The complementation patterns of the adjective *certain*

in two corpora of British English

University of Tampere

School of Language, Translation and Literary Studies

English Philology

Pro Gradu Thesis

Satu Männistö-Luomanen

Spring 2014

_____

Tämä pro gradu -tutkielma tarkastelee englannin kielen adjektiivin *certain* komplementaatiota brittienglannissa 1700-luvulla sekä nykypäivänä. Tarkoituksena on tutkia onko tämän adjektiivin komplementatiossa tapahtunut muutoksia ajan myötä sekä selvittää onko kielen rakenteilla yhteyttä adjektiivin eri merkityksiin.

Aineisto on kerätty kahdesta eri korpuksesta. Lähteenä historialliselle aineistolle on The Corpus of Late Modern English Texts (CLMET) ensimmäinen osa, joka koostuu pääosin kaunokirjallisista teksteistä aikavälillä 1710-1780. Nykyenglannin lähteenä toimii The British National Corpus (BNC), joka kattaa aikavälin 1960-1993. Haku tästä korpuksesta rajattiin kaunokirjallisuuteen, jotta aineistot olisivat keskenään vertailukelpoiset.

Tutkielman alkuosassa käsitellään korpustutkimusta sekä selvitetään millaista tietoa adjektiivista on ennestään saatavilla sanakirjojen ja kielioppiteosten avulla. Lisäksi tarkastellaan komplementaatiota koskevia teorioita.

Teoriaosuuden jälkeen korpusaineistoa analysoidaan alkaen historiallisesta aineistosta. Kutakin komplementtityyppiä tarkastellaan erikseen ja adjektiivin merkitys kussakin kontekstissa määritellään. Aineistosta selviää, että *that*-lauseke on adjektiivin *certain* yleisin komplementti, mutta sen käyttö on vähentynyt nykypäivään tultaessa. Lisäksi selviää, että adjektiivin tietyt merkitykset ovat kytköksissä eri komplementteihin ja että adjektiivin käyttö eri lauserakenteissa vaikuttaa komplementaation valintaan.

Asiasanat: certain, komplementaatio, korpus, adjektiivi

TABLE OF CONTENTS

1 Introduction

Certain is a very common adjective and it is used in various ways and meanings in the English language. The main interest of this thesis is in the different complementation patterns that the adjective *certain* can take. Consider the following examples taken from the Corpus of Late Modern English Texts and the British National Corpus:

> (1) As to the third sense of the word, it is ***certain***, <u>that</u> both vice and virtue are equally artificial, and out of nature. (Hume 1739-40, *Treatise of Human Nature*)
>
> (2) He was ***certain*** <u>to ask</u> her to dance when she looked like this! (BMW 1718)

The underlined parts in the above sentences represent some of the complement patterns that occur with *certain*. It has to be noted, however, that not all linguists agree on the nature of the *that*-clause illustrated in (1). In fact, according to Quirk et al. (1985, 1223) and Huddleston and Pullum (2002, 964) the construction in question is an extraposed subject instead of a complement. The point of view adopted in this thesis, however, is the one of Herbst et al. (2004, xvii), who claim that a clause of the type of (1) is a complement.

In addition to examining whether patterns like these have changed over time I aim to discover whether there can be found any correspondence between the different complements and meanings of the adjective *certain*. After all, according to Bolinger's Principle, "a difference in syntactic form always spells a difference in meaning" (Bolinger 1968, 127).

This thesis focuses on the adjectival *certain* and therefore its other grammatical functions, such as determiner, as in (3), and phrasal constructions, as in (4), will not be studied. Their frequency is calculated, however, in order to portray the relative usage of the adjective *certain*. Here are illustrations of these disregarded constructions taken from the *Collins Dictionary* online:

> (3) Commissioner Byrne said one focus was on the activities of *certain* off-licences. (*Irish Times* 2002)
>
> (4) Vanrin was someone and she knew *for certain* she had heard the name before. (Will Davenport 2003, *The Painter*)

Complementation is an interesting field of research in linguistics as it covers a wide range of syntactic properties in a sentence. Therefore, I find it particularly fruitful for a speaker of English as a foreign language, such as the author of this thesis, to study different complementation patterns as a way of improving their language skills. In fact, Hunston (2002, 174) argues that instead of learning just one word the knowledge of patterns helps a student of language particularly with respect to fluency and accuracy.

By concentrating on the patterns of one predicate, *certain*, the purpose of this study is not only to deepen the knowledge of its usage and meanings in different contexts but to contribute to the literature on the complementation of adjectives in general as it has received less attention compared to the study of verbs in the field.

The data used for analysis is taken from two different corpora. In order to be able to conduct a diachronic study they have been chosen to represent different periods of time in the use of the English language: the Corpus of Late Modern English Texts for historical usage and the British National Corpus for contemporary one.

First, I will talk about using corpora in linguistic research and present the corpus data used in this thesis. Then, the dictionaries will be consulted concerning the etymological origin as well as the meanings and complementation patterns of the adjective *certain*. After that, I will discuss the theoretical aspects, of complementation in general and related to *certain*, found in grammars and related articles on linguistics. Finally, the data will be analysed and the conclusions will be presented.

2   Corpus research

2.1   Corpora in linguistic research

A corpus in general can be described as a large collection of texts. However, a corpus that is suited for linguistic research needs to be further defined. Lindquist (2009, 21) and Tognini-Bonelli (2001, 2) emphasize the fact that a corpus has to be compiled in such a manner that it can be said to represent a given language as a whole to be qualified for linguistic analysis. Hunston (2002, 2) and Biber (2010, 159) point out the authenticity of language usage that can be found in corpora, which makes them suitable for empirical linguistic research. In addition, Svartvik (1992, 7) as well as Hunston and Francis (2000, 15) refer to electronic storage and readability of corpora. In fact, the era of corpora available in electronic form began in the 1960's and nowadays corpus research can be said to be digital by default.

Corpora can be used in at least two different ways for linguistic research projects. Firstly, linguistic descriptions can be made based on the corpus findings, which is called a corpus-driven approach (Biber 2010, 162). Secondly, pre-existing linguistic theories can be tested in the light of corpus evidence, which refers to a corpus-based approach (ibid.). The method of research adopted in this thesis is specifically the corpus-based approach since, according to Biber (2010, 163), it is ideal for empirically identifying the frequencies of particular linguistic patterns.

2.2   Issues in using corpora

There are certain issues of which one needs to be aware when using corpora in linguistic research. Some linguists doubtful of corpus use state, as Mukherjee (2004, 114), that a corpus can never be absolutely representative of a language. On the other hand, Leech (1968, 94) argues that corpus data can be used as empirical evidence of specific language usages, nevertheless. Furthermore, corpora can contain various kinds of mistakes (Lindquist 2009, 10), which seems only natural since they are, after all, man-made. In fact, mistakes or other imperfections should be expected especially in

corpora that contain transcriptions of spoken texts. With respect to these, Leech (1968, 89) mentions false starts and hesitations.

Svartvik (1992, 10) points out, regarding spoken corpus texts in particular, that they risk being used without taking into account the context when the data is gathered and transcribed by someone else than the user of the corpus. Hunston (2002, 23) also brings up the importance of context which is sometimes insufficient in electronic display of corpus data. These days, however, the major online corpora, the ones used in this thesis included, offer the possibility to see more context if wanted.

In addition to human errors, corpora can also contain digital ones caused by the tagging systems of electronic corpora (Biber et al. 1998, 262). These tagging programs categorize every word of the corpus in order to enable electronic retrieval of specific items in it. Sometimes the programs can be misled by the complexities typical of human languages (ibid.). This leads us to the central issues in any electronic corpus research, namely *precision* and *recall* which are used to measure the effectiveness of data search (Ball 1994, 295). In fact, a query from a tagged corpus is likely to deliver some unwanted tokens as well, which decreases the precision of the retrieval. Recall, in turn, evaluates whether all relevant tokens have actually been retrieved or not (ibid.). As Ball (ibid.) points out, the precision can easily be dealt with simply by discarding the unwanted items but it is almost impossible to know whether any relevant data have been missed in the search. While going through a large corpus by hand cannot often be even considered (to improve recall), one should not rely solely on computerized handling of the data either. Svartvik (1992, 10) emphasizes the fact that in many cases manual analysis is indispensable. For this thesis, the corpus search is done electronically but the retrieved data is then carefully examined by hand.

2.3   Corpus data used in this study

This is a diachronic study that compares the historical complements of the adjective *certain* to the ones used today. According to Bas et al. (2011, draft, ch. 54, 1), it is most useful to recur to the Late

Modern English when examining the linguistic constructions of Present-day English, particularly syntactic ones. The historical evidence is therefore taken from the original version of the Corpus of Late Modern English Texts (CLMET) compiled by Hendrik de Smet. It contains texts that are written by native speakers of British English between 1710-1920 comprising approximately 10 million words. (de Smet 2005, 71). For this thesis, I have made use only of the first part of the CLMET which consists of texts from 1710 to 1780 equivalent to slightly over 2 million words. In that part of the CLMET there were 888 instances of *certain* of which 25 per cent were collected for the purposes of this study, which means 222 tokens altogether.

The modern usage of British English is represented by the British National Corpus (BNC) created by the BNC Consortium led by Oxford University Press. According to Hoffmann et al. (2008, 27) the BNC is compiled so that it reflects modern British English in a balanced way. The BNC comprises 100 million words from written (90%) and spoken (10%) texts published between 1960 and 1993. In order to be comparable with the CLMET sample, the BNC data is selected by choosing the text type of Imaginative Prose, which contains 16.5 million words. The quiery from this section of the BNC returned 2579 hits. This amount was thinned into 10 per cent with the method of random selection which yealded 257 tokens in total.

It must be taken into account, however, that the subsections chosen from the two corpora for comparison differ considerably in size. In consequence, it is evident that comparing the raw counts of the frequencies would not give an accurate result. The relative frequencies can be obtained by norming the frequency counts. Biber et al. (1998, 263) explain how this *normalization* is done by dividing the raw frequency by the total number of words in the text (or in the subcorpus in this case). The result is then multiplied with the basis chosen for norming (ibid.). The basis should be selected according to the typical length of the texts in the corpus (ibid., 264). In this study, the counts are therefore normed to a basis of one million words. Consequently, the normed count of *certain* in the CLMET is: (222 instances/500000 words) x 1 million = 444 and in the BNC: (257

instances/1.65 million words) x 1 million = 156. As we can see by looking at the normalized frequencies, the adjective *certain* occurs much more frequently in the historical texts.

3   *Certain* in the dictionaries

3.1   Etymological background

According to the *Oxford Dictionary of English Etymology* (1966, 159), the origin of the adjective *certain* lies in the Latin verb *cernere* (sift, separate, decide) of which the past participle is *certus* (settled, sure). Partridge (1966, 89), on the other hand, claims that *certus* is the result of transposing the letters within *cretus*, the actual past participle of *cernere*. Furthermore, according to Partridge (ibid), the verb *cernere* originally had a very concrete meaning of passing through a sieve from which the more abstract senses of distinguishing and deciding have developed. Consequently, the past participle *certus* became to mean decided, determined, sure (ibid.), which very much correspond to the senses given to *certain* today as can be seen in the following section. The form *certain* has come to the English language through Old French (*The Oxford English Dictionary*).

3.2   Meanings and patterns

This thesis relies chiefly on the *Oxford English Dictionary* (*OED*) in determining the different meanings of the adjective *certain*. However, the *Collins Cobuild English Language Dictionary* (*Cobuild*) is also consulted to supplement the information found in the *OED*. The complement patterns are derived from the dictionary illustrations.

The *OED* lists a total of seven different meanings with subcategories for the adjective *certain*. For the purposes of this study, the obsolete senses have been disregarded since they have mostly occurred before the earliest time span under scrutiny here, i.e. before 1710. The relevant senses can be observed in table 1 below. Some key words in the meanings-column have been underlined to be later used as reference to that particular sense.

Meanings	Illustrations	Patterns
1.a. <u>Determined, fixed</u>, settled; not variable or fluctuating; unfailing.	1. (1845 H. J. Stephen, *New Comm. Laws Eng.* II. 111) Payment of money on a day **certain**. 2. (1866 A. Crump: *Pract. Treat. Banking* vii. 146) Paris is said to give to London the 'uncertain' for the '**certain**' price, when a [varying] number of francs and cents are exchanged for the £ sterling.	attributive use (the adj. is sometimes put after its noun) attributive use
b. <u>Definite, exact</u>, precise. arch.	1. (1736 Bp. J. Butler, *Analogy of Relig.* i. i.18) No way of determining..what is the **certain** Bulk of the living Being each man calls himself. 2. (1788 J. J. Powell, Ess. *Learning Devises* (1827) II. 75) It is of more importance that rules of this description should be **certain**.	attributive use Ø
2.a. Sure, unerring, to be depended upon; wholly trustworthy or <u>reliable</u>.	(1834 M. Somerville, *Connex. Physical Sci.* (1849) xv. 141) A **certain** indication of a coming tempest.	attributive use
b. <u>Sure to</u> come or <u>follow</u>; inevitable.	(1884 A. C. J. Gustafson, *Found. Death* Pref. 6) Truth's laborious but **certain** advance.	attributive use
c. Sure in its operation or effects; 'unfailing; that always produces the <u>expected effect</u>' (Johnson).	1. (1702 R. Mead *Mech. Acct. Poisons* in C. Hitch et al., *Medical Wks.* (1762) 87,) I have often wished that I knew as **certain** a remedy for any other distemper. 2. (1771 'Junius', *Stat Nominis Umbra* (1772) II. lxi. 283) The abuse of a valuable privilege is the **certain** means to lose it.	Ø attributive use
3. <u>Established</u> as a truth or <u>fact</u> to be absolutely received, depended, or relied upon; not to be doubted, disputed, or called in question; indubitable, sure.	1. (1705 S. Clarke, *Being & Attrib. God* i. 19 One of the **certainest** and most evident Truths in the World. 2. (1726 Bp. J. Butler, *15 Serm.* xv. 295 It is **certain that** Effects must have a Cause. 3. (1877 E. R. Conder, *Basis of Faith* iv. 175) It appears to me not only conceivable, but probable, if not **certain**.	use as comparative attribute *that*-clause Ø
4. Of persons: <u>Fully confident</u> upon the ground of knowledge, or other evidence believed to be infallible; sure (= 'subjectively certain').	1. (1796 H. Hunter tr. J. H. B. de Saint-Pierre, *Stud. Nature* (1799) I. 6) We are **certain**, at least, **of** the existence of those beings. 2. (1837 T. Carlyle *French Revol.* II. iv. vi. 241) Besides one is not sure, only morally-**certain**. 3. (1864 Tennyson *Grandmother* xxi, in *Enoch Arden* 124) I am not always **certain if** they be alive or dead.	*of* + NP Ø *wh*-clause
5. By a change of construction, a person or agent is said to be <u>certain to do</u> a thing, when the fact that he will do it is certain.	(1889 *N.E.D.* at *Certain*, *Mod.*) We are **certain to** meet him in the course of our rambles.	*to* − inf.

6.a. Used to define things which the mind definitely individualizes or <u>particular</u>izes from the general mass, but which may be left without further identification in description.	(1710 R. Steele, *Tatler* No. 173. ⁋ 3) There are **certain** faces for **certain** Painters, as well as **certain** Subjects for **certain** Poets.	attributive use
b. Of positive yet restricted quantity, or <u>degree</u>; of some extent at least.	(1860 J. Tyndall, *Glaciers of Alps* i. 123) The ice is disintegrated to a **certain** depth.	attributive use
c. Sometimes <u>euphemistic</u>ally: Which it is not polite or necessary further to define.	(1958 B. Nichols, *Sweet & Twenties* viii.103) 'Syphilis' had always been described as 'a **certain** disease', just as an attempt at rape had been described as 'a **certain** suggestion', and the result of the rape on the lady was described as leaving her in 'a **certain** condition'.	attributive use
d. With a <u>proper name</u>, it implies that the person so indicated is presumed to be unknown except by name. (often conveying a slight shade of disdain)	(1785 W. Cowper, *Let*. 7 Feb. (1981) II. 323) A **certain** Lord Archibald Hamilton has hired the House of Mr. Small..for a Hunting seat.	attributive use

Table 1. Meanings given for the adjective *certain* in the *OED*, obsolete senses excluded.

Cobuild mentions all the same meanings for *certain* as the *OED* except for 1.a (determined, fixed), 1.b (definite, exact), 2.a (unerring, reliable) and more interestingly 6.c (the euphemistical meaning) although the categorisation differs considerably from the one used in the *OED*. For example, *Cobuild* groups the meanings in 2 (sure to follow and expected effect) and 5 (certain to do) under the title of 'sure to happen'. In addition to the complementation patterns given in the *OED*, *Cobuild* offers several different ones in reference to the same meanings.

   In correspondence to the meaning in 2.b (sure to follow) *Cobuild* presents the illustration "It is almost certain that he will be elected…" thus introducing a *that*-clause pattern. For the meaning in 2.c (expected effect) *Cobuild* offers the *of –ing* pattern as in the example sentence "No treatment more certain of arousing American sympathy could have been devised…"

Meaning 4 (fully confident) is defined as 'no doubt in one's mind' in *Cobuild* and exemplified with

"He felt certain that she would disapprove…" Therefore the additional complementation pattern to

express this meaning is a *that*-clause.

In order to get the big picture of all the patterns that *certain* can take, according to the two

dictionaries, a synthesis is illustrated in table 2.

Meanings in the OED	Patterns in the OED	Patterns in Cobuild
1.a. determined, fixed b. definite, exact	attributive attributive, Ø	

2.a. to be depended upon, reliable
   b. sure to follow
   c. expected effect
	attributive
	attributive
	attributive, Ø
		
		attributive, *that*-clause
		of -ing

3. established fact	attributive, *that*-clause, Ø	attributive
4. fully confident	*wh*-clause, *of* + NP, Ø	*that*-clause, *of* + NP, Ø
5. certain to do	*to* –inf.	*to* –inf.

6.a. particular
   b. degree
   c. euphemistic
   d. proper name
	attributive
	attributive
	attributive
	attributive
		attributive
		attributive
		
		attributive

Table 2. Complement patterns and attributive uses of *certain* in the *OED* and in *Cobuild*.

When examining table 2 it is evident that senses 4 (fully confident) and 5 (certain to do) are clearly

different in that they do not exhibit the attributive use at all. Of these two, meaning 5 seems to be

connected to only one complementation pattern, the *to*-infinitive.

Curiously enough, the definition of the first meaning for *certain* in *Cobuild* is "If you are

certain about something, you have no doubt in your mind about it." However, there is no example

sentence to be found with the prepositional complement *about* + NP in either of the dictionaries.

As we can see *certain* is used as an attribute in most of its meanings. In this thesis, however,

the attributive uses will be counted but not studied any further since they do not take part in

complementation nor interfere with it in any way. Consequently, now that we have defined the

different senses and patterns in the two dictionaries, we can set the attributive uses aside. Moreover,

since meaning 6 is clearly connected to attributive use only, it can be ignored from now on as it will

not be relevant to the analysis of complement patterns and senses related to them.

4 Theoretical framework

4.1 Valency theory

One of the most relevant concepts concerning complementation is valency. According to Huddleston and Pullum's (2002, 218-219) interpretation, valency takes into account only the number of the complements a verb can take and no other properties. This is a notion generally known as transitivity and Huddleston and Pullum do make a comparison between transitivity and valency (ibid, 219).

Valency forms an essential part of valency theory that covers a broader range of complement features. Herbst et al. (2004, xxv) describe valency theory as an analysis of the specific complement patterns that a verb, adjective or noun can take and classify complements with respect to their obligatoriness, form and semantics. These aspects will all be clarified in the following sections as, in this thesis, I will apply the principles of valency theory.

First, I will define the concept of complement by comparing it with an adjunct, which explains the obligatoriness property of a complement in general. In fact, the distinction between a complement and an adjunct is at the core of valency theory. Then, I will continue with the argument and thematic structures of a predicate, which, in turn, account for formal and semantic characteristics of complements.

4.1.1 Complements versus adjuncts

According to Huddleston (1984, 178), complements are elements that are needed, in addition to the predicator, to complete the predicate. This implies that a complement is an indispensable part of a predicate. In fact, Huddleston continues that omitting a complement either changes the meaning or makes the sentence ungrammatical. Huang (1997, 75) and Quirk et al. (1985, 65) also refer to the completing of a specific meaning of a predicate in their definition of a complement. Therefore, it is important to recognize a complement and not to confuse it with other elements in a given sentence.

Adjuncts, that Haegeman (1991, 32) calls "optional phrasal constituents" are sometimes difficult to distinguish from complements. Huddleston (1984, 177) states that, unlike a complement, an adjunct can always be omitted and illustrates it with the examples (1) and (2). Haegeman (1991, 32) elaborates that adjuncts often refer to manner, time or place as can be seen in (1).

> (1) Unfortunately, my uncle was using an electric drill at that very moment.
> (2) My uncle was using an electric drill.

Sentence (1) contains adjuncts *unfortunately* (manner) and *at that very moment* (time) whereas sentence (2) is stripped of them, which shows that, from a syntactic point of view, adjuncts are omissible. Consequently, *an electric drill* is a complement of the predicator *was using* since omitting it would leave us with at least a change in the meaning of the verb. Consider sentence (3):

> (3) My uncle was using.

However, the distinction between a complement and adjunct is not always clear-cut. In fact, Herbst et al. (2004, xxviii) note that there are borderline cases and therefore the distinction can be seen as gradient.

4.1.2 Argument structure

According to Huddleston and Pullum (2002, 226), arguments are syntactic elements involved in the action or state expressed by the predicate. They continue that complements are arguments whereas adjuncts are not. In fact, Huang (1997, 76) claims that arguments "include all the complements". However, Huddleston and Pullum (ibid.) observe that they do not always go hand in hand and mention the pleonastic *it* or dummy, as they call it, as an element that is not an argument because it is a semantically empty NP which does not contribute to the meaning of the clause. Consequently, arguments have a semantic function in relation to the predicate. Huang (1997, 76) calls them place-holders of the predicate-argument structure and refers to the number of arguments required by a predicate (ibid, 64). This is comparable to the transitivity feature of verbs.

While transitivity concerns verbs, Haegeman (1991, 40) makes a connection between verbs and adjectives when comparing the argument structure of *know* and *aware* in (4) and (5):

(4) know: verb;   1   2
                              NP   NP
(5) aware: adjective;   1   2
                              NP   PP

The schema above tells us the number of arguments each predicate minimally takes and could be illustrated with the following sentences:

(6) John *knows* the situation.
(7) John is *aware* of the situation.

In (6), the predicate *know* takes the NP *John* as one argument and the NP *the situation* as its second argument. In (7), the adjectival predicate *aware* also takes two arguments: the NP *John* and the PP *of the situation*. Huang (1997, 72), however, has another take on the argument structure of *aware*:

(8a) John is *aware*.   (one argument; NP)
(8b) John is *aware* of Bill's success.   (two arguments; NP, PP)
(8c) John is *aware* that Bill won the prize.   (two arguments; NP, *that*-clause)

According to Huang, *aware* is a one or two-place predicate (ibid). In other words, the number of arguments that *aware* minimally takes is only one (the subject argument) for Huang whereas for Haegeman it is two. This raises the question whether the meaning of *aware* changes depending on the number of arguments (cf. (3) in 4.1.1). *Cobuild* and the *OED* claim that *aware* does not occur without a complement unless preceded by an adverb. *The Longman dictionary of contemporary English*, however, argues that in addition to a PP and a *that*-clause, a zero complement can be used for the same meaning of *aware* and gives the example: As you are *aware*, a fee will be charged annually. In conclusion, *aware* can be used as a one-place predicate without a change in the meaning but not very commonly or in contexts where the whole meaning can be understood even without a complement.

By saying one or two-place predicate, Huang (1997, 72) refers to the optionality of the second argument, which is a characteristics typical of adjectives. In fact, adjectival predicates differ from

verbal ones in that their complements are mostly optional (Huddleston and Pullum 2002, 542) and can often be implicitly understood although omitted (Haegeman 1991, 39). Consider illustrations (9) and (10) (ibid) where the parenthesis refer to the optionality of the prepositional phrase:

(9) envious: adjective;   1   (2)   [John is envious (of her).]
                                NP   PP

(10) restless: adjective;   1   [John is restless.]
                                NP

4.1.3 Thematic structure

In addition to argument structure, all predicates have a thematic stucture as well. This means that a predicate assigns different semantic roles to its arguments (Haegeman 1991, 41). In theta theory these semantic roles are called (thematic or) theta roles (ibid.) Huang (1997, 78-79) emphasizes the fact that these roles are independent from other linguistic classifications such as grammatical functions (subject, object, adjunct) or categories (NP, VP) and illustrates it with the following sentences:

(11a) <u>The boy</u> kicked the ball.   (**Agent**: initiator of the action)
(11b) <u>The boy</u> was scolded by his sister. (**Theme**: undergoer of the action)
(11c) <u>The boy</u> was given a toy.   (**Goal**: toward which the action is directed)
(11d) <u>The boy</u> enjoyed the game.   (**Experiencer**: experiencer of a state)

These examples show that while *the boy* remains the subject in all the sentences the thematic role varies from one sentence to another. The descriptions of the roles are adapted from Haegeman (1991, 41-42). In addition to the ones mentioned above, Huddleston and Pullum (2002, 230) enlist theta roles such as Instrument, Patient, Recipient, Location and Beneficiary. Although the importance of theta theory is generally acknowleged among linguists, there is no consensus over the number or terminology of theta roles (Dowty 1991, 574). In fact, the role called Theme here is termed Dative by Fillmore (1968, 24) or used alternatively with Patient. Furthermore, the role of the one who initiates the action expressed by the predicate can be called Agent or Actor in case it refers to an animate, intentional entity (Haegeman 1991, 41-42) but some grammarians distinguish

between an Agent and a Causer which can be either volitional or inanimate (Huddleston and Pullum 2002, 230).

Despite the fuzziness of the roles, the main principle of theta theory remains that an argument can have only one theta role and vice versa. This principle is called the Theta Criterion (Chomsky 1981, 36):

> Each argument is assigned one and only one theta role.
> Each theta role is assigned to one and only one argument.

4.2 Category selection and *certain*

As established earlier (in 4.1.2), every predicate has its own argument structure stating the number of the arguments. Instead of random adoption the predicate selects its arguments for their syntactic category (NP, PP etc.) and this is called Category selection (or C-selection) (Huang 1997, 68-73). Therefore, adjectives, as well as verbs, can be categorized according to the complementation they take. Huddleston and Pullum (2002, 219) call this phenomenon subcategorisation and, in their terms, a predicate licenses its complements. According to Quirk et al. (1985, 1220), there are six different complementation patterns that adjectives can license: prepositional phrase, *that*-clause, *wh*-clause, *to*-infinitive clause, *than*-clause and *–ing* participle clause. The four first ones, plus the zero complement found in the dictionaries, concern the adjective *certain* and will be discussed more in detail next.

4.2.1 Non-sentential complements

4.2.1.1 Zero complement

Consider the sentence taken from Huddleston and Pullum (2002, 964):

(12) That he is being victimized is ***certain***.

The absence of a complement after the predicate, as in (12) above, is called a zero complement (Ø). Rissanen (1991, 288) claims that the use of zero complements diminished in the 18th century

compared to what it had been before that. Nevertheless, according to Bas et al. (2011, 10), the *that-*clause and the zero complement continue to be the major complementizers in Present-day English.

4.2.1.2 Prepositional phrase

Adjectives do not generally license NPs as their complements without any prepositional connection (Huddleston and Pullum 2002, 659). This is true with *certain* as well. The only prepositional phrase found in both of the dictionaries consulted was *of* + NP (see table 2 in 3.2). In fact, Huddleston and Pullum (ibid.) state that whenever an adjectival predicate is related to a subordinate NP, the preposition is always *of*. On the other hand, they specify that predicates belonging to the semantic domain of trust and confidence can select various prepositions and thereby *certain* also selects *about* as an alternant of the preposition *of* (ibid., 658). Quirk et al. (1985, 1221) only mention *of* as a possible preposition for *certain*.

4.2.2 Sentential complements

4.2.2.1 *That*-clause

A *that*-clause is a sentential complement when it "functions as an argument with respect to a governing element or head" (Fanego 2004, 3) There are, however, two schools of thought concerning the *that*-clause after an adjective. Quirk et al. (1985, 1223) argue that in many cases it is in fact an extraposed subject and so do Huddleston and Pullum (2002, 964) who illustrate this with the following sentences:

> (13) Max is ***certain*** <u>that he is being victimized</u>. (complement)
> (14) It is ***certain*** <u>that he is being victimized</u>. (extraposed subject)
> (15) <u>That he is being victimized</u> is ***certain***. (no extraposition)

Certain has two arguments in (13) whereas in (14) and (15) it has only one argument. Huddleston and Pullum continue that usually adjectives licencing both constructions, a complement and an extraposed subject, prefer one over the other. However, *certain* is the only adjective to occur commonly in both constructions (ibid.).

A different view is represented by Herbst et al. (2004, xvii). They consider extraposed subjects introduced by *that* as complements exemplifying a *that*-clause complement as follows:

(16) It is awkward <u>that a third party is taking a close interest in the reorganisation</u>.

In this thesis, the approach of Herbst et al. is adopted.

On a semantic note, Quirk et al. (1985, 1223) point out that the indicative *that*-clause refers to an established fact. This is in concordance with the *OED* where the meaning 'established fact' (number 3 in table 1) was the only one to have a *that*-clause as a possible complement pattern.

With finite *that*-clauses there is often the choice between an explicit *that* complementizer and its omission. Rohdenburg (1996, 164) mentions a factor that specifically tends to favor the omission of *that*: the subject *it* introducing the subordinate clause. There are, however, also factors that motivate the retention of *that*. Consider the following sentences given by Rohdenburg (1996, 160):

(17) He told me (yesterday) that John had gone away.
(18) He told me (yesterday) John had gone away.

According to Rohdenburg (1996, 161), any material inserted between the matrix predicate and the subordinate clause triggers the need of an explicit *that* as a signal of subordination. The insertion makes the sentence more complex thus requiring more transparency from the other elements of the sentence. This concept is called the Complexity Principle formulated by Rohdenburg (1995, 368):

> The less directly the dependent clause is linked to its superordinate clause, or the more complex the dependent clause turns out to be, the greater is the need to make its sentential status more explicit.

In addition to discontinuous constructions (such as insertions) Rohdenburg refers here to the complexity of the dependent clause, which can be caused simply by its length (1996, 164). Other factors causing cognitive complexity include passive constructions and the length of the subjects and objects (ibid., 149). There are many other cases as well where a choice between two grammatical alternatives can be made. In fact, Rohdenburg (1996, 151) later reformulated the Complexity Principle into more general terms:

> In the case of more or less explicit grammatical options the more explicit one(s) will tend to be favored in cognitively more complex environments.

4.2.2.2 *Wh*-clause

Huddleston and Pullum (2002, 976) classify the adjective *certain* in the semantic group of knowing that licenses subordinate interrrogatives i.e. *wh*-clauses as complements. Furthermore, Quirk et al. (1985, 1225) note that adjectives tend to occur with a *wh*-clause particularly in non-assertive contexts as in (19) and (20):

> (19) Are you ***certain*** how much the machine costs?
> (20) It was not ***certain*** how far the westernization process would go.

4.2.2.3 *To*-infinitive clause

According to Huddleston and Pullum, (2002, 1258) *certain* takes a raised subject instead of an ordinary one. They argue that in the case of raising adjectives the complement is obligatory since omitting the complement either changes the meaning or makes the sentence ungrammatical:

> (21) She is sure to win. (It is certain that she will win.)
> (22) She is sure. (She is not in any doubt.)

Sentences (21) and (22) show that the meaning of the adjective *sure* changes when the complement is omitted. The adjective *sure* in these sentences can be replaced with *certain* as they are semantically and syntactically parallel. In fact, the dictionary findings for *certain* also suggest a similar change in the meaning: *to*-infinitive complement implies that a person is 'certain to do' something while zero complement is used when a person is 'fully confident' (see table 2 in 3.2).

A deeper look at the syntactic properties of raising adjectives reveals that the subject of the infininitive clause is the same as in the main clause (Quirk et al. 1985, 1226):

> (23) Bob is ***certain*** to agree with you.

In (23) the original subject of the lower clause verb (Bob) has been moved to its final position through Raising (Postal 1974, 33). Since the argument moving is always a noun phrase (NP), as in this case Bob, this procedure is also called NP-movement. In general, a predicate takes either a raising or a control construction. However, according to Carnie (2002, 262), some predicates can take both.

Carnie (2002, 263) as well as Davies and Dubinsky (2004, 8) suggest an idiom chunk as a method of distinguishing between raising and control constructions. In this idiom the cat denotes a secret instead of an actual animal and being out of the bag refers to the secret being revealed. The prediction is that when the idiomatic reading is retained, the predicate tested takes a raising construction and vice versa: in case the meaning is literal instead of idiomatic, it involves a control predicate (ibid.):

(24) The cat *is likely* to be out of the bag. (idiomatic meaning → raising)
(25) The cat is *eager* to be out of the bag. (non-idiomatic meaning → control)

When inserting our adjective *certain* in the cat idiom, we get the following result:

(26) The cat is ***certain*** to be out of the bag.

The idiomatic meaning is clearly retained, which further confirms that *certain* is a raising predicate.

Davies and Dubinsky (2004, 7) introduce yet another method to test whether an item is a raising or control predicate. This test, in turn, relies on the fact that the semantically empty expletive *it* cannot function as the subject of a control predicate. The same is true with the existential *there* as illustrated by Davies and Dubinsky (ibid.):

(27) It *seemed* to be raining.
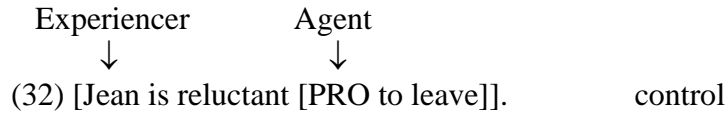(28) There *seems* to be a unicorn in the garden.

While these sentences are acceptable with a raising predicate, they would be ungrammatical with a control predicate. When the predicate *certain* is applied we get the following sentences:

(29) It is ***certain*** to be raining.
(30) There is ***certain*** to be a unicorn in the garden.

The sentences are not ill-formed, which shows that *certain* behaves the same way as *seem* which is a one-place predicate.

The following diagrams by Carnie (2002, 261) summarize the difference between raising and control constructions:

```
        no θ role                Agent
            ↓                      ↓
(31) [ _____ is likely [Jean to leave]].        raising
            ↑_____|
```

```
      Experiencer            Agent
            ↓                  ↓
(32) [Jean is reluctant [PRO to leave]].          control
```
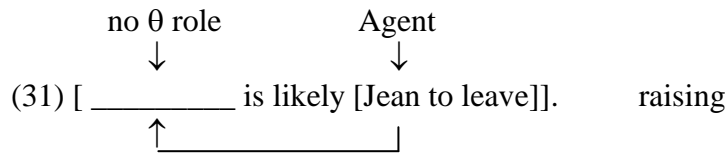
In raising constructions the matrix predicate does not assign an external argument whereas in control constructions it does (ibid.). The external theta role of the lower predicate is assigned to PRO which is a Caseless NP required in order not to violate the Theta Criterion (ibid., 260). PRO occurs in non-finite clauses only (ibid.). Sentence (32) is a case of subject control where the subject NP of the matrix clause controls the PRO or, in other words, is co-referential with it (ibid., 267).

5 Analysis of the corpus data

The data for this thesis is drawn from two different corpora and I will discuss the findings in them in a separate section devoted to each corpus. In each section, we will first look at the numbers concerning all the tokens of *certain* found in that corpus to see the overall usage of the adjective. The instances that are irrelevant to complementation will then be identified and discarded from further analysis. After that the complement patterns found in the data will be analyzed each in their own subsection following the same order as established in the previous chapter on theoretical framework.

5.1 *Certain* in the CLMET

The corpus data from the CLMET yielded expected results with respect to the amount of attributive use of the adjective *certain*. In fact, in almost three quarters (72 %) of the instances *certain* was used attributively. The *that*-clause was the most common complementation pattern (19 %) and the zero complement the second (4,5 %). The frequencies of the complement patterns and other instances found in the data are illustrated in table 3 below.

Pattern in the CLMET	Number of tokens	Frequency as percentage	Frequency normed per million words
Attributive	159	72 %	318
that-clause	43	19 %	86
(of which *that* omitted)	(13)	(6 %)	(26)
Ø	10	4.5 %	20
of + NP	5	2 %	10
of -*ing*	1	0.5 %	2
in + NP	1	0.5 %	2
to -inf.	1	0.5 %	2
Phrasal construction	2	1 %	4
TOTAL	**222**	**100 %**	**444**

Table 3. The frequencies of all the different tokens of *certain* found in the CLMET data.

The phrasal constructions mentioned in the table above are of the type *for certain*:

> (1) This is all I know **for certain**, for I will not send you the thousand lies of every new day (Walpole 1735-1748, *Letters*).

These constructions along with the instances of attributive use are not complements and they will therefore not be discussed any further.

The number of different patterns found in the data was sufficient to give an idea of the overall usage of the adjective *certain* in Late Modern English. However, the distribution of different senses related to the complementation patterns in the corpus data differs from the one in the dictionaries. This is shown in table 4 below:

Meanings in the *OED*	Patterns in the *OED* and *Cobuild*	Patterns in the CLMET
1.a. determined, fixed b. definite, exact	attributive attributive, Ø	attributive attributive

2.a. reliable
  b. sure to follow
  c. expected effect
	attributive
	attributive, *that*-clause
	attributive, *of -ing*, Ø
		attributive
		attributive, Ø
		attributive, *of* + NP

3. established fact	attributive, *that*-clause, Ø	*that*-clause, *of* + NP, Ø
4. fully confident	*that*-clause, *wh*-clause, *of* + NP, Ø	*that*-clause, *of* + NP, *of -ing*, Ø, *in* + NP
5. certain to do	*to* -inf.	*to* -inf.

Table 4. Meanings and patterns in the dictionaries and patterns found in the CLMET data.

The complementation patterns identified in the data will now be analyzed in the following subsections. As there were no *wh*-clauses in the data of this corpus that pattern will not be discussed here. On the other hand, the data findings include one complement pattern that was not mentioned in the previous literature, namely the *in* + NP. This pattern will be looked at in its own subsection under non-sentential complements.

5.1.1 Non-sentential complements

5.1.1.1 Zero complement

The zero complement is the second most frequent pattern in the CLMET data after the *that*-clause.

All ten instances of zero complement found in the data fall under three different senses which are

exemplified here:

> (2) What are Canons, or Bombs, or clashing of Swords? For Death is more ***certain***
>   by Witnesses Words. (Gay 1728, *The Beggar's Opera*) [sure to follow]
> (3) "_Quare non?_" answered Partridge, "it is possible, and it is ***certain***." (Fielding
>   1749, *Tom Jones*) [established fact]
> (4) A thought which I am ***certain*** could never have entered into any mind but that of
>   a lover. (Fielding 1749, *Tom Jones*) [fully confident]

These three meanings were quite equally represented, only the sense 'established fact' had one

instance more than the others. In sentence (2), *certain* is followed by the preposition *by*, which

could indicate a possible case of a complement. However, the meaning of *certain* does not change

when the PP *by Witnesses Words* is omitted. It can therefore be considered as an adjunct, and more

specifically in this case, an adjunct of manner. Sentence (3) represents a typical example of the

meaning 'established fact' in that it has the pleonastic *it* as subject whereas the sense 'fully

confident' inherently requires an animate subject as can be seen in (4).

5.1.1.2 *of* + NP

There were five instances of the complement pattern *of* + NP in the data. Three of them illustrated

the sense 'fully confident' as in (5):

> (5) He then turned to the child, and asked her if she was ***certain*** of that circumstance.
>   (Fielding 1751, *Amelia*) [fully confident]
> (6) … and that whatever we discover to be true of the one species, may be concluded
>   without hesitation to be ***certain*** of the other. (Hume 1739-40, *Treatise of Human
>   Nature*) [established fact]
> (7) …but as no human policy can form measures ***certain*** of success, an irreconcilable
>   hatred was nearly produced… (Johnson 1740-1, *Parliamentary Debates 1*)
>   [expected effect]

The subject argument in (6) is *whatever* and therefore not a sentient Agent, which is typical of the sense 'established fact'. Similarly in (7) the Agent is inanimate *measures* but the meaning of *certain* here is clearly 'expected effect'. Both of these meanings are new to the pattern *of* + NP compared to the dictionary findings.

5.1.1.3 *in* + NP

This PP is a candidate for a new complement pattern for *certain* since it was not mentioned in the literature. There was only one instance of this kind in the corpus data so it is a rare pattern in any case if it is used at all. Since the context becomes of more importance when determining the meaning of the prepositional phrase, the whole sentence, although lengthy, is presented:

> (8) Nor is it only proper we should in general indulge our inclination in the most elaborate philosophical researches, notwithstanding our sceptical principles, but also that we should yield to that propensity, which inclines us to be positive and **certain** in particular points, according to the light, in which we survey them in any particular instant. (Hume 1739-40, *Treatise of Human Nature*) [fully confident]

At first glance, the PP *in particular points* appears to be an adjunct of place just like the PP *in any particular instant* is an adjunct of time. However, the definition depends on the meaning that can be given to *points*. If it refers to mental viewpoints as in 'believe in one's opinions' it would be more diffucult to leave it out without altering the meaning. In that case it could be argued that the PP *in particular points* is in fact a complement to *certain* here.

5.1.2 Sentential complements

5.1.2.1 *That*-clause

The *that*-clause is by far the most frequent complement pattern to occur with *certain* in the corpus data. In fact, one fifth of all instances of *certain* represent this type including the ones where *that* is omitted. Most of these instances (38 out of 43) have the pleonastic *it* as subject like the one illustrated below:

(9) It is ***certain*** that their insolences was the cause of his death. (Gibbon 1776, *The Decline and Fall of the Roman Empire*)

The remaining five instances of a *that*-clause complement have an animate subject as illustrated in (10) and (11):

(10) A man may easily find means of being ***certain*** that he has offended no law in being, … (Johnson 1740-1, *Parliamentary Debates 1*) [fully confident]

(11) I am not, indeed, ***certain*** that those who now call so loudly for information would be prevailed on by any degree of evidence… (ibid) [fully confident]

The argument structure in (10) and (11) is different from the one in (9) in that the sentient subject can now be given the theta role of an Experiencer whereas the semantically empty *it* in (9) could not receive a semantic role. Consequently, the predicate *certain* in (10) and (11) has one more argument: a *that*-clause with the theta role of a Theme.

In both (10) and (11) the *that*-clause is introduced with an explicit *that*, which is in concordance with the Complexity Principle since they both have factors that motivate expliciteness. In (10) there is quite a long distance between the matrix subject *a man* and the predicate *certain,* which can be considered as a complex matrix clause as a whole thus requiring an explicit signal of subordination. In (11) there is a short insertion *indeed* in the superordinate clause but the major complexity factor is probably the longer one, *who now call so loudly for information,* between the lower verb and its subject.

In the three other instances with a *that*-clause as a complement, *that* is omitted:

(12) I am ***certain***, sir, I have not asked you so much out of the way. (Fielding 1749, *Tom Jones*)

(13) If that be the case, as I am ***certain*** it must be, I have fifty pieces at your service. (Fielding 1751, *Amelia*)

(14) I am ***certain*** I can persuade him to his duty. (Goldsmith 1773, *She Stoops to Conquer*)

As we can see, sentences (12) - (14) are quite short and therefore simple compared to (10) and (11). The only discontinuous construction is in (12) where there is a short insertion *sir* between the matrix predicate and the subordinate clause but it does not seem to make the sentence complex enough to require an explicit *that*. Furthermore, the subject in all matrix and lower clauses is a

pronoun, which, compared to multi-word noun phrases, keeps the sentences simple. Moreover, in (13) the subject of the lower clause is *it* which, according to Rohdenburg (1996, 164), often triggers the omission of *that*.

As can be seen by examining all of these five example sentences (10) - (14), the meaning of the predicate *certain* with a *that*-clause complement is clearly 'fully confident'. This concurs with the findings in *Cobuild* whereas the *OED* does not list the *that*-clause as a possible complement pattern with the meaning 'fully confident'.

5.1.2.2 *To*-infinitive

In the corpus data, the complement pattern *to*-infinitive was represented by the following sentence:

> (15) … but fellows who excel in some little, low, contemptible art, are always ***certain*** to despise those who are unacquainted with that art. (Fielding 1749, *Tom Jones*)

This is an NP-movement construction and the syntactic subject of the predicate *certain* is *fellows*. However, the only argument for *certain* in this case is the whole subordinate clause which gets the theta role of a Theme. In fact, the same NP *fellows* has originally been the subject for the verb *despise* in the lower clause and has now been raised to its final position. Before movement, the lower predicate *despise* has assigned the theta role of an Experiencer to the NP *fellows*.

In this sentence the matrix predicate *certain* and the lower one *despise* both express a feeling or a state of mind which are quite stative and can therefore be easily associated with an Experiencer as their subject. In consequence, this sentence does not demonstrate the NP-movement as evidently as in the case where the lower predicate expresses a more active action and would therefore assign the role of an Agent to its subject argument.

The meaning of *certain* here is clearly 'certain to do' as the idea is that the despisement will eventually be done (by the subject *fellows*), which emphasizes more the certainty of the action instead of the confidence of the subject for instance. The latter would, in turn, refer to the sense 'fully confident'.

5.1.2.3 *Of –ing*

There was only one instance of the pattern *of –ing* in the corpus data:

> (16) …to which they retreated with great precipitation; so that the pursuer was now ***certain*** of having housed his pray. (Smollett 1751, *The Adventures of Peregrine Pickle*)

In order to discover whether the construction represented in (16) is raising or control we can apply the cat idiom to it as illustrated in (17):

> (17) The cat is ***certain*** of being out of the bag.

As we can see, (17) cannot be interpreted idiomatically. Therefore, the construction involved here is control.

The thematic stucture is illustrated in (18) where the sentence from the data is inserted in the control construction diagram:

> Experiencer   Agent
> ↓   ↓
> (18) [The pursuer was now ***certain*** [PRO of having housed his pray]].

In (18) *the pursuer* gets the theta role of an Experiencer from the predicate *certain* and there is no movement involved. PRO is in the subject position of the non-finite clause and gets the theta role of an Agent. The NP of the matrix clause, *the pursuer*, controls PRO, which means that they are co-referential. Therefore the construction illustrated in (16) is more specifically subject control.

The meaning of *certain* in this case is not evident. It could be argued that *housing his pray* was an 'expected effect' for *the pursuer*. In fact, the only instance of the pattern *of*-ing in the literature was found in *Cobuild* and the sense given there was 'expected effect' (see 3.2). However, a more plausible interpretation here would be that *the pursuer* was now 'fully confident' of a particular result. *The pursuer* is, after all, the Experiencer of *certain*(ty).

5.2 *Certain* in the BNC

There was a total of 257 instances of the adjective *certain* in the BNC data. Half of them represented attributive use. The pattern that occurred most frequently was the *that*-clause (24.5 %) and the second most common pattern was the zero complement (11.7 %). The frequencies of all patterns found in this corpus data are illustrated in table 5 below.

Pattern in the BNC data	Number of tokens	Frequency as percentage	Frequency normed per million words
attributive	130	50.6 %	79
that-clause	63	24.5 %	38
(of which *that* omitted)	(28)	(10.9 %)	(17)
Ø	30	11.7 %	18
of + NP	9	3.5 %	6
wh-clause	8	3.1 %	5
to –inf.	5	1.9 %	3
of –ing	2	0.8 %	1
Phrasal construction	10	3.9 %	6
Total	**257**	**100 %**	**156**

Table 5. Patterns and their frequencies in the BNC data.

There were two kinds of phrasal constructions in the corpus data including the one in (1) that was also defined as such despite the emphasizing insertion in the middle of it:

> (1) I'm not staying either, that's **for** good and **certain**. (CCM 1960)
> (2) I've spent four years **making** very **certain** I don't care. (H8F 1348)

In addition to *for certain*, which was the most common of the phrasal constructions, there were three instances of *make certain*. These constructions as well as the attributive uses can now be set aside as we move on to the complementation patterns found in the data.

To see whether the senses given to *certain* with different patterns in the dictionaries match with the ones found in the data, table 6 is offered:

Meanings in the *OED*	Patterns in the *OED* and *Cobuild*	Patterns in the BNC
1.a. determined, fixed	attributive	attributive
b. definite, exact	attributive, Ø	attributive
2.a. reliable	attributive	attributive, Ø
b. sure to follow	attributive, *that*-clause	attributive,
c. expected effect	attributive, *of -ing*, Ø	attributive, *that*-clause
3. established fact	attributive, *that*-clause, Ø	*that*-clause, *wh*-clause, Ø
4. fully confident	*that*-clause, *wh*-clause, *of* + NP, Ø	*that*-clause, *wh*-clause, *of* + NP, *of -ing*, Ø
5. certain to do	*to* -inf.	*to* -inf.

Table 6. Meanings and patterns of *certain* in the dictionaries and patterns in the BNC data.

As can be seen from table 6, the patterns found in the corpus data are the same as in the dictionaries but there is variation in relation to the senses they refer to.

In the data of this corpus there were no new patterns. On the other hand, the findings include several *wh*-clauses compared to the CLMET data where there were none of those. They will be discussed under sentential complements. In the following subsections, each complement pattern will be analyzed along the meanings related to it. All the illustrations from the BNC corpus will have a reference code in parenthesis at the end of the sentences as in (1) and (2).

5.2.1 Non-sentential complements

5.2.1.1 Zero complement

Most of the instances of zero complement found in the corpus data have an animate subject. There were 20 sentences of that kind:

> (3) She had been worried for a bit but now she was ***certain***. (CCM 146)
> [fully confident]

In some cases the sentence containing *certain* as a predicate is inserted in the middle of a longer sentence as in (4):

> (4) At one time Carrie thought she heard Anna crying, but she could not be ***certain*** and Melody had insisted that she wanted to be alone with Anna for a while. (HHC 1128) [fully confident]

The meaning of *certain* in all 20 instances of zero complement with an animate subject is 'fully confident'.

There is, however, another substantial group of zero complement with 10 instances containing an inanimate subject as in the following sentences:

> (5) No deal is ***certain*** until the signatures are on the dotted line and the money comes through. (FRS 2750) [established fact]
> (6) "Everything is ***certain***," Hope said. (FPI 2009) [established fact]

In most of the instances the inanimacy was represented by subjects containing *thing*; one thing, nothing or everything as in (6). All of the cases with an inanimate subject fall under the sense 'established fact'.

There is a third meaning that can be related to *certain* with zero complement. It was represented by one instance in the data:

> (7) His world was still ***certain***. (HTX 3846) [reliable]

Instead of 'established fact' for instance, *certain* is more likely to describe *his world* as sure, unerring, not liable to fail; to be depended upon; wholly trustworthy or, in fact, 'reliable'.

5.2.1.2 *of* + **NP**

All instances of *certain* containing an *of* + NP complement have an animate subjet, a person, as in the following examples:

> (8) It is in Amaryllis's best interests and she has told us that she is ***certain*** of happiness with you…', (FP1 1874) [fully confident]
> (9) Of that Theodora was ***certain***. (HA2 1663) [fully confident]

Curiously, the most common NP after the preposition *of* was *that* as in (9). Complemented with *of* + NP *certain* seems to have an animate subject and consequently the sense 'fully confident'.

5.2.2 Sentential complements

5.2.2.1 *That*-clause

The most common pattern with *certain* in the corpus data is the *that*-clause. In the majority of these cases *certain* had an animate subject argument as the Experiencer and the *that*-clause as the Theme. The meaning of *certain* in these cases is 'fully confident'. This is illustrated in (10) – (12). The underlined parts, in turn, represent the factors motivating the use of an explicit *that* to mark the beginning of the lower clause.

> (10) She <u>reluctantly followed Nicole outside</u>, ***certain*** that she was verging on dragging her steps and generally acting like a twelve-year-old delinquent on the way to the principal's office. (H8H 2807) [fully confident]
> (11) She doubled up as matron, treating every illness, injury, headache or toothache with determined suspicion, ***certain*** <u>in her own mind</u> that the pupil was trying to get away with something. (BPD 60) [fully confident]
> (12) In the long months after her return from Paris, she had felt ***certain*** that <u>after the birth (if all went well – and it must, it would)</u> she would be herself again. (FPH 3837) [fully confident]

It can be noted that the Complexity Principle is at play here. In (10), the subject itself is not long but the distance between it and the predicate *certain* is, which creates an environment that can be defined as cognitively complex. The same is true with the lower subject and the predicate *certain* in (12). In (11), it is the insertion between the matrix predicate and the subordinate clause that triggers the use of an explicit *that*.

In the following sentence, *that* has been omitted:

> (13) Hilary Robarts said impatiently: "But the body, when you first went back for the torch and saw her, you were ***certain*** she was dead?" (C8T 1491)
> [fully confident]

Compared to (10) – (12), it is clear why *that* can easily be omitted in (13): there is no discontinuity in constructions, subjects and objects are short and clear and the whole sentence is therefore simple.

There were only seven instances where the matrix clause had the pleonastic *it* as subject as in (14):

> (14) It soon became quite ***certain*** that the plane was in trouble. (B0B 102) [established fact]

The argument structure here is different from the one in (10) – (13) as the clausal subject is the only argument for *certain*. In five of the cases the meaning of *certain* is 'established fact' as in (14) while in the two remaining instances it is 'expected effect':

> (15) It's almost ***certain*** that the enemy <u>will</u> attack the Residency from the north, very likely at dawn tomorrow. (EFW 1359) [expected effect]
> (16) He added sadly, "It's not yet ***certain*** that Elinor <u>will</u> ever again be well enough to live without nursing support." (FPB 2543) [expected effect]

What makes (15) and (16) interesting is the fact that they both have the lower verb in the future tense. This evidently has an effect on the meaning of *certain* by shifting the Theme into irrealis compared to realis assertions that typically occur in the past tense as in (14).

5.2.2.2 *Wh*-clause

Most of the instances containing a *wh*-clause represent a structure where *certain* has two arguments and the meaning of *certain* is 'fully confident' as in the following:

> (17) The police don't seem to be entirely ***certain*** whether or not it was Burrows, that's all. (C8D 3390) [fully confident]

Only one of the eight instances of *wh*-clause has the dummy *it* as subject followed by a change of the meaning of *certain* into 'established fact':

> (18) "It's not ***certain***," he quavered, "whether it is allowable to issue a challenge by proxy – " (HA3 3662) [established fact]

In all cases where *certain* is connected to a *wh*-clause it occurs in a sentence expressing negative assertion as can be seen in (17) and (18). This evidence supports the same observation made by Quirk et al. (1985, 1225) concerning all adjectives.

5.2.2.3 *To*-infinitive

There were five instances of *to*-infinitive complements connected to *certain* in the corpus data. Four of them had the same structure as (19):

(19) He was ***certain*** <u>to ask</u> her to dance when she looked like this! (BMW 1718)
[certain to do]

Sentence (19) can be analyzed by applying Carnie's diagram of raising:

no θ role   Agent
↓   ↓
(20) [_____ was *certain* [he to ask her to dance]].
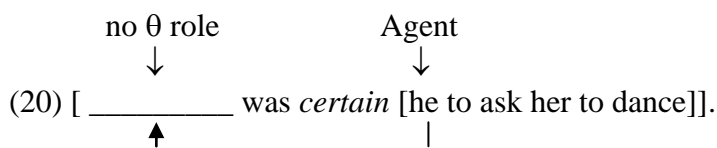
The lower predicate *ask* first assigns the role of an Agent to the NP *he* which is then raised to its final position. The whole subordinate clause is the only argument of *certain* and gets the theta role of a Theme. In fact, it is the idea of *he asking* that was certain instead of just *he* being certain. The meaning of *certain* here is therefore 'certain to do'. The following three instances of *to*-infinitive complement represent the same structure and meaning as (19):

(21) He was also the one good horse Kelly had been ***certain*** <u>to ride</u> – while Bill owned him. (BP7 199) [certain to do]
(22) Very soon, he knew, he would have to choose sides, and despite the success of tonight's endeavour, and the security he'd won with it, he was by no means sure that he belonged amongst the ranks of the purgers, even though they were ***certain*** <u>to carry</u> the day. (CRE 3032) [certain to do]
(23) He was glad to see him, for he knew him for a tough, sturdy fellow, who was considered ***certain*** <u>to get</u> into the Owsla as soon as he reached full weight. (EWC 425) [certain to do]

There was one instance of *certain* followed by *to*-infinitive with a more complex structure:

(24a) He said if anything bad ever happened in the school, it was ***certain*** <u>to be</u> his daughter who did it. (CH4 1068) [certain to do]

Instead of an animate subject (24a) has an expletive *it* in the subject position of the matrix clause. The canonical word order SVO has been modified by a type of extraction called clefting (Postal 1994, 162; Vosberg 2003b, 201). Consider (24b) where the clefting is illustrated:

(24b) [   ] was ***certain*** [it to be his daughter who did it.]

The whole lower clause is the subject argument of *certain*. Based on this approach, the thing that is certain here is *being his daughter* instead of the one *who did it*. Thus, with the emphasis on the action, *certain* in (24a) gets the same sense 'certain to do' as with the other *to*-infinitive constructions (19) – (23).

5.2.2.4 *Of –ing*

Consider the two instances of *certain* licensing an *of -ing* complement found in the data:

(25a) After waiting some time, instead of going to my bedroom and packing again, as before dinner I'd been so ***certain***, at last, <u>of being</u> able to do, I went in search of him. (FPF 3533) [fully confident]

(26) I couldn't see half my customers – and then I had to leave early to be ***certain*** <u>of getting</u> back. (FRJ 1713) [fully confident]

On the surface the construction in these two sentences seems similar to raising in the previous subsection. In the following, the cat idiom is used to demonstrate the difference between the two:

(27) The cat is *certain* <u>to be</u> out of the bag. (idiomatic meaning → raising)

(28) The cat is *certain* <u>of being</u> out of the bag. (non-idiomatic meaning → control)

As can be seen in (28), it is the *cat* that is certain and not *being out* as in (27). Thus, the meaning of *certain* is 'fully confident'. Based on this comparison (25a) and (26) involve a control construction.

Now we examine sentence (25a) more in detail by applying the control diagram to it:

(25b) [I had been so *certain* [PRO of being able] [PRO to do]].
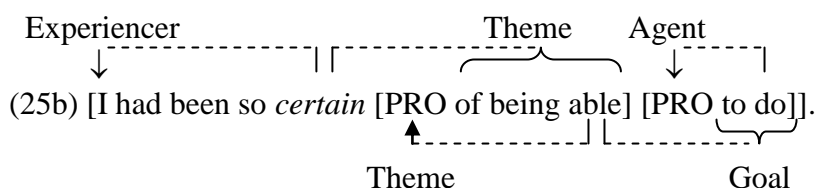
In (25b) we can see how each predicate assigns theta roles to their arguments. The matrix predicate *certain* has two arguments in this construction. The subject argument *I* gets the theta role of an Experiencer and the external argument *of being able* gets the role of a Theme. There is no NP movement, only one special NP, namely PRO, is added in the subject position of the non-finite clause. Now the lower predicate can assign the theta role of a Theme to PRO and not to *I,* which would violate the Theta Criterion. PRO is co-referential with the NP *I* that controls it. In this sentence, there happens to be another control construction with the predicate *able*. It also assigns two roles: a Theme and a Goal. PRO is added in the subject position of *to do* which assigns it the role of an Agent. This PRO is controlled by the higher PRO (cf. Manzini 1983) although ultimately they are both coreferential with the matrix subject. In other words, both constructions involve subject control.

6 Conclusions

The aim of this thesis was to discover the different complement patterns that occur with the adjective *certain* and whether there has been a change in their usage between Late Modern English and Present-day English. The purpose was also to find out if there is any connection between form and meaning i.e. between different complement constructions and senses of *certain*. These research questions have been answered in this thesis and in this final chapter, I will present concluding remarks on the findings.

First, we shall look at the numbers. There were nine different types of construction connected to *certain* found in the two corpora, the CLMET and the BNC. Seven of them were complements of which four were sentential and three non-sentential. Their distribution between the two corpora is shown in table 7 below:

	Pattern	Normed frequency CLMET	Normed frequency BNC
Sentential complements	*that*-clause (of which *that* omitted)	86 (26)	38 (17)
	wh-clause	-	5
	to -inf.	2	3
	of -ing	2	1
Non-sentential complements	Ø	20	18
	of + NP	10	6
	in + NP	2	-
Disregarded constructions	Attributive	318	79
	Phrasal	4	6
	TOTAL	**444**	**156**

Table 7. The normalized frequency of different patterns found in the CLMET and the BNC.

When examining table 7 we notice that the most frequent pattern in both corpora was the *that*-clause. The usage of a *that*-clause complement with *certain* is, however, considerably less frequent in Present-day English. On the other hand, the appearance of *wh*-clauses in the BNC data shows an increase in the usage of another sentential complement. However, the overall use of sentential complements with *certain* has decreased towards modern times.

One candidate for a new complement pattern, the *in* + NP, emerged in the CLMET data with only one illustration. Since there were no other instances, and the context was not clear enough to be able to make definite conclusions, the probability of this being a new pattern is quite non-existent.

As regards the constructions that were not in the focus of this thesis, it is worth noting that the attributive use of *certain* is reduced to half today compared to what it was in Late Modern English times. The reasons for this trend cannot be found within this study but it would be an interesting topic for another research.

Turning now to the relationship between sense and structure, six different meanings of *certain* could be found in co-occurrence with its complement constructions. The distribution of the complement patterns across these senses in the dictionaries and in the two corpora is presented in table (8):

	'reliable'	'sure to follow'	'expected effect'	'established fact'	'fully confident'	'certain to do'
OED & Cobuild						
Sentential		*that*-cl	*of* -ing	*that*-cl	*that*-cl, *wh*-cl	*to*-inf
Non-sentential		Ø	Ø	Ø	Ø, *of* + NP	
CLMET						
Sentential				*that*-cl	*that*-cl, *of* -ing	*to*-inf
Non-sentential		Ø	*of* + NP	Ø, *of* + NP	Ø, *of* + NP, *in* + NP	
BNC						
Sentential			*that*-cl	*that*-cl, *wh*-cl	*that*-cl, *wh*-cl, *of* -ing	*to*-inf
Non-sentential	Ø		Ø	Ø	Ø, *of* + NP	

Table 8. Meaning – pattern compatibility in the dictionaries and in the CLMET and the BNC.

As can be seen in table (8), one sense, namely 'fully confident', is compatible with all types of complement with the exception of the *to*-infinitive which is restricted to the sense 'certain to do'. The sense 'fully confident' was also the most frequent in both corpora. The second most common sense was 'established fact' the other meanings remaining rather marginal.

One fundamental difference was discovered in the occurrence of the two most common senses. The meaning 'fully confident' occurs exclusively in environments where the subject argument is animate whereas the sense 'established fact' is connected to constructions with an inanimate subject.

Certain is a one or two-place predicate and there is a difference in the complementation it selects with each construction. When *certain* has only one argument the possible patterns are zero complement, *to*-infinitive, *that*-clause and *wh*-clause. In case *certain* has two arguments the complement patterns it can license are *that*-clause, *wh*-clause, *of −ing* and *of + NP*.

REFERENCES

Primary sources:

BNC: The British National Corpus

CLMET: The Corpus of Late Modern English Texts (Original version)

Secondary sources:

Ball, C. N. 1994. "Automated Text Analysis: Cautionary Tales, "*Literary and Linguistic Computing* 9, 295-302. 71

Bas, Aarts, Maria José Lopez-Couso, Belén Mendez Naya. 2011. Draft version.

Biber, et al. 1998. *Corpus Linguistics*. Cambridge: Cambridge University Press.

Biber, Douglas. 2010. *Corpus-based and corpus-driven analyses of language variation and use.* In The Oxford Handbook of Linguistic Analysis, ed. Heine, Bernd and Heiko Narrog, 159-192. Oxford: Oxford University Press.

Bolinger, Dwight. 1968. "Entailment and the Meaning of Structures," *Glossa* 2, 119-127.

Carnie, Andrew. 2002. *Syntax*. Oxford: Blackwell.

Chomsky, Noam. 1981. *Lectures on Government and Binding.* Dordrecht: Foris

Collins Cobuild Dictionary online, available at: http://www.collinsdictionary.com

Davies, William and Stanley Dubinsky. 2004. *The Grammar of Raising and Control.* Oxford: Blackwell.

Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection", *Language* 67, 547-619.

Fanego, Teresa. 2004. *Is Cognitive Grammar a usage-based model? Towards a realistic account of English sentential complements.* Miscelánea: A Journal of English and American Studies 29, 23-58.

Fillmore, Charles. 1968. "The Case for Case", in E. Bach and Robert Harms, *Universals in Linguistic Theory.* New York: Holt, Reinhart and Winston.

Haegeman, Liliane. 1991. *Introduction to Government and Binding Theory*. Oxford: Basil Blackwell.

Herbst, Thomas et al. eds. 2004. *A Valency Dictionary of English*. Berlin: de Gruyter.

Hoffmann, Sebastian, Stefan Evert, Nicholas Smith, David Lee and Ylva Berglund Prytz. 2008. Corpus Linguistics with BNCweb-a Practical Guide. Frankfurt am Main: Peter Lang.

Huang, James. 1997. *Introduction to Syntax*. 1997 LSA Linguistic Institute.

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

Huddleston, Rodney and Geoffrey Pullum. 2002. *The Cambridge Grammar of English.* Cambridge: Cambridge University Press.

Hunston, Susan and Gill Francis. 2000. Pattern Grammar: A Corpus Driven Approach to the Lexical Grammar of English. Amsterdam: John Benjamins Publishing Company.

Hunston, Susan. 2002. "Pattern grammar, language teaching, and linguistic variation: Applications of a corpus-driven grammar." In *Studies in Corpus Linguistics: Using Corpora to Explore* Linguistic Variation, ed. R. Reppen, S.M. Fitzmaurice, D.Biber, 167-183. Amsterdam: John Benjamins Publishing Company.

Leech, Geoffrey. 1968. "Some Assumptions in the Metatheory of Linguistics," *Linguistics* 39, 87-102.

Lindquist, Hans. 2009. Corpus Linguistics and the Description of English. Edinburgh: Edinburgh University Press.

Manzini, Rita. 1983. "On control and control theory," *Linguistic Inquiry*, Vol. 14, No. 3, pp. 421-446. Cambridge. The MIT Press.

Mukherjee, Joybrato. 2004. *The state of the Art in Corpus linguistics: Three Book-length perspectives*. English Language and Linguistics 8 (1): 103-119.

Onions, C.T. et al. eds. 1966. *The Oxford Dictionary of English Etymology.* Oxford. Oxford University Press.

Partridge, Eric. 1966. 4th ed. *Origins. A Short Etymological Dictionary of Modern English.* London and Henley: Routledge & Kegan Paul

Postal, Paul M. 1974. On Raising: One Rule of Grammar and Its Theoretical Implications. Cambridge, Massachusetts and London: The MIT Press.

Postal, Paul M. 1994. "Contrasting extraction types", *Journal of linguistics* 30, 159-186.

Rissanen, Matti. 1991. On the history of *that*/zero as object clause links in English. In: Karin Aijmer and Bengt Altenberg (eds.), *English Corpus Linguistics: Studies in Honour of Jan Svartvik*, 272-289. London: Longman.

Quirk, R., S. Greenbaum, G. Leech, & J. Svartvik. 1985. *A Comprehensive Grammar of the English Language.* London: Longman.

Rohdenburg, Günter. 1995. "On the replacement of finite complement clauses by infinitives in English." *English Studies* 4: 367-388.

Rohdenburg, Günter. 1996. "Cognitive complexity and increased grammatical explicitness in English." *Cognitive Linguistics* 7, 2: 149-182.

Simpson, J.A. and Weiner, E.S.C. eds. 1989. *The Oxford English Dictionary.* 2nd ed. Oxford: Clarendon Press.

Sinclair, John et al. eds. 2006. *Collins Cobuild English Language Dictionary*. 5th ed. Glasgow: HarperCollins Publishers.

Smet, Hendrick de. 2005. A Corpus of Late Modern English Texts," *ICAME Journal* 29, 69-82.

Svartvik, Jan. 1992. "Corpus linguistics comes of age." In Directions in corpus linguistics. Proceedings of Nobel Symposium 82. Stockholm 4-8 august 1991, ed. J. Svartvik, 7-16. Berlin - New York: Mouton de Gruyter.

Tognini-Bonelli, Elena. 2001. *Corpus Linguistics at Work.* Amsterdam: John Benjamins Publishing Company.

Vosberg, Uwe. 2003. "The role of extractions and *horror aequi* in the evolution of –ing complements in Modern English." In Determinants of Grammatical Variation in English, ed. G. Rohdenburg and B. Mondorf, 305-328. Berlin: Mouton de Gruyter.