drain problems

With reference to sequencing patterns, *perhaps* can be found at the beginning of a sentence, after the main verb or between the modal verb and the base form: in my datasets, among the verbs which occur followed by *perhaps*, we find *to be* and *to think*, which occur 10 times (5%) in the *BNC* and 17 times (8.5%) in the *CWo*.

Fig. 5.49 Screenshot of *perhaps* preceded by verbs  
in the *BNC* dataset

133	EWB	is not arbitrary. The matrix equation $A^{-1}b$ arises < perhaps > most commonly in the study of the natural frequencies and
164	G0K	one sprinkler system. As a basis for this calculation it is < perhaps > appropriate to assume that a single sprinkler system should
171	B7C	explain the failure of his attempts to do so. Although he might < perhaps > have done better to observe Wittgenstein's adage, 'whereof
197	JA9	Yes, cert Then. in the week before somebody comes then I think < perhaps > a note saying, so and so's starting. Okay. Mhm. We haven't
209	FRL	categories - to talking about the text as a unit of meaning, it would < perhaps > be useful to explain briefly what a text is and why we

In my datasets, among the right collocates of *perhaps* we find that the adverb can also be the first of a series of sentence adverbs; in particular, as shown in Fig. 5.50, it can be immediately followed by *more/most* modifying other adverbs or adjectives (*BNC*: 7 occurrences, 3.5%; *CWo*: 4 occurrences, 2%). Other right collocates of the term are conjunctions such as *because* (*BNC*: 5 occurrences, 2.5%; *CWo*: 2 occurrences, 1%); personal pronouns (see the paragraph above) and NPs (*BNC*: 56 occurrences, 28%; *CWo*: 60 occurrences, 30%) indicating the subject of the sentence.

Fig. 5.50 Screenshot of the right collocates of *perhaps*  
in the *CWo* dataset

77	usbooks	air was colored such a sweet blue he could not stop looking. < Perhaps > he had dreamed and perhaps she had really come to him.
99	usbooks	for Ruffles than the ridged chips of Frito-Lay's competitors. < Perhaps > it is the fancy package Frito Lay uses, the 'Ruffles have
114	times	as both teams looked to put width on the ball -- New Zealand < perhaps > more convincingly than Wales. Graham Henry, the former Wales
116	usnews	another drive. Likewise, save your data to a different drive. And, < perhaps > most important, change the swap file for windows to a
156	brbooks	Listen! Even Nature seems suddenly to have been gagged.' < Perhaps > the lady is lurking in the dark, waiting for us.' The water

With reference to the meanings conveyed, *perhaps* frequently expresses uncertainty about (the truthfulness of) something (in such cases the adverb frequently occurs with modal verbs such as the conditional *would* and the verb *to suppose*); when expressing uncertainty, *perhaps* is also used to make hypothesis on the past or

suppositions on the future (*BNC*: 185 occurrences, 92.5 %; *CWo*: 186 occurrences, 93%; e.g. “[...] it would *perhaps* be useful to explain briefly what a text is [...]”, *BNC*: FRL; “Melissa stopped short, uncertain what to do. *Perhaps* he had stepped behind a tree to relieve himself”, *BNC*: GVP; “[...] from a financial aspect I suppose *perhaps* being a little more more callous I mean [...]”, *CWo*: brspok; “I don’t know what to call it, an identity *perhaps*”, *CWo*: usbooks). In addition, within the instances which express uncertainty, there is a subset of sentences in which the subject noun phrase denotes someone who gives suggestions or advice to someone else (*BNC*: 36 occurrences, 18%; *CWo*: 32 occurrences, 16%), for example, in association with the verb *to think* (*BNC*: 5 occurrences, 2.5%; *CWo*: 6 occurrences, 3%; e.g. “*Perhaps* you should drink this”, *BNC*: GVP; “[...] I think, the easier it is *perhaps* to get a job [...]”, *BNC*: KRH; “I think *perhaps* no one has warned you [...]”, *CWo*: usbooks; “It is time that you should marry, *perhaps*”, *CWo*: usbooks).

The adverb *perhaps* appears also to be used in the answer to questions when one interlocutor does not want to say explicitly if he/she agrees or not with the other interlocutor (*BNC/CWo*: 2 occurrences, 1%; e.g. “You’ve found something, haven’t you? She shrugged. *Perhaps*. Give me the file”, *BNC*: ECK; “Is it coincidence that the fields dominated by black Americans - basketball, jazz, running backs in football - all have this improvisational decision-making, with numerous factors being decided in an instant under emotional pressure? *Perhaps* - and perhaps not”, *CWo*: usbooks). Finally, *perhaps* appears to be used when someone is making polite requests (*BNC/CWo*: 6 occurrences, 3%) and also use modal and conditional verbs not to impose him/herself on the requestee (e.g. “*Perhaps* you could ‘enter’ them Kev?”, *BNC*: H9U).

In conclusion, the adverb *perhaps* has its highest frequency scores in fiction. The term can be found at the beginning, at the end or in the middle (after the main verb) of the clause, and can be used both as a clause adverb (*BNC*: 98%; *CWo*: 97.75%) and as an adverb modifying a number (*BNC*: 2%; *CWo*: 2.25%). Finally, the concordances in my datasets show that the adverb is frequently used to express uncertainty on the truthfulness of something.

### 5.5.3 *Possibly*

The adverb *possibly* occurs about 6,974 times in the *BNC* and 24,093 times in the *COCA*, that is about 70 times pmw in the former and 50 times pmw in the latter (see Table 5.41). A big difference is visible in the distribution of *perhaps* across genres in





The immediate right collocates of *possibly* are verbs such as *to have*, *to be*, *to go*, *to know* and others (construction: subject + *possibly* + verb), which, in most of the cases, are instantiated with the epistemic modal verb *can/could* (*BNC*: 66 occurrences, 33%; *CWo*: 59 occurrences, 29.5%) and *may/might* (*BNC*: 15 occurrences, 7.5%; *CWo*: 5 occurrences, 2.5%). In the latter construction, the adverb stresses the epistemic meaning of the adverb (see Fig. 5.52.1 and 5.52.2; see also section 4.2.4.3).

Fig. 5.52.1 Screenshot of the right collocates of *possibly* in the *BNC* dataset:  
verbs

107	KCP	three four five You you couldn't think six of clubs. you could < possibly > go all that way and not pick a hand up that would make you
112	KRP	advertising by advertising space in newspapers and so forth couldn't < possibly > have done, and she'd done it free. And she got it. She
177	AOC	are notes on subjects such as fruit crystallisation and what is < possibly > the best explanation published of how and why jams set.
45	FCJ	flag rights. Consequently, even if a non-member country may < possibly > be entitled not to recognise a flag granted in a manner
110	G3D	family. Guilt that the family member himself or herself may < possibly > have been in part a cause of the problem in the primary

Fig. 5.52.2 Screenshot of the right collocates of *possibly* in the *CWo* dataset:  
verbs

40	safrica	History. Some of the rarer guns in the storage facility will < possibly > be reprieved and salvaged. Nobody is sure where these will
92	brnews	agrees: 'Even if a firm has only one operating system, you can't < possibly > get all the skills you need at one time.' Acceptance of an
167	brnews	foreseeable future. But there was more here than anyone could < possibly > take in at one go - not least, a fascinating tension
120	brspok	Pronunciation right. <tc text="pause"/> know how to spell it they < possibly > know how to use it correctly but they can't pronounce it
153	times	investment in what's happening in the story'. But how can they < possibly > recapture for a live theatre audience the crowd on the

*Possibly* is rarely used to express uncertainty on numbers or values in the meaning of 'it may be' (see Fig. 5.53). In particular, *possibly* precedes numbers or values only in 4 occurrences (2 %) in the *BNC* and in 10 occurrences (5%) in the *CWo*.

Fig. 5.53 Screenshot of *possibly* in the *CWo* dataset:  
numbers and values

88	usbooks	supplied to keep pumping out electricity for three more weeks, < possibly > four. The problem was, it was not on a seaward terminus
158	brregnews	performance right, if we are to win." Solomons was expected to make < possibly > six changes to the side that defeated Connacht at Ravenhill
170	brbooks	King ended badly but in the days of his pomp created thousands, < possibly > tens of thousands of jobs, and was fted by kings and
193	brbooks	have been responsible?" 'I think it's narrowed down to two, < possibly > three or four possibilities. Hietanen, who'd had his rôle
201	oznews	well balanced team. I think we're looking at at least one and < possibly > two wins, if not three.' 'Talent scouts will be on the

*Possibly* can occur in questions expressing polite requests, for example with *can/could* (*BNC*: 8 occurrences, 4%; *CWo*: 6 occurrences, 3%), and once in the *CWo* with *would you mind* (e.g. "could you *possibly* put flowers on Will's grave sometimes for me?", *BNC*: CA0; "Sweetie, if you're not too busy would you mind *possibly* remembering to bring me the book?" *CWo*: usbooks). The term *possibly* is also used in sentences in which people give explanations on something which they are not sure is true, when they want to express uncertainty (*BNC*: 189 occurrences, 94.5%; *CWo*: 184 occurrences, 92%; e.g. "Short, layered cuts are *possibly* the easiest and most versatile", *BNC*: CDH; "The books could *possibly* arrive in Seattle-area bookstores by June 29", *CWo*: usnews).

To sum up, the adverb *possibly* appears to be mainly used in academic writing in British English and in spoken American English, and can be placed at the beginning or in the middle of the sentence (after the main verb). In my datasets, the term can occur both as a clause adverb (average percentage: 96.5% of the instances) and before a number (average percentage: 3.5% of the instances). *Possibly* is also largely used to express uncertainty on something, that is when people is not totally sure of what they are saying or suggesting and also when a number is considered more or less accurate. Other meanings refer to the adverb used as intensifier of the modal verbs *can/could*, *may/might* and as adverb used to make polite request.

#### **5.5.4 Conclusion**

In conclusion, the analysis of the occurrences of the three adverbs *maybe*, *perhaps* and *possibly* in the corpora shows that they have some differences and also similarities in their frequencies and patterns of occurrence. In particular, *perhaps* has the highest frequency of occurrence both in the *BNC* and in the *COCA*, and share with *maybe* the highest frequency scores in fiction, whereas *possibly* has higher score in the academic writing in the *BNC* and in the spoken language in the *COCA*.

Moreover, as Table 5.42 shows, the three of them can occur at the beginning or in the middle of a sentence, as sentence adverbs and before a number; *possibly* is more likely to be used as an intensifier of the modal verbs *can/could*, *may/might* and to make polite requests. Finally, the three adverbs are used to convey a meaning of uncertainty, especially when people make suppositions/hypothesis on something that could happen and give suggestions/advice.

Table 5.42 Patterns in which *maybe*, *perhaps* and *possibly* are likely to occur

		<i>Maybe</i>	<i>Perhaps</i>	<i>Possibly</i>
<b>Position within the clause</b>	At the beginning	x	x	x
	In the middle	x	x	x
	At the end	x	x	-
<b>Kind of adverbs</b>	Clause adverb	x	x	x
	Adverb before a value	x	x	x
<b>Meanings/Use</b>	To express uncertainty	x	x	x
	To make polite requests	-	x	x
	Intensifier of modal verbs	-	-	x
	To make suggestions/to give advice	x	x	-
	To give options	x	-	-
	To answer to questions	x	x	-

Therefore, as Table 5.42 shows, the three adverbs do not seem to be completely interchangeable. As a matter of fact, while *maybe* and *perhaps* occur at the end of a sentence, *possibly* does not; in addition, although the three of them are used to express uncertainty, only *possibly* and *perhaps* appear to be used to make polite requests and only *possibly* is used as intensifier of modal verbs such as *can/could* and *may/might*, only *maybe* and *perhaps* appear to occur when making suggestions or giving advice, only *maybe* is used to give options and finally, only *maybe* and *possibly* are used to answer to questions.

## 5.6 Conclusion

To sum up, in this chapter I have considered some corpus occurrences of the terms under analysis; in particular, I examined their frequency scores, their collocational and colligational patterns and the meanings they appear to convey in their immediate co-text. The findings suggest that the sets of words considered share some, but not all, phraseological patterns and therefore that they are only partly interchangeable.

In the next chapter we will first comment on the most important findings emerging from my research, then I will point out its limitations and make suggestions on how to further research in this field.

## CHAPTER SIX

### CONCLUSION

#### 6.1 Introduction

In this dissertation I looked at some sets of English near-synonyms (the nouns *assassin*, *murderer* and *killer*; the verbs *to disturb* and *to bother*, the adjectives *compulsory*, *obligatory* and *mandatory*; the adverbs *maybe*, *perhaps* and *possibly*) in order to determine if they are fully interchangeable in all contexts of use or not, and in the latter case, what different patterns of use may be attributed to them.

To achieve this goal, after having considered some studies carried out in the field of synonymy, I started my research project, which comprised two phases:

- in the first phase (see chapter 4), I consulted some monolingual, bilingual and synonym dictionaries, and I took into considerations the etymology, the meanings, the Italian translation equivalents and the synonyms of the terms under study in order to give an overview of their attested similarities and/or differences;
- in the second phase (see chapter 5), I consulted four corpora in order to collect information about the actual use of the above-mentioned terms which I could not easily retrieve from a dictionary survey. In particular, first I used two corpora (the *BNC* and the *COCA*) to identify the frequencies of occurrence and the distribution across text types of the terms under study; then, by examining subsets (i.e. *my datasets*) extracted from other two large general corpora, I retrieved their preferred collocational and colligational patterns.

In this concluding chapter I will summarize the findings relevant to the sets of near-synonyms that I took into consideration, then I will point out the strengths and weaknesses of my research and finally, I will give some suggestions for further research in this field.

#### 6.2 The findings

In the following sub-sections I will give a brief overview on the findings obtained from the consultation of dictionaries and corpora. In particular, I will compare and integrate the findings retrieved from the former with those from the latter in order to be able to outline the main semantic and syntactic features that characterise each term.

### 6.2.1 The nouns

The first noun I considered was *assassin*, which, according to the dictionary definitions, was first used to denote Muslims who used hashish before being sent on suicidal mission by their leaders. In the *BNC* and the *COCA*, the term appears to occur almost 5 times pmw and is mainly used in fiction. Both the dictionary definitions and the corpus data showed that *assassin* is used when referring to a person who is paid and hired to kill someone, the victims mainly being important political and religious figures; in addition, the corpus data also showed that the term occurs when the victims are ordinary people.

My datasets also revealed the following collocational and colligational patterns characterizing the term:

- *assassin* can be preceded by attributive adjectives and past participles which denote personal characteristics or contextual circumstances which are relevant to the assassin being talked about and which, in particular, tend to portray the assassin as a professional;
- in the genitive form (i.e. *assassin's*), the term appears to be followed by head nouns which identify objects commonly found at a crime scene, or alternatively, parts of the body of the assassin him/herself;
- *assassin* often occurs with coordinated noun phrases denoting negative concepts (i.e. illegal acts) or bad people (i.e. other types of criminals); these collocations reinforce the negative meaning already conveyed by *assassin*;
- when *assassin* is a subject noun phrase, the predicate indicates the action of killing or describes the physical movements of the assassin; when *assassin* is a direct object, the predicate indicates actions carried out by police or other authorities against criminals;
- *assassin* also occurs in a prepositional phrase or as a subject complement; in the former case, the noun occurs preceded by the following prepositions: *of*, *for*, *to* *by*, which denote *assassin* as a beneficiary, a recipient or an agent (in passive clauses); and in the latter, it is used to ascribe an identity (i.e. that of an assassin) to the referent of the subject noun phrase.

The second noun I considered was *murderer*, a term which dates back to the 12<sup>th</sup> century. In the first corpora consulted (the *BNC* and the *COCA*), it occurs about 10 times pmw and is instantiated both in fiction and in the spoken language. According to the dictionary definitions, *murderer* is used to denote someone that deliberately kills someone else after having planned the criminal act.

My datasets shows that *murderer* occurs in the following collocational and colligational patterns:

- it often occurs with adjectival pre-modifiers that indicate multiple killing actions (e.g. *mass, multiple, serial*);
- as in the case of *assassin, murderer* too can occur in its genitive form (*murderer's*) followed by nouns which indicate the murderer's belongings or way of behaving;
- the victims of the murderous act are encoded as genitive pre-modifiers of the head noun *murderer* (i.e. *John's murderer*) or as the object of the preposition *of* following the term *murderer* (i.e. *the murderer of John*);
- *murderer* is instantiated in coordination with other terms denoting criminals and terrorists;
- the noun can be the subject of predicates that indicate the action of killing and the direct object of verbs that identify actions carried out by authorities against criminals;
- *murderer* also occurs with the following prepositions: *of, by, to, about, with, on for, from*, which identify *murderer* as the possessor of something, the agent that carries out an action, the source from which someone get something and so on. The term is also used as the subject complement in constructions where people are thus categorized as member of the class 'murderers'.

The third noun I considered was *killer*, which is of Germanic origin. Its frequency of occurrence in the *BNC* and the *COCA* is between 17 and 31 times pmw, and the term is distributed both in the spoken language and in newspapers. The noun appears to have a wider range of meanings than the other two nouns under study, in the sense that it identifies both people that kill and specific kinds of animals or things that kill or destroy (*killer whale, weed killer*). In addition, the dictionaries point out that the term can have a positive, adjectival meaning, namely that of 'formidable' or 'sensational'; however, the corpus data considered do not instantiate it.

Frequent collocational and colligational patterns of *killer* instantiated in my datasets are the following:

- the term can be used with adjectival pre-modifiers that describe the referent of *killer*, that is they refer to a quality or property characterizing them either temporarily or permanently (e.g. *silent, notorious*);

- the term can also be used with nominal pre-modifiers that identify sub-types of human agentive killers, or inanimate entities affected by the action of the killing agent (i.e. patients);
- *killer* can be used as the head noun in a noun phrase whose pre-modifier is a proper name or noun in the genitive form which denotes the victim of the killer;
- when the term occurs in a nominal compound, it is used as a pre-modifier of nouns referring to dangerous animals, diseases or natural disasters, and it qualifies them as ‘something that kills’;
- when used as the subject of a clause, *killer* occurs with predicates that indicate the action of killing or other deliberate actions (e.g. movements) carried out by the referent of *killer*; on the other hand, when used as the direct object, the preceding predicates denote actions carried out by the police and other authorities which affect the referent of *killer*;
- *killer* is also used as the object of the prepositions *of*, *by*, *in*, *to*, *against*, *with*, *on*, *for*, *from*, which indicate the term *killer* as the agent, the beneficiary, the opponent and so on; *killer* can also be used as a subject complement identifying people who behave as killers.

In conclusion, of the three nouns, the most frequent one is *killer*, while the least frequent is *assassin*; *assassin* and *murderer* mainly occur in fiction and *killer* in the spoken language and newspapers. Both the dictionary survey and the corpus data show (a) that the three nouns convey the meaning of ‘people that kill’, (b) that only *killer* can be used when reference is made to aggressive types of animals or diseases or other entities, and (c) that only *assassin* is used when the victim involved is an important political or religious figure. The collocational patterns identified show that only *killer* does not appear to be used in its genitive form (*killer’s*), whereas *assassin’s* precedes nouns that identify objects commonly found at a crime scene or parts of the body, and *murderer’s* precedes nouns identifying properties or ways of behaving. The three terms appear to have the same colligational patterns, that is they occur as subjects of predicates that denote the action of killing someone, and as the object of predicates that indicate actions carried out by some law-enforcing body hunting for criminals. They can be followed by prepositions such as *of*, *for*, *to* and *by*, which identify the noun as the direct object, indirect object, agent and so on. Finally, they are also used as subject complements, that is they identify the referent of the subject noun phrase as a person that kills.



### 6.2.2 The verbs

The first verb I considered was *to disturb*, which dates back to the Middle Ages. In the BNC and the COCA, it occurs between 13 and 20 times pmw and mainly in fiction. According both to the dictionaries and the corpus findings, the verb is used to convey five main meanings: (1) to make something move or change its position; (2) to interrupt the quiet of a situation; (3) to interrupt someone's activity; (4) to upset or to make someone feel worried; and (5) to frighten animals.

In addition, the analysis of my datasets showed that *to disturb* occurs in the following collocational and colligational patterns:

- it can be pre-modified by adverbs (adverbs of intensity, of frequency and adverbs that identify a connection with the realm of emotions);
- the verb is often used both in the active and the passive voice with subject noun phrases denoting noisy events, physical conditions (mental insanity), human agents and inanimate entities (both concrete and abstract) presented as the agents causing disturbance in active sentences or as the patients affected by the action of disturbing in passive sentences;
- in the active voice, *to disturb* is followed by direct objects denoting the patients affected by the action of disturbing, and these can be realized as personal pronouns, common nouns denoting people, nouns denoting animals or nouns denoting calmness and tranquillity;
- the passive voice of the verb can be followed by prepositions such as *by*, *about* and *to*, which are followed by nouns/verbs denoting entities or actions as the cause of disturbance;
- the verb can also occur in the phraseological pattern '*sorry to disturb*', which is an apologetic formula.

The second verb I analysed was *to bother*, whose origin is unclear, although it is said to be of Anglo-Irish origin. The term occurs almost 40 times pmw in the BNC and the COCA and mainly in the spoken language or in fiction. According to both the findings from the dictionary survey and the corpus analysis, the verb appears to mainly have the following meanings: (1) to annoy and interrupt someone; (2) to upset or to make someone feel worried; (3) to stalk someone, to harass or persecute someone with unwanted and obsessive attention; (4) to cause physical pain; (5) not to make the effort to do something; (6) to take the trouble to do something. In addition, the dictionaries point out that the term in its base form (i.e *bother*)

can be used in exclamations or imprecations; however, no instances of this usage were found in my datasets.

With reference to the collocational and colligational patterns of the verb, the following findings emerged:

- *to bother* is frequently preceded by adverbs of manner, of intensity, of frequency or adverbs that identify the circumstances in which the action of bothering is carried out;
- it is also used both in the active and the passive voice as a predicate of subject noun phrases denoting concrete or abstract entities, parts of the body, pronouns, human and animal agents, that is, a variety of entities and participants that can be responsible for the act of bothering, when the verb is used in the active voice, or that can be affected by the action of bothering, when the verb is used in the passive voice;
- the direct objects of *to bother* can be personal pronouns or nouns that denote sentient beings or experiencers who can be affected by the action of bothering;
- the verb (both active and passive) can be followed by prepositions such as *about*, *by*, *with*, *to*, which are followed by nouns or verbs that indicate the topic or event, the issue, the agent or the action that causes annoyance;
- finally, as in the case of *to disturb*, *to bother* too can be used in the phraseological pattern '*sorry to bother*' as an idiomatic apologetic formula.

In conclusion, the verb that occurs more frequently is *to bother*, which is frequently instantiated in fiction (as *to disturb* is) and also in the spoken language. From the point of view of semantics, both verbs indicate the act of interrupting someone's activity or that of making someone feel worried or upset, but only *to disturb* can be used to signal a change in the position or shape of something, the frightening of animals and mental derangement, while only *to bother* can be used to indicate deliberate annoyance and to refer to annoyance due to physical pain or to the action of stalking. With regard to their collocational and colligational patterns, both verbs are preceded by adverbs of intensity and of frequency and can be used with pronouns and subject noun phrases that indicate animate and inanimate entities, both abstract and concrete; the direct objects of both verbs can be personal pronouns, but only *to disturb* is used with animals, whereas *to bother* cannot be used when the direct object are inanimate entities. The two verbs can be followed by prepositions and in particular, both occur with *about*, *by* and *to* followed by nouns or verbs that indicate events, agents or activities that cause annoyance, and only *to bother* occurs with the preposition *with*, which in

turn is followed by a noun denoting issue that causes disturbance. Finally, both verbs are used as part of common apologetic formulas in association with *sorry*.

### 6.2.3 The adjectives

The first adjective I took into consideration was *compulsory*, which derives from Medieval Latin. The term occurs almost 17 times pmw in the *BNC* and only twice pmw in the *COCA* and in both cases, mainly in academic writing. My datasets confirmed the information provided in the dictionaries, that is, the term is used to indicate an obligation meant for others' good, such as keeping people safe or improving their education.

According to my datasets, the adjective occurs in the following collocational and colligational patterns:

- it can be used both as an attributive adjective and as a predicative adjective: in the former case, it is used as a modifier of nouns denoting entities or aspects relevant to the fields of education, administration, law and bureaucracy; in the latter case, it qualifies nouns referring to education, policies and official institutions. The term is also used when reference is made to employment issues or the military service;
- *compulsory* is often followed by prepositions such as *for* + a noun denoting people who are subjects to the obligation, *in* + placename or year identifying the circumstances of the obligation, or the *to*-infinitive form indicating the action that must be carried out.

The second adjective I took into consideration was *obligatory*, which is attested in post-classical Latin, Anglo-Norman and Middle French. In the *BNC* and the *COCA*, the term occurs about 2-3 times pmw and mostly in academic writing. According to the data retrieved in the dictionary survey and confirmed through the corpus analysis, the term appears to be used to indicate 'something that must be done to obey a law or rule'. In addition, the data also revealed that *obligatory* is used to represent something that is considered morally, rather than legally, necessary; that is *obligatory* indicates something which is expected to be done or the proper thing to do rather than a binding obligation; the term is also used to identify something that is typical or traditional (e.g. "For centuries, it was the second major oasis outside China, the starting point for the grassland Silk Road, and an *obligatory* stop for travelers", CWO: brbooks).

The collocational and colligational patterns relevant to *obligatory* that are instantiated in my datasets include the following:

- *obligatory* can be used both as an attributive adjective and as a predicative adjective: in the former case, it precedes nouns denoting requirements, or obvious or unavoidable things; in the latter case, the adjective modifies nouns relevant to the semantic fields of education, grammatical rules, protective clothing, law and bureaucracy;
- the term appears to occur followed by prepositions such as *for* + a noun denoting the people who are subjects to the obligation, *in* + a placename indicating the place in which the obligation takes place, or a *to*-infinitive form signalling the action that has to be carried out.

The third adjective I considered was *mandatory*, which derives from the post-classical Latin. The term occurs about 10-11 times pmw in the *BNC* and the *COCA* and mainly in academic writing. The meaning that the term appears to convey is ‘to obey a command, a commission, a rule or a law’; other meanings given in the dictionary definitions but not attested in the corpus data refer to the use of the adjective before a State name to indicate State mandates and territories that are subjects to rule by mandate.

My datasets reveals the following collocational and colligational patterns:

- the term can occur in informatics in the standard phrases *Mandatory Input/Mandatory Input Field/Mandatory Display Field*;
- *mandatory* can both occur as an attributive adjective and a predicative adjective: in the former case, it is followed by nouns referring to the legal field; in the latter case, it qualifies subject noun phrases denoting policies, injunctions, equipment items, duties and legal consequences;
- the adjective can be followed by prepositions such as *in* + a placename or a year signalling the circumstances of the obligation, and *for* + a noun denoting people who have to obey the obligation.

In conclusion, the adjective that is most frequently instantiated in the corpora is *compulsory* and the one that is least frequently instantiated is *obligatory*; the three adjectives are used in academic writing. The three adjectives are used to indicate ‘something that must be done because it is imposed by an authority or law’ and are instantiated both as attributive adjectives and predicative adjectives. Nonetheless, they appear to be used in different semantic fields, that is only *compulsory* and *obligatory* occur when the relevant discourse topic deals with education and bureaucracy, *compulsory* and *mandatory* with policies,

*obligatory* and *mandatory* with protective clothing and equipment, *compulsory* with the military and the workplace, *obligatory* with grammar, and *mandatory* with legal obligations and in informatics. In addition, the three adjectives can be followed by the preposition *for* + a noun indicating the agents involved in the obligation, or by the preposition *in* + year or the name of the place in which the obligation takes place (*in* + year is not attested for *obligatory*), but only *compulsory* and *obligatory* appear to be followed by the *to*- infinitive with a verb denoting the action that must be carried out.

#### 6.2.4 The adverbs

The first adverb I considered was *maybe*, apparently already in use in the 17<sup>th</sup> century. In the *BNC* and the *COCA*, the term occurs between 102 and 271 times pmw and mostly in fiction. According to the dictionary definitions, the meanings that *maybe* appears to convey are the following ones: (1) to express uncertainty on something that may happen; (2) to make suggestions or give advice without being certain; (3) to indicate the meaning of ‘sometimes’; (4) to guess a number; and in American English (5) before a negative, as an emphatic assertion of the corresponding positive statement or (6) after a statement to indicate that the facts are indisputable, or that the conditions are not negotiable. The findings retrieved from the corpus analysis appear to confirm most of the definitions previously mentioned; in particular, there are many instances which exemplify the meanings of the definitions number (1), (2) and (4). On the one hand, my datasets do not seem to show any exemplification on the use of *maybe* in American English or with the meaning of ‘sometimes’; on the other hand, the instances from my datasets show other two meanings in addition to those given in the dictionaries, that is (1) to give options and (2) to answer to questions without agreeing or disagreeing with it. According to the dictionary definitions and the corpus findings, the term can occur both as a sentence adverb and before a number.

The second adverb I considered was *perhaps*, which derives from Anglo Norman and in the *BNC* and the *COCA*, it occurs between 210 and 345 times pmw, with higher frequency scores in fiction. The term is listed in the dictionary definitions with the following meanings: (1) to express uncertainty on the truthfulness of something; (2) to politely make opinions, requests or give answers without being certain on what is said; (3) to give suggestions or advice; (4) to guess a number. In my datasets, the term appears to be mainly used to convey uncertainty (see meaning (1)), to make polite requests or give answers without showing 100% certainty (see meaning (2)) and to give advice or suggestions (see meaning (3)). In addition, my datasets show that *perhaps* is also instantiated to make hypotheses on a past event or

suppositions on a future event, and that it is both used as sentence adverb and before a number. The corpus findings also show that the term is mainly used at the beginning of a sentence or clause but that it can also occur within the sentence after the main verb.

The third adverb I considered was *possibly*, which derives from Latin. The term is instantiated in the BNC and the COCA between 51 and 72 times pmw with higher frequency scores both in academic writing and spoken language. According to the dictionary definitions, this adverb is likely to occur (1) to convey a meaning of uncertainty on something that may happen; (2) to emphasize that someone has tried their best to do something; (3) to express astonishment, shock or upset on something; (3) with modal verbs to make polite requests; (5) to guess a number. My datasets give many instances of the term used as an intensifier of the modal verbs *can/could*, *may/might* (as reported also in the dictionary definitions). Moreover, in addition to the definitions given in the dictionaries, the corpus findings show that *possibly* appears to be used ‘to give explanations or make suppositions on something that is uncertain or not completely true’.

Even in this case, we cannot consider the three adverbs as completely interchangeable. As a matter of fact, they have only a few features in common: they are used as sentence adverbs and before a number (although *possibly* with lower frequency scores than the other two adverbs), and all express uncertainty. On the other hand, *maybe* and *perhaps* appear to be used to make suggestions or give advice, whereas *perhaps* and *possibly* can be used to make polite requests. Finally, only *maybe* is used to give options, while only *possibly* is used as intensifier of the modal verbs *can/could*, *may/might*.

### **6.3 Strengths and weaknesses of the research**

This research was carried out with a view to being of some help to the learners of foreign languages and to teachers who may need to explain to their students how to use the sets of near-synonyms analysed. As a student of English as foreign language, I believe that this work may be useful to understand the contexts in which the terms under study can be and should be used. In particular, this research was based on the analysis of some near-synonyms taken as case studies; this gave me an opportunity to explore several aspects of the terms under study. In particular, by consulting etymological dictionaries, I learnt about the origin and the semantic evolutions of the terms; then, by consulting and comparing other lexicographic sources, I was able to collect and systematize the information available on those terms; finally, by consulting corpora, I was able both to check the validity of the data found in dictionaries and to retrieve additional information about the terms. In particular, the corpus

findings enabled me to identify some of the semantic, syntactic and textual specificities of the terms and highlight some of their similarities and differences, which are displayed in the comparative tables at the end of every major section of chapters 4 and 5.

Now, looking at the findings of my research, I may say that, on one hand, I am more confident on the use of this terms; for example, I know that in newspapers I will easily find the term *killer* rather than *assassin* or *murderer*, and that when reference is made to important political figures as victims of an act of killing, the default term to refer to the criminal is *assassin*. I also know that I can use the term *to bother* when I want to convey the notion of intentionally annoying someone, and that it is better to use *to disturb* when referring to unintentional annoyance. I also know that in informatics I will find the term *mandatory*, rather than *compulsory* or *obligatory* and that when I am guessing a number it is better to use *maybe* or *perhaps*, rather than *possibly*.

On the other hand, I still have some uncertainty as to the choice of one term vs the others. More specifically, when the terms seem to be used indifferently to convey a given specific meaning, I am not completely sure that they are fully interchangeable in every their nuance or if, on the contrary, there are some other characteristics that make the choice of one term more advisable than the other one, characteristics which, however, I was unable to detect through my analysis (e.g., both *bother* and *disturb* occur in the phraseological pattern ‘*sorry to*’; both *compulsory* and *obligatory* occur when talking about education, and both *maybe* and *perhaps* are used to make suggestions and give advice). Moreover, when one term is preferred (i.e. *assassin* with political figures as victims), does it mean that a near-synonym in that context would sound wrong in any case?

My research method suffered from some limitations, which include the following:

- 1) I mainly considered the immediate collocates of the terms under study and only in few cases did I also look at the larger context in which they occur; as a result, I may have missed out some phraseological patterns characterizing the terms under study, but visible only in a larger context, such as a whole sentence.
- 2) Additional near-synonyms could have been analysed in order to better understand the semantic and grammatical space occupied by a given term in opposition to its near-synonyms (additional terms that could have been analysed include the noun *slayer*, the verb *to annoy*, the adjective *necessary* and the adverb *potentially*).
- 3) At the beginning of my research I thought that my decision to first look at the dictionary definitions and then to consult corpora was appropriate because this way I could become aware of what was already known about the terms under study

before setting about to discover something new, if any, about them. (Also, according to the literature, this was a standard procedure; see chapter 2.) But in my case, I think that this method did not work very well; as a matter of fact, in the second phase of the research, I was aware of the meanings previously retrieved from the dictionary definitions, and as a result, when I looked at the instances in my datasets, I first instinctively tried to match the meanings previously retrieved with the terms instantiated. It would probably have been better if I had first consulted the corpora and then the dictionaries.

#### **6.4 Recommendation for further work**

This research has also highlighted some topics on which further research would be needed: for example, the literature review showed that there are not many studies on the meaning, the actual use and the contextualization of specific terms; therefore, it could be interesting to select terms in common use and to conduct research on their meanings together with their use in the real world.

Moreover, in this research the use of more data sources and bigger datasets could have been helpful to identify other meanings or patterns which were not given in the dictionaries and corpora consulted (for instance, when dealing with the term *killer*, the corpus findings did not show any positive connotation of the term, unlike what the dictionary definitions highlighted); I could therefore suggest to people who want to develop a similar research that with the use of a wider range of corpora and/or instances, they are more likely to retrieve all the kind of information that are useful to detect what is typical and what is untypical of the terms under study.

In addition, it could also be interesting to look at possible semantic gaps in the term sets I considered; these gaps refer to words that does not exist in a language but which, given the structures and the grammatical rules of that language, could hypothetically exist. The non-existence of such words to represent a specific meaning could explain why certain terms are preferred rather than others. For example, although the verb *to manslaughter* indicates “the crime of killing a human without malice aforethought, or in circumstances not amounting to murder” (*Oxford Dict.* definition), the noun *\*manslaughterer*, which should indicate the person/criminal that ‘kills without intending to’, does not exist.

Finally, it could also be interesting if lexicographers include some basic data referring to the frequency of occurrence of terms, their typical registers and preferred contexts of use besides providing definitions on terms. This would give a more in-depth outline of the



semantic profile of terms. The inclusion of such data in dictionary definitions could be easily achieved thanks to the widespread availability of on-line dictionaries, which can be continually updated and enriched and, unlike printed dictionaries, have not space constraints.

## **6.5 Final remarks**

The findings of this research could be used in school by teachers of English as foreign language to explain to their students some rules to be used in order to determine when and why to use one or the other near-synonyms. For this purpose, I think that the tables at the end of every major section of chapters 4 and 5 are an useful way to exemplify to the students the main characteristics that portray each near-synonym in the sets. In addition, teachers could also use the findings and the examples of this work to create some testing activities such as cloze tests<sup>30</sup>, in which the terms under study could be removed and students would put the right word in the right place. The use of cloze tests could be both the start for the explanations or the final proof to determine if students have understood the main differences among the terms.

This research was conducted to provide a contribution to the field of English lexical semantics, by focusing in particular on some sets of near-synonyms which are likely to cause problems to foreign English learners. This research has revealed that the near-synonyms taken into considerations do not represent cases of absolute synonymy, which, according to scholars, is really rare, and therefore that they cannot be considered completely replaceable one with another in all contexts in which they tend to occur – each term seems to be characterized by some features that are not shared with the other near-synonyms in the same set.

---

<sup>30</sup> A cloze test is a text from which words have been removed and replaced by blank spaces. The person taking the test tries to 'close' the text by putting a correct word into each of the blank spaces.



## References

- Atkins B. T. S. & Levin B. (1995) "Building on a corpus: a linguistic and lexicographical look at some near-synonyms", *International Journal of Lexicography*, 8(2): 85-114.
- Brodda B. & Karlgren H. (1969) "Synonyms and synonyms of synonyms", *Statistical Methods in Linguistics*, 5: 3-17.
- Cambridge Dictionaries Online, <http://dictionary.cambridge.org/> (last accessed 15 May 2013)
- Cappuzzo B. (2010) "Patterns of use of English synonymous medical terms: the case of *disease* and *illness*. A corpus- based study", *Rivista della Facoltà di Scienze Motorie dell'Università degli Studi di Palermo*, 3(1): 19-40.
- Church K.W., Gale W., Hanks P., Hindle D. & Moon R. (1994) "Lexical substitutability", in Atkins B. T. S. (ed.) *Computational Approaches to the Lexicon*, 16: 1-35.
- Clift R. (2003) "Synonyms in action: a case study", *International Journal of English Studies (IJES)*, 3(1): 167-187.
- Cobuild (2003) *Collins Cobuild Advanced Learner's English Dictionary*, Glasgow: Harper Collins Publishers.
- Collins English Dictionary*, <http://www.collinsdictionary.com/> (last accessed 22 July 2013)
- Cruse D. A. (1986) *Lexical semantics*, Cambridge: Cambridge University Press.
- De Jonge B. (1993) "The existence of synonyms in a language: two forms but one, or rather two, meanings?", *Linguistics. An Interdisciplinary Journal of the Language Sciences*, 31(3): 521-538.
- Divjak D. (2010) *Structuring the lexicon: a clustered model for near-synonymy*, Berlin: Mouton de Gruyter.
- Duffley P. J. & Joubert J.-F. (1999) "The gerund and the infinitive with the verbs *intend*, *mean*, *purpose* and their close synonyms", *Canadian Journal of Linguistics*, 44(3): 251-266.
- Edmonds P. G. (2000), "Semantic representations of near-synonyms for automatic lexical choice", *Dissertation Abstracts International, Section B: The Sciences and Engineering*, 61:1.
- Edmonds P. G. & Hirst G. (2002) "Near-synonymy and lexical choice", *Computational Linguistics*, 28(2): 105-144.
- Filippov E. N. (1971) "Coexistence and rivalry of American and British synonyms in Modern British English", *Inostrannye Jazyki v Škole*, 4: 113-117.

- Gesuato S. (2007) “How (dis)similar? Telling the difference between near-synonyms in a foreign language”, in Hidalgo E., Quereda L. & Santana J. (eds.) *Corpora in the Foreign Language Classroom, Selected Papers from the Sixth International Conference on Teaching and Language Corpora (TaLC6), University of Granada, Spain, 4-7 July, 2004*, Language and Computers Series, Amsterdam/New York: Rodopi, 175-190.
- Gottschalk K.D. (1992) “What about How about? Or, The non-synonymy of How about and What about”, in Tracy R. (ed.) *Who climbs the grammar tree*, Tübingen, Niemeyer, 237-255.
- Hayakawa S. & Fletcher P. J. (1971) *Cassell’s modern guide to synonyms and related words*, London: Cassell.
- Hazon M. (2010) *Garzanti Hazon: inglese-italiano, italiano-inglese*, Milano: Garzanti Linguistica.
- Hoey M. *What’s in a word?* [http://www.onestopenglish.com/support/methodology/grammar-vocabulary-and-skills/whats-in-a-word/155130.article#ID\\_5](http://www.onestopenglish.com/support/methodology/grammar-vocabulary-and-skills/whats-in-a-word/155130.article#ID_5) (last accessed 12 June 2013)
- Hudson R., Rosta A., Holmes J. & Gisborne N. (1997) “Synonyms and syntax”, *Journal of Linguistics*, 33(1): 227-234.
- Inkpen D. Z. & Hirst G. (2004) “Building a lexical knowledge-base of near-synonym differences”, *Dissertation Abstracts International, Section B: The Sciences and Engineering*, 65(5).
- Kalisz R. (1988) “Prototypes and invariants in linguistic categorization”, *Studia Anglica Posnaniensia: International Review of English Studies*, <http://www.thefreelibrary.com/Prototypes+and+invariants+in+linguistic+categorisation.-a093027785> (last accessed 16 March 2012).
- Kearns K. (2000) *Semantics*, London: MacMillan.
- Kent R. K. (1984) *Webster’s New World dictionary of synonyms*, New York: Simon and Schuster.
- Lyons J. (1995) *Linguistic semantics: an introduction*, Cambridge: Cambridge University Press.
- Longman Dictionary of Contemporary English*, <http://www.ldoceonline.com/> (last accessed 22 July 2013)
- Longman (1993) *Language activator: the world’s first production dictionary*, Harlow: Longman.
- Macmillan Dictionary*, <http://www.macmillandictionary.com/> (last accessed 22 July 2013)

- Merriam Webster Dictionary*, <http://www.merriam-webster.com/> (last accessed 22 July 2013)
- Moss H. & Motta V. (2000) *Using Italian synonyms*, Cambridge: Cambridge University Press.
- Niepokuj M. (1997) "Differentiating synonyms: some Indo-European verbs of cutting", in Adams D. Q. (ed.) *Festschrift for Eric P. Hamp*, Washington DC: Institute for the Study of Man, 97-109.
- OED*, <http://www.oed.com/> (last accessed 04 June 2013)
- Oxford Dictionaries* (British and World English), <https://oxforddictionaries.com/> (last accessed 22 July 2013)
- Palmer F. R. (1981) *Semantics*, Cambridge: Cambridge University Press.
- Picchi F. (2010) *Dizionario Hoepli inglese: inglese-italiano, italiano-inglese*, Milano: Hoepli.
- Ragazzini G. (2013) *Dizionario Zanichelli: Il Ragazzini*, <http://dizionarionline.zanichelli.it/dizionariOnline/#ragazzini> (last accessed 19 July 2013)
- Randolf Q. & Douglas B. (1999) *Longman-grammar of spoken and written English*, London: Longman.
- Saeed A. T. & Fareh S. (2006) "Some contextual considerations in the use of synonymous verbs: the case of *steal*, *rob*, and *burglarize*", *Studia Anglica Posnaniensia: An International Review of English Studies*, 42: 323-336.
- Saeed J. (1997) *Semantics*, Oxford: Blackwell.
- Sinclair J.M. (1991) *Corpus, concordance, collocation*, Oxford: Oxford University Press.
- Soul R. (1938) *A dictionary of English synonyms and synonymous expressions*, London: Frederick Warne & Co.
- Strang B. M. H. (1971) *Modern English structure*, London: Edward Arnold (2 ed.).
- Stubbs, M. (2001) *Words and phrases. Corpus studies of lexical semantics*, Oxford: Blackwell.
- Suarez J. A. (1971) "A case of absolute synonyms", *International Journal of American Linguistics*, 37(7): 192-195.
- Sullivan W. J. (1986) "Six troublesome synonyms", *The thirteenth lacus forum*, 13: 302-314.
- Von Schneidemesser L. (1980) "Miscellany: *purse* and its synonyms", *American speech*, 55(1): 74-76.
- Taylor J. R. (2003) "Near synonyms as co-extensive categories: 'high' and 'tall' revisited", *Language Sciences*, 25(3): 263-284.

- Tsuji K. & Kageura K. (1998/1999) "An analysis of medical synonyms – The word-structure of preferred terms", *Terminology*, 5(2): 229-249.
- Watkins C. (2000), *The American heritage dictionary of Indo-European roots*, Boston: Houghton Mifflin.
- Wierzbicka A. (1996) *Semantics, primes and universals*, Oxford: Oxford University Press.
- Wilding J & Mohindra N. (1983) "Preferred synonyms of each noun of 279 synonym pairs", *British Journal of Psychology*, 74: 91-106.