

## The English *It*-Cleft

# Topics in English Linguistics

## 79

*Editors*

Elizabeth Closs Traugott  
Bernd Kortmann

De Gruyter Mouton

# The English *It*-Cleft

A Constructional Account  
and a Diachronic Investigation

*by*

Amanda L. Patten

De Gruyter Mouton

ISBN 978-3-11-027780-7  
e-ISBN 978-3-11-027952-8  
ISSN 1434-3452

*Library of Congress Cataloging-in-Publication Data*

A CIP catalog record for this book has been applied for at the Library of Congress.

*Bibliographic information published by the Deutsche Nationalbibliothek*

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;  
detailed bibliographic data are available in the Internet at <http://dnb.dnb.de>.

© 2012 Walter de Gruyter GmbH, 10785 Berlin/Boston

Cover image: Brian Stablyk/Photographer's Choice RF/Getty Images

Printing: Hubert & Co. GmbH & Co. KG, Göttingen

⊗ Printed on acid-free paper

Printed in Germany

[www.degruyter.com](http://www.degruyter.com)

## Acknowledgements

This book is a revised version of my 2010 University of Edinburgh dissertation *Cleft Sentences, Construction Grammar and Grammaticalization*. I would first like to thank my PhD supervisor, Nikolas Gisborne, and the examiners of my thesis, Geoff Pullum and Kersti Börjars. I am immensely grateful to the AHRC for funding my postgraduate studies. Without this financial support, I could not have continued in education.

Among those who have discussed my work with me, I would especially like to thank Graeme Trousdale and Regina Weinert. I am also grateful to Roger Higgins for taking the time to correspond with me early on in my studies and for encouraging me in my work on clefts. I must also thank Catherine Ball, Nancy Hedberg, Bettelou Los and Javier Pérez-Guerra for helping me to access unpublished or not widely distributed works, as well as Susan Pintzuk, Ann Taylor and Anthony Warner for introducing me to the English historical parsed corpora series and for their guidance in its use. Thanks also to Linda van Bergen and Meg Liang for help with translating some of the historical data.

Above all, I thank Elizabeth Traugott for showing interest in my work and for her guidance throughout the writing of this book. I am very grateful for her encouragement and for her valuable comments on my written work at various stages in the process. Many thanks also to Kirsten Börgen, Julie Miess and Angelika Hermann at De Gruyter Mouton for help in producing the finished piece.

ALP  
Newcastle, 2012



# Contents

|  |     |
|--|-----|
| <i>Acknowledgements</i> . . . . .  | v   |
| <i>Contents</i> . . . . .  | vii |
| <b>Chapter 1</b> Introduction and background. . . . .                              | 1   |
| 1. An outline of the project. . . . .  | 1   |
| 2. An overview of the literature on cleft sentences . . . . .                      | 5   |
| 2.1. The expletive approach. . . . .   | 5   |
| 2.2. The extraposition approach. . . . .   | 7   |
| 3. A constructional approach to <i>it</i> -clefts. . . . .                         | 9   |
| 4. A diachronic approach to <i>it</i> -clefts . . . . .                            | 12  |
| 5. Methodology. . . . .  | 14  |
| <b>Chapter 2</b> A model of language structure and language change. . . . .        | 16  |
| 1. Some basic assumptions. . . . .   | 16  |
| 2. A constructional model of language structure. . . . .                           | 17  |
| 3. A constructional model of language change . . . . .                             | 21  |
| 4. The application to <i>it</i> -clefts and copular constructions . . . . .        | 25  |
| <b>Chapter 3</b> Specificational copular constructions. . . . .                    | 27  |
| 1. Different and competing analyses . . . . .                                      | 27  |
| 1.1. The equative approach . . . . .   | 28  |
| 1.2. The inverse approach . . . . .  | 30  |
| 1.3. A less formal approach . . . . .  | 32  |
| 2. Specification as (the inverse of) nominal predication . . . . .                 | 34  |
| 2.1. Specification and definite NP predicates . . . . .                            | 34  |
| 2.2. Specification and inversion. . . . .  | 38  |
| 2.3. Capturing this account in cognitive and constructional<br>frameworks. . . . . | 42  |
| 3. Accounting for the behaviour of indefinite NPs. . . . .                         | 47  |
| 3.1. Specification and indefinite NP predicates. . . . .                           | 48  |
| 3.2. An account based on discourse requirements. . . . .                           | 49  |
| 3.3. An account based on definiteness. . . . .                                     | 51  |
| 4. Summarizing and extending the account. . . . .                                  | 56  |
| 4.1. An overview of specificational NP <i>be</i> NP sentences . . . . .            | 56  |
| 4.2. Positioning this account in relation to the literature . . . . .              | 58  |

|   |     |
|---|-----|
| 4.3. Other specificational copular constructions . . . . .                        | 62  |
| 4.3.1. <i>Th</i> -clefts as specificational copular sentences. . .                | 63  |
| 4.3.2. <i>Wh</i> -clefts as specificational copular sentences . .                 | 64  |
| 4.3.3. <i>All</i> -clefts as specificational copular sentences. . .               | 68  |
| 4.3.4. A family of specificational copular sentences. . .                         | 69  |
| <b>Chapter 4</b> <i>It</i> -clefts as specificational copular sentences. . . . .  | 71  |
| 1. The English <i>it</i> -cleft. . . . .  | 71  |
| 1.1. A “discontinuous constituent” account of <i>it</i> -clefts . . . . .         | 72  |
| 1.2. Explaining the <i>it</i> -cleft’s pragmatic properties. . . . .              | 78  |
| 1.2.1. Focus . . . . .  | 79  |
| 1.2.2. Presupposition . . . . .   | 80  |
| 1.2.3. Exhaustiveness. . . . .  | 82  |
| 1.2.4. Contrast . . . . .   | 85  |
| 1.3. Explaining the <i>it</i> -cleft’s structural properties . . . . .            | 87  |
| 1.3.1. The behaviour of the cleft clause. . . . .                                 | 88  |
| 1.3.2. The evidence for VP constituency. . . . .                                  | 93  |
| 1.3.3. The evidence from agreement . . . . .                                      | 96  |
| 1.4. Interim summary. . . . .   | 101 |
| 2. A comparison with expletive accounts of <i>it</i> -clefts. . . . .             | 102 |
| 3. A comparison with other extraposition accounts of <i>it</i> -clefts . . . . .  | 107 |
| 3.1. The early extraposition accounts of the 1970s . . . . .                      | 107 |
| 3.2. The more recent discontinuous constituent accounts . . . . .                 | 110 |
| 3.3. A different extraposition account? . . . . .                                 | 114 |
| 4. A comparison with other constructional accounts of <i>it</i> -clefts . . . . . | 115 |
| <b>Chapter 5</b> Other varieties of <i>it</i> -cleft . . . . .                    | 120 |
| 1. Beyond the archetypal <i>it</i> -cleft . . . . .                               | 120 |
| 2. Predicational (and proverbial) <i>it</i> -clefts. . . . .                      | 122 |
| 2.1. An expletive approach to predicational <i>it</i> -clefts. . . . .            | 123 |
| 2.2. Predicational <i>it</i> -clefts and the inverse approach . . . . .           | 127 |
| 2.3. Predicational <i>it</i> -clefts and the equative approach . . . . .          | 131 |
| 3. <i>It</i> -clefts with non-nominal foci. . . . .                               | 134 |
| 4. Informative-presupposition (IP) <i>it</i> -clefts . . . . .                    | 140 |
| 5. Summary. . . . .   | 146 |
| <b>Chapter 6</b> The <i>it</i> -cleft and earlier periods of English . . . . .    | 148 |
| 1. Beyond the present-day language system. . . . .                                | 148 |
| 2. The early history of the English <i>it</i> -cleft . . . . .                    | 149 |
| 3. A restrictively modified pronoun? . . . . .                                    | 152 |



|   |     |
|---|-----|
| 4. An obligatorily extraposed relative clause? . . . . .                                      | 157 |
| 5. An unusual pattern of agreement? . . . . .   | 162 |
| 6. The evidence from Old English gender agreement . . . . .                                   | 170 |
| 7. The <i>it</i> -cleft as a relic from an earlier time . . . . .                             | 177 |
| 8. Summary . . . . .  | 183 |
| <b>Chapter 7</b> The <i>it</i> -cleft's development over time. . . . .                        | 184 |
| 1. A diachronic investigation . . . . .   | 184 |
| 2. The corpora, the search and the selection process . . . . .                                | 184 |
| 2.1. OE presentational/impersonal sentences. . . . .  | 186 |
| 2.2. Existential sentences with <i>it</i> . . . . .   | 190 |
| 2.3. The pattern <i>I it am</i> . . . . .   | 191 |
| 2.4. Other constructions mistaken for clefts. . . . .   | 192 |
| 2.5. Interim summary. . . . .   | 193 |
| 3. Frequency information. . . . .   | 194 |
| 4. Changes to the clefted constituent . . . . .   | 196 |
| 5. Changes to the cleft clause. . . . .   | 204 |
| 6. Summary . . . . .  | 211 |
| <b>Chapter 8</b> The <i>it</i> -cleft and constructional change. . . . .                      | 212 |
| 1. The two kinds of constructional change. . . . .  | 212 |
| 2. A grammatical constructionalization account . . . . .                                      | 213 |
| 3. Some alternative explanations . . . . .  | 218 |
| 3.1. An impersonal account . . . . .  | 219 |
| 3.2. A Celtic account . . . . .   | 221 |
| 3.3. A word order account . . . . .   | 222 |
| 3.4. Interim summary. . . . .   | 224 |
| 4. Why do <i>it</i> -clefts undergo a construction-specific development? . . . . .            | 225 |
| 4.1. Why do <i>it</i> -clefts develop a construction-specific range of<br>foci? . . . . .     | 225 |
| 4.2. Why do <i>it</i> -clefts develop construction-specific discourse<br>functions? . . . . . | 232 |
| 4.3. Summary. . . . .   | 241 |
| <b>Chapter 9</b> Conclusions . . . . .  | 243 |
| <i>Corpora and data sources</i> . . . . .   | 248 |
| <i>References</i> . . . . .   | 249 |
| <i>Index</i> . . . . .  | 266 |



# Chapter 1

## Introduction and background

### 1. An outline of the project

*It*-clefts are interesting for a number of reasons. For one thing, they have a non-standard structure which appears not to conform to the general rules of the language. If we take a look at the example in (1), we can see that *it*-clefts have four main components: the introductory pronominal *it*, a form of the copular verb *be*, a postcopular phrasal element and a sentence-final clause.

(1)     [*It*] [*was*] [*Frank*] [*that complained*]

From this example, we can see that (for the linguist) the *it*-cleft's syntactic configuration is difficult to make sense of. The *that*-clause is structured internally like a restrictive relative: it contains a gap (in subject position) which corresponds to the constituent that precedes the clause. However, proper names, such as *Frank*, are full noun phrases. As such, they cannot normally be modified by restrictive relative clauses. So how does this clause relate to the postcopular element, if at all? Can we really call this a restrictive relative clause? If so, what does it modify? If not, are we dealing with a clausal structure that is unique to the *it*-cleft? Equally problematic is the role of initial *it*. Is this an expletive dummy subject and if so, why is it there? Does it operate as a syntactic placemaker and if so, for which element? Or is the constituent *it* related in a different way to other elements in the sentence?

In addition, *it*-clefts have a number of unusual semantic, pragmatic and discourse-functional properties. These are particularly interesting since it is not immediately clear which elements in the cleft structure contribute to the meaning of the construction. For example, the *it*-cleft is a focusing construction. The primary informational content is placed in the syntactically marked postcopular focal position and is often given primary stress, see

(2).<sup>1</sup> However, it is not at all obvious why this particular syntactic configuration should be chosen as a focusing device. Is this its primary function?

(2) *It was FRANK that complained*

A further property of the *it*-cleft is that it exhibits exhaustiveness effects. For example, in (2) we assume that *Frank* was the only person who complained on this occasion. They are also presuppositional; the information in the sentence-final clause is not asserted and is preserved under negation. For instance, in example (3), we are told that *Frank didn't complain* but we are left with the presupposition that *somebody (else) did complain*. This begs the question, where do these pragmatic meanings come from? Which elements contribute to them?

(3) *It wasn't Frank that complained*

Cleft sentences also have a specificational (or identifying) meaning. For some authors, specificational meaning involves a *value-variable* relationship (see Higgins 1979; Declerck 1988). For example, (2) identifies *Frank* as the value for the variable described by the sentence-final clause, *x complained*. However, for others, specificational meaning is attributed to a special use of the copular verb. So does *be* have a specificational meaning in the *it*-cleft? If not, where does the specificational meaning of *it*-clefts come from?

To a large extent, how these questions are answered (and perhaps whether they are even asked at all) depends on how we think *it*-clefts relate to other constructions, or configurations, in the language. Most approaches to *it*-clefts fall into two broad categories: those that understand *it*-clefts in relation to simple subject-predicate sentences, such as (4), and those that relate *it*-clefts to other specificational copular sentences, such as (5).

(4) *Frank complained*

(5) *The one that complained was Frank*

---

1. I use the term *focus* to refer to a unit of information structure where the assertion differs from the presupposition (see Chapter 4, Section 1.2.1). Although focus is often marked by intonation, Lambrecht (1994: 208) observes that “accent placement and focus marking are not to be equated”. Where focus marking is unclear or ambiguous in my examples, I make use of small capitals to indicate the marking of focus by intonation.

I discuss these two approaches and the analyses that result from them in Section 2. Essentially, justification for the first approach comes from the truth-conditional equivalence between *it*-clefts and simple subject-predicate sentences. From this perspective, *it*-clefts are viewed primarily as a means of marking focus syntactically. The second approach, on the other hand, builds on the fact that the *it*-cleft is a copular construction with a specificational meaning. So which is the better approach? The answer to this question depends upon what we think is the primary function of *it*-clefts and asking which perspective can best explain the range of properties that *it*-clefts display.

There are also different varieties of *it*-cleft which are sometimes regarded as separate structures. One domain of variability involves the category of elements that can occur in the postcopular (focal) position. Although, most frequently, the focal element is a noun phrase, *it*-clefts allow a range of other constituents to occur as the complement of *be*, such as the prepositional phrase in (6). So should these examples be analysed in the same way as those with nominal foci or do they require a separate analysis? Can the sentence-final clause still be analysed as a restrictive relative if the immediate antecedent is not nominal?

- (6) *It's in October that he's leaving*

Another domain of variation relates to the information status of the sentence-final clause. In *it*-clefts, the clausal component is typically associated with expressing discourse-old information, as in (7). In this example, we know from the prior discourse that a letter has been written and so the open proposition *x wrote it* is given information. However, in other cases, such as (8), the information expressed by the sentence-final clause is not given by the previous discourse and the proposition, that *someone once said "Laws are silent at times of war"*, does not even have to be known (or familiar) to the intended audience. Do these examples represent two different types of *it*-cleft? And if so, how are they related? Is one more basic, or prototypical, than the other?

- (7) A: *Did Max write the letter?*  
 B: *No. It was Walter **who wrote it**.*
- (8) (Start of lecture)  
*It was Cicero **who once said, "Laws are silent at times of war"***

A further domain of variation concerns the relationship between specificational and predicational tokens. *It*-clefts are usually identified as having a specificational meaning (see above). However, superficially similar proverbial sentences, such as (9), have a predicational meaning. For instance, (9) is most closely paraphrased by the predicational copular sentence in (10). In both sentences, the postcopular element describes, rather than identifies, the referent (*the road that has no turning*) as *long*.

(9) *It is a long road that has no turning*

(10) *The road that has no turning is a long one*

How does this structure relate to the specificational *it*-cleft, if at all? Can proverbial sentences, such as (9), really be called clefts? *It*-clefts also seem to resemble extraposed sentences, such as (12). For instance, on the surface, the *it*-cleft in (11) differs only in that it contains a gap in the sentence-final clause. Do *it*-clefts share more than just an apparent likeness with this extraposed structure?

(11) *It was the Colonel [that \_\_ survived]*

(12) *It is a miracle [that he survived]*

In this book, I provide answers to these questions by examining *it*-clefts within the framework of construction grammar (cf. Croft 2001; Goldberg 1995, 2006; Lakoff 1987). Construction grammar was developed with a view to providing full and explanatory accounts of specialized linguistic patterns as well as broad generalizations. On this model, constructions are not considered the epiphenomenal byproducts of a combination of componential meaning and highly general rules. Instead, aspects of form and meaning can be encoded by the construction itself. Since much of the *it*-cleft's structure and use cannot be predicted from general patterns of correspondence, it is well-suited to treatment within a constructional approach. I come back to this issue in Section 3.

In addition to providing a synchronic account of the English *it*-cleft, I also examine the construction's diachronic development. Relevant questions here include: What is the origin of the *it*-cleft construction? How has the *it*-cleft (and its relationship to other constructions) changed over time? How did the different varieties of *it*-cleft emerge? This study contributes to the recent literature on diachronic construction grammar. As a complex and specialized linguistic pattern, the English *it*-cleft provides an example of

how larger (multi-word) constructions undergo change. I discuss this aspect of my investigation in Section 4.

In this section, I have provided a very brief outline of the different types of phenomena discussed and the kinds of questions addressed throughout this book. In the next section, I provide an introductory background into the literature on cleft sentences. Section 3 asks why construction grammar is helpful in the treatment of *it*-clefts. Here, I present an overview of the particular kind of cleft analysis argued for in this book and I compare it to other constructional accounts proposed in the literature. The historical facts of the English *it*-cleft, along with my diachronic construction grammar account, are sketched in brief in Section 4. I explain how I use this diachronic evidence to both support and inform my synchronic analysis. The methodology employed in this study is discussed in Section 5.

## 2. An overview of the literature on cleft sentences

As I noted in Section 1, authors tend to view *it*-clefts either from the perspective of their relationship to truth-conditionally equivalent subject-predicate sentences or from the perspective of their relationship to other specificational copular constructions. In this section, I outline these two approaches and the analyses that result from them. Although the individual proposals differ, these opposing viewpoints lead to two different kinds of analysis: those that treat the postcopular phrase as the preposed argument of the proposition expressed in the sentence-final clause and those that consider the sentence-final clause to be associated in some way with the initial element *it*. The purpose of this section is not to provide an exhaustive and comprehensive review of the literature, but simply to highlight intellectual trends in the history of the analysis of the construction.

### 2.1. The expletive approach

For many authors, *it*-clefts are considered primarily as information structure variants of syntactically more basic sentences. From this perspective, *it*-clefts do not differ dramatically in their semantic content from canonical subject-predicate sentences, but are marked by the way that this informational content is presented (Ward, Birner, and Huddleston 2002). Unlike their canonical counterparts, *it*-clefts have a fixed information structure: the information that is to be foregrounded is placed in the postcopular position

while the remaining semantic content is backgrounded into a sentence-final clause, as shown in (13).

(13) *It was* [[*Frank*]<sub>i</sub> [*that* \_\_\_\_<sub>i</sub> *complained*]] [it-cleft]

(14) *Frank complained* [canonical]

The analyses resulting from this approach assume that the focal element in *it*-clefts enters into a predication relationship with the information in the sentence-final clause; this accounts for their truth-conditional equivalence with simple subject-predicate sentences. From this, it follows that the initial element *it* and (in most accounts) the copular verb *be* are semantically empty, serving only to introduce, or foreground the postcopular element. As a result, in the cleft literature, these analyses are referred to cumulatively as the expletive approach; common to all such accounts is the assumption that the initial pronoun *it* does not play an essential role in the interpretation of the sentence.

An early example of an expletive account is detailed by Jespersen (1937: 83–89). He suggests that *it*-clefts are syntactically identical to their noncopular counterparts except for the addition of a “lesser subject and verb” and a “connective word”. So, for instance, the elements *it*, *be* and *that* in (15) are semantically empty, with *Frank* and *complained* entering into a predication relationship. The example is formalized using Jespersen’s notation.<sup>2</sup>

(15) [*It was*] *Frank* [*that*] *complained*  
[sv]        S    [3<sup>c</sup>]        V

According to Jespersen, this analysis explains why *it*-clefts are used as a means of marking focus syntactically. He notes, “A cleaving of a sentence by means of *it is*...serves to single out one particular element of the sentence and very often, by directing attention to it and bringing it, as it were, into focus, to mark a contrast” (Jespersen 1949: 147).

A number of similar analyses were developed within the generative tradition of the 1980s. The details of these analyses differ. For example, Rochemont (1986) suggests that the postcopular element is situated within

2. The use of square brackets in Jespersen’s (1937: 86) formalism indicates that this information is “extraposed” relative to the sentence proper. Lower case *s* and *v* indicate the “lesser” (i.e. expletive) subject and verb and 3<sup>c</sup> represents a “tertiary connective”.



the sentence-final clause at deep structure, while for Williams (1980) and Heggie (1988) these components are coindexed at the level of surface structure by a predication rule. Chomsky (1977) claims that the *it*-cleft is a type of topicalization construction involving *wh*-movement. The relationship between the postcopular element and the sentence-final clause is therefore akin to that between topic and comment. For Delahunty (1982, 1984) on the other hand, these components enter into a predication relation at the level of logical form. Delahunty converts the sentence-final clause into a function (via lambda abstraction) which takes the postcopular element as its argument; after a reduction operation, the Logical Structure of cleft sentences is equivalent to that of their noncopular counterparts.

Although they use different mechanisms to achieve it, these authors assume that *it*-clefts and truth-conditionally synonymous sentences share a level of representation. Common to all of these accounts then is the treatment of the initial pronoun *it* as an expletive element, the analysis of the sentence-final clause as being in some way related to the postcopular element, and the understanding that the primary function of *it*-clefts is as a focusing device.

## 2.2. The extraposition approach

For others, *it*-clefts are considered foremost as specificational copular sentences. From this perspective, *it*-clefts are analysed in relation to corresponding pseudocleft sentences and sometimes to noncleft copular constructions with an identifying function. The term *pseudocleft* is commonly used to encompass both *wh*-clefts, which are introduced by *wh*-words, and *th*-clefts, which are introduced by the definite article and one of a small number of semantically general head nouns such as *the one* or *the thing*.<sup>3</sup>

- |      |   |                            |
|------|---|----------------------------|
| (16) | <i>It's grape soda that I like best</i>         | [ <i>it</i> -cleft]        |
| (17) | <i>What I like best is grape soda</i>           | [ <i>wh</i> -cleft]        |
| (18) | <i>The thing that I like best is grape soda</i> | [ <i>th</i> -cleft]        |
| (19) | <i>My favourite drink is grape soda</i>         | [noncleft NP <i>be</i> NP] |

---

3. My use of the term *th-cleft* is from Collins (1991a, 1991b). This term is used in a different sense in Ball (1977) and Hedberg (1990, 2000) to refer to cleft sentences introduced by demonstratives, such as *this* or *that*.

Each of these examples has the function of identifying (or specifying) the postcopular element *grape soda* as matching a certain description. Like the *it*-cleft in (16), pseudoclefts also contain clausal elements. However, in (17) and (18) these clauses are in subject position. This suggests that the *it*-cleft is an extraposition construction: the sentence-final clause is not connected to the focal element; instead, it is related in some way to the initial *it*.

Again, an early example of such an approach is provided by Jespersen (1927). Prior to his (1937) account, outlined above, he proposed a “transposition analysis” of *it*-clefts. In the following passage, Jespersen suggests that *it*-clefts are paraphrased most closely by other specificational copular constructions. Here the sentence-final clause is analysed as a restrictive relative, modifying the constituent *it*. He notes,

...it is not really the antecedent (or what looks like the antecedent) that is restricted by a relative clause. When we say ‘it is the wife that decides’ or ‘it was the Colonel I was looking for’ what we mean is really ‘the wife is the deciding person’ and ‘the Colonel was the man I was looking for’: the relative clause thus might be said to belong rather to ‘it’ than to the predicative following after ‘it is’  
(Jespersen 1927: 88)

Other accounts that view *it*-clefts in relation to specificational copular sentences can be found in the transformational analyses of the 1970s, which derive *it*-clefts from pseudoclefts (see Akmajian 1970; Gundel 1977), or from the same source as pseudoclefts (see Wirth 1978). For these authors, the clausal element in subject position is extraposed, leading to the manifestation of *it* as either a placemaker or a pronominal copy. Again, the details of these analyses differ. For example, Akmajian and Wirth derive *it*-clefts via extraposition rules that are particular to cleft sentences, whereas Gundel suggests that this process is an instance of ordinary right-dislocation. For Gundel, the initial *it* is a pronominal copy of the right-dislocated clause, whereas for Akmajian, *it* seems to be an expletive element.

Bolinger (1972) takes an approach that is more in line with Jespersen’s (1927) original proposal, in which the relative clause restrictively modifies the constituent *it*. He suggests that analytic compound relatives, as in (20), “provide an ideal source” for *it*-clefts (Bolinger 1972: 110). Such sentences can undergo what he labels “inversion” (extraposition-from-NP), whereby the restrictive clause is extraposed but the nominal head remains *in situ*, as in (21). For *it*-clefts however, this “inversion” is obligatory rather than optional. On this analysis then, the constituent *it* is neither a placemaker, nor a pronominal copy, but a restrictively modified pronoun.

- (20) *That which he stole was money* [analytic compound]  
 (21) *That was money which (that) he stole* [inverted compound]  
 (22) *It was money which (that) he stole* [*it*-cleft]  
 (examples from Bolinger 1972: 109)

Consequently, for these authors, there is little consistency as to the exact role of *it* or how the relationship between *it*-clefts and other specificational sentences works. Nevertheless, what these extraposition analyses share is a concern for recognizing *it*-clefts primarily as specificational copular sentences.

### 3. A constructional approach to *it*-clefts

So which of these approaches is the better one? From the perspective of construction grammar, the choice is straightforward. In this section, I explain why construction grammar is a useful framework for representing and accounting for the unusual properties of *it*-clefts and why an approach that examines clefts in relation to specificational sentences is more compatible with the principles of construction grammar. I go on to provide an overview of my account of *it*-clefts before showing how this improves on the previous constructional analyses proposed in the literature.

In construction grammar, larger linguistic patterns are represented as symbolic pairings of form and meaning, much like individual lexical items. These complex constructions are made up of smaller units, which are also form-meaning pairs. However, since these correspondences are internal to the larger construction, such compositional meanings may nevertheless be construction-specific. Furthermore, on this model, constructions can sometimes encode meanings which are not compositional; that is, which cannot be attributed to its individual components. In recognizing the *construction* as a theoretical symbolic object, construction grammar therefore anticipates the existence of (and can represent) idiosyncratic grammatical information which cannot be predicted on the basis of highly general grammatical rules. Within this framework, all aspects of form and meaning (including aspects of *use*) are listed inside the construction and so form part of the speaker's grammatical knowledge. The theory of construction grammar is therefore ideally suited to the task of providing full and comprehensive accounts of the properties of more specialized linguistic patterns, such as the English *it*-cleft construction.

As Goldberg (2003: 120–121) notes, the explanatory power of constructional accounts comes from the requirement that each construction must be *motivated*; that is, there must be some reason as to why this particular construction should exist in the language. For the most part, the motivation for a construction comes from within the grammar. On a usage-based constructional theory, a speaker's grammatical knowledge is represented as a network of constructions (form-meaning pairs). Specialized linguistic patterns inherit properties from more general patterns. The more properties a construction inherits, the more it can be said to be motivated by the language system. Constructions that are related to one another are shown to inherit properties from the same general pattern, forming a *family* of constructions. A more detailed introduction to construction grammar is provided in Chapter 2.

The way that grammatical knowledge is organized in construction grammar suggests that analyses of *it*-clefts which are based on their relationship to structurally simple noncopular sentences will result in a less satisfactory account than an approach which views *it*-clefts in relation to other specificational copular sentences. On this model, inheritance links are posited between constructions that are both formally and functionally related, with an emphasis placed on similarities of *surface* form (Goldberg 2006: 23) and aspects of meaning that go beyond truth-conditional synonymy (Goldberg 1995: 103). Goldberg (1995: 108) says, “The intuition is that the existence of a given form with a particular meaning in no way motivates the existence of a different form with a closely related meaning”. Therefore, while noncopular subject-predicate sentences can often be used to paraphrase *it*-clefts, their truth-conditional synonymy is not necessarily expressed (as closely) in the grammatical system. As we might expect, accounts which view *it*-clefts in relation to structurally less complex sentences leave a number of questions unresolved: Why should focus be marked using this particular structure? Why do *it*-clefts have so many semantically empty elements? Where do the existential presuppositions and the property of exhaustiveness come from?

I view the *it*-cleft foremost as a member of the family of specificational copular constructions. *It*-clefts, *wh*-clefts, *th*-clefts, *all*-clefts and certain noncleft copular sentences all inherit properties from a more general, schematic, specificational copular construction. But what is a specificational copular construction? And where does specificational meaning come from? The answers to these questions are not obvious and a number of different analyses have been proposed in the literature. In order to understand the larger schema, or category, of copular constructions, I examine the nature

of these sentences in Chapter 3. I argue that in specificational copular sentences, specificational meaning results from a class-membership predication relation associated with the concept of definiteness. I show that this type of analysis is able to account for data which has eluded alternative approaches to specificational sentences.

On this account, many of the *it*-cleft's properties are shown to be motivated; that is, they are simply inherited from the more basic specificational copular construction. Following the extraposition accounts of Jespersen (1927) and Bolinger (1972), I analyse the sentence-final clause as a restrictive relative, modifying the initial *it*. In particular, I argue in Chapter 4 that *it* and the relative clause together operate like a discontinuous definite description (see also Hedberg 1990, 2000; Percus 1997; Han and Hedberg 2008). Since definite descriptions exhibit existential presuppositions and are associated with exhaustiveness (or inclusiveness), this analysis explains why these properties are found in *it*-clefts too. It also reduces the number of semantically "dummy" elements. For instance, on this account, the initial *it* is not expletive and is instead shown to perform an important function. Where my account advances the current literature is in providing a reason as to *why* definite descriptions are a fundamental component of *it*-clefts (and other specificational constructions). Furthermore, as I explain in Chapter 5, this particular analysis allows for a more straightforward account of the relationship between specificational and predicational/proverbial *it*-clefts.

Alternative constructional accounts of *it*-clefts have been put forward by Lambrecht (2001) and Davidse (2000). However, while both authors extend their accounts to other types of cleft sentence, neither makes use of a system of inheritance. Consequently, they treat the *it*-cleft as a highly idiosyncratic construction. For example, Lambrecht (2001) views *it*-clefts in relation to simple noncopular sentences, and so presents an expletive analysis. Thus, while his account is able to accommodate the *it*-cleft's unusual properties (through invoking the concept of the construction), it cannot identify *how* they come about. Davidse (2000), on the other hand, analyses the *it*-cleft as a highly complex structure involving two clauses (one of which is unique to cleft constructions) which enter into different semantic relationships with the postcopular element. Again, it is not clear how this structure is motivated by the language system.

The constructional approach outlined here is therefore, in some ways, an improvement on those of Lambrecht (2001) and Davidse (2000) since it makes full use of the tools employed in construction grammar for making *generalizations*. By examining *it*-clefts in relation to the taxonomy of

specificational sentences and exploiting an appropriate inheritance hierarchy, the motivation for this construction is maximized. Only after examining *it*-clefts in relation to the rest of the grammar are the exceptional or truly construction-specific characteristics isolated. As Goldberg (2003: 118) observes, “a given construction often shares a great deal with other constructions that exist in a language; only certain aspects of its form and function are unaccounted for by other constructions”.

#### 4. A diachronic approach to *it*-clefts

We have seen then that construction grammar tolerates (although nevertheless seeks to limit) idiosyncrasies in the language system. However, ideally, even exceptional properties should be provided with an explanation of some sort. According to Goldberg (2003: 121), in such cases, motivation can be provided by factors external to the grammar. In this section, I ask whether historical evidence can provide motivation for, and so account for, some of the construction-specific properties of the *it*-cleft.

From my synchronic analysis of the *it*-cleft as a type of specificational copular construction, certain structural aspects remain a puzzle, such as the modification of *it* by a restrictive relative clause and the extraposition of the relative clause. By examining the *it*-cleft's structural idiosyncrasies in relation to the language system of earlier periods of English, I show in Chapter 6 that although these properties are no longer motivated by the language system, they are likely to have been inherited from formally related constructions existing at earlier periods of the language. In this way, the *it*-cleft shows how the retention or *entrenchment* of once-motivated form-function pairings can lead to construction-specific properties which are no longer productive in other areas of the grammar.

For most types of *it*-cleft then, their seemingly idiosyncratic properties become much less mysterious when examined in relation to the grammar of earlier periods of English. However, there are subtypes of *it*-cleft which exhibit properties that cannot be attributed to inheritance at any period of the language. In Section 1, I introduced two varieties of *it*-cleft which are sometimes treated as separate constructions from the *it*-cleft proper: those with non-nominal foci and those with new information in the sentence-final clause. The particular range of elements found to occur in the postcopular position of the *it*-cleft is not shared by other specificational copular constructions. Likewise, the *it*-cleft seems to be the only kind of specificational sentence to express brand-new information in the presuppositional clause

(see Prince 1978; Collins 1991a). As a result, the range of non-nominal foci and the ability to express hearer-new information are properties which are not inherited from the wider specificational construction (see Chapter 5). This begs the question, where did these more idiosyncratic properties come from?

In Chapter 7, I conduct a diachronic investigation, using data (from four historical English parsed corpora) which spans from Old English to Modern English. I find that the *it*-cleft occurs with an increasingly wide range of foci and appears in a greater variety of discourse contexts over time. These idiosyncratic properties of the *it*-cleft are therefore shown to be an outcome of the construction's historical development. In Chapter 8, I ask how and why the *it*-cleft construction has developed in this particular way. I interpret the changes to the function and use of the specificational *it*-cleft as an example of *schematization*. On this account, novel instances are formed by extension from the prototype, overriding inheritance from more basic patterns. As these new types of instance become more conventional, the *it*-cleft, in turn, becomes a more abstract and schematic construction. The *it*-cleft's construction-specific development is therefore shaped by its prototype, which differs in subtle ways from that of other specificational copular constructions. I exemplify this with a short comparison of *it*-clefts and *wh*-clefts.

The historical evidence therefore demonstrates that the *it*-cleft was once fully motivated by inheritance from the language system. Over time, the construction has acquired a range of idiosyncratic properties via conventional pathways of change – both fossilization and schematization. Therefore, while these construction-specific characteristics are not inherited from more general patterns of correspondence, they are nevertheless shown to be motivated by general principles of language change. A more comprehensive overview of usage-based approaches to constructional change is provided in Chapter 2.

With the inclusion of a substantial diachronic component, the present study is able to contribute to the somewhat limited literature on the history of the English *it*-cleft. Until very recently, Ball's (1991, 1994a) work has dominated the literature on this topic (see Filppula 2009; C. Johansson 2008; Los 2009; Los and Komen forthcoming; Patten 2010; Pérez-Guerra 1999, forthcoming for more recent contributions). However, in Chapters 6, 7 and 8, I present a very different account of the specificational *it*-cleft's origin and subsequent diachronic development. The approach outlined here is unique in that it reviews the historical data in light of an extraposition (from-NP) account of *it*-clefts. As I explain, the historical evidence actually

provides considerable support for this particular synchronic approach. Finally, the present study contributes to the recent literature on diachronic construction grammar, in that it provides an example of how more complex (multi-word and partially schematic) constructions undergo change.

## 5. Methodology

This book examines the English *it*-cleft from both a synchronic and diachronic perspective. In the synchronic part, I rely largely on examples that are either invented or taken from the literature. I have chosen to exemplify my discussion in this way for several reasons. First, I am engaging with a literature where the use of invented examples is common practice. Second, since the issues surrounding the data are often complex, I have made an effort to keep examples brief and to choose examples that highlight the relevant features without requiring unnecessary explication. Furthermore, in these chapters, the focus of my discussion is on the prototypical *it*-cleft subtype. As a result, I am not always interested in detailing qualitative differences between individual instances.

However, where my concern is to highlight variation in the *it*-cleft's structure or to demonstrate aspects of use, I provide attested examples from the British component of the International Corpus of English (ICE-GB), extracted using the corpus utility program ICECUP 3.1. A search on the CLEFTIT annotation produces an output of 430 hits within 422 text units. From this data, I discounted incomplete tokens, tag questions and truncated clefts (see Chapter 4, Section 1.1 for an analysis of such structures), as well as a handful of instances which appear to have been mistakenly tagged as *it*-clefts.<sup>4</sup> This amounts to a data-set of 404 tokens, from which frequency counts and proportions are measured. For more comprehensive quantitative studies of present-day *it*-clefts in the ICE-GB see Gómez-González (2004, 2007), Hasselgård (2004), and Nelson (1997).

---

4. For example, in the utterance below, *work* is not specified as the thing she most enjoys; instead, *working full-time as a nursing auxiliary* (referred to by *it*) is described as enjoyable work. This is not an *it*-cleft, despite its mark-up in the ICE-GB.

(i) *At the time of the accident she was thirty-nine years old, married, with children, and working full-time as a nursing auxiliary at the Pembury Hospital near Tunbridge Wells. It was work which she much enjoyed a and to which she was fully committed.* (S2A-062 008, Legal Presentation)



The diachronic investigation makes use of data from four independent, yet related, historical English corpora: the *York-Toronto-Helsinki Corpus of Old English Prose* (YCOE), the *Penn-Helsinki Parsed Corpus of Middle English*, second edition (PPCME2), the *Penn-Helsinki Parsed Corpus of Early Modern English* (PPCEME) and the *Penn Parsed Corpus of Modern British English* (PPCMBE). These four corpora form part of the same series of syntactically annotated historical English corpora from the University of Pennsylvania and the University of York. As a result, they all use the same system of syntactic annotation and are accessed by the same search engine: CorpusSearch2. Together, the corpora comprise over 5 million words of running text and span the entire history of British English up until 1914. This amounts to over 500 clear *it*-cleft tokens, dating from the mid-tenth to the early-twentieth century. I discuss the method of extraction and selection in detail in Chapter 7, Section 2.

The study undertaken here is *corpus-based* rather than *corpus-driven*; that is, my cleft analysis informs my corpus investigation and, in turn, the empirical evidence serves to support my constructional account of *it*-clefts. Unlike in corpus-driven studies then, I approach the data with already established hypotheses. One reason for adopting this approach is that *it*-clefts are notoriously difficult to identify and separate from superficially similar but structurally distinct sentence-types, such as those containing extraposed subject clauses (see Calude 2008a; Haugland 1993). Furthermore, since there are different ways of analysing *it*-clefts, linguists may differ over which examples count as clefts. It therefore seems preferable to have a clear understanding of the criteria used to extract the relevant data, so that at least the approach is consistent. It also means that I can address the issue of how well my theory accounts for the data.<sup>5</sup>

---

5. See Tognini-Bonelli (2001) for an outline of the differences between corpus-based and corpus-driven approaches to corpus study. While I argue for a corpus-based approach in this instance, Tognini-Bonelli (2001: Chapter 5) discusses the merits of a corpus-driven approach.

## Chapter 2

# A model of language structure and language change

### 1. Some basic assumptions

This book examines the synchronic structure and diachronic development of the English *it*-cleft within the framework of construction grammar. In particular, I adopt the fundamental principles of *usage-based* theories of construction grammar, and make use of concepts from Cognitive Grammar (Langacker 1987, 1991), Cognitive Construction Grammar (Lakoff 1987; Goldberg 1995, 2006) and Radical Construction Grammar (Croft 2001). In this chapter, I outline some of the basic claims underlying the present study and introduce some of the machinery that I make use of in my analysis of cleft and copular sentences. In Section 2, I explain how grammatical knowledge is organized on a usage-based, constructional theory and discuss the rationale behind this model of language structure. In Section 3, I show how, together with usage-based assumptions, the conceptualization of language as a hierarchical network makes a number of predictions regarding the diachronic development of constructions which are compatible with well-attested pathways of change, such as grammaticalization. In Section 4, I go on to provide a (very brief) indication of how I make use of these various concepts in my analysis of the English *it*-cleft construction.

It should be noted that not all versions of construction grammar adopt this same set of principles. The Construction Grammar of Kay and Fillmore (1999), for example, is not usage-based. As a result, it differs from other construction grammars in that it adopts a complete (rather than a default) mode of inheritance, which licenses a non-redundant system of grammar. As Goldberg (2006: 214, 216) notes, this version of construction grammar has developed somewhat separately, in that it more closely resembles the formalist theory Head-driven Phrase Structure Grammar (HPSG). In what follows, I indicate where this version of construction grammar differs from the other relevant theories and, as a result, from the fundamental assumptions of the present study.

## 2. A constructional model of language structure

Construction grammar was developed with a view to providing full and explanatory accounts of specialized linguistic patterns. Such configurations are problematic for the componential model of grammatical knowledge that underlies generative theories of grammar (see Croft 2007). On a componential model, each type of linguistic knowledge (syntax, semantics and so on) makes up a separate component, with aspects of meaning and form being mapped on to one another through general linking rules. The only idiosyncratic and item-specific mappings between these components are found in lexical items, which are both arbitrary and conventional. On this model then, complex structures are built out of discrete, atomic elements in accordance with the combinatorial rules specific to each component. This reductionist perspective lends itself to the stronger hypothesis that constructions (more complex grammatical structures) are purely epiphenomenal. As Chomsky (2000: 8) says, “grammatical constructions are taken to be taxonomic artifacts, useful for informal description perhaps but with no theoretical standing”.

The componential model works well for regular syntactic expressions, since, once we know the meanings typically associated with each of the lexical items, we can determine the meaning of the construction as a whole. However, the meaning and function of more specialized linguistic patterns cannot be determined from the general rules of semantic interpretation for their constituents as they exist outside of the particular construction; that is, the meaning of specialized linguistic patterns is *conventional* in the sense that it must be learned. Such structures are also sometimes *noncompositional*; that is, they express meanings which cannot be entirely broken down into components and attributed to their individual formal elements.<sup>1</sup> These, less regular expressions are therefore problematic for a strictly componential model of language structure, since aspects of their meaning and/or use cannot be generated from the application of highly general linking rules.

Such phenomena represent strong evidence for the need to recognize the *construction* as having an independent theoretical status. On this model of grammar, constructions are symbolic pairings of form and meaning, much like lexical items. The syntactic elements and semantic components that are particular to each complex construction are related via symbolic links that

---

1. See Nunberg, Wasow, and Sag (1994) for a detailed discussion of the differences between the concepts of *conventionality* and *noncompositionality*.

are internal to that construction. This allows the representation of meanings that are at once compositional (attributable to individual elements of form) and conventional (construction-specific or irregular). In addition to housing these smaller pairings of form and meaning, complex constructions contain a further symbolic link which relates the entirety of the construction's form to the construction's conventional meaning. This allows the representation of noncompositional meanings, which are associated with the construction as a whole, but cannot be attributed to any of its individual parts.

A model of grammar which recognizes the construction as a theoretical object is therefore able to handle the existence of idiosyncratic grammatical information. However, construction grammar takes this a step further, proposing that, in fact, all grammatical knowledge, including the predictable or regular patterns of the language, should be represented as constructions. As Fillmore, Kay, and O'Connor (1988: 534) suggest, the machinery needed to describe more peripheral constructions can be "generalized to more familiar structures".

Indeed, this follows directly from the usage-based assumptions of most construction grammars. On a usage-based model, humans are not innately programmed with linguistic knowledge; instead, all of language (and not just the periphery) is learned inductively from the input, or rather, from the speaker's linguistic experiences. It follows from this that all grammatical knowledge (both specialized linguistic patterns and broad generalizations) should be given a uniform representation. Therefore, in usage-based constructional theories, the entire language system is made up of constructions (form-meaning pairs). As long as a linguistic pattern is sufficiently frequent that the speaker is likely to induce an abstract mental schema, then it may be stored as a symbolic unit. Thus, on a usage-based framework, constructions are simply conventionalized chunks of linguistic knowledge (Goldberg 2006: 05).

These constructions form a structured inventory which makes up the speaker's knowledge of the language. This inventory is represented as a taxonomic network of constructions with each construction constituting a separate node (Croft and Cruse 2004: 262). The network is hierarchical, showing that some constructions are more basic or general than others. Lower-level constructions inherit attributes from higher-level constructions. That is to say, more specialized and substantive (lexically filled) constructs are instances of more general and schematic (lexically open) constructions. Usage-based construction grammars adopt the *default* mode of inheritance. On the *complete* inheritance model, an inheriting construction inherits all of the information which is specific to a dominating construction. This means