It-clefts in the meta-informative structure of the utterance in Modern and Present-day English

Chapter - January 2013
DOI: 10.1075/slcs.143.07mar

CITATIONS
2

READS
74

2 authors:

Anaelina Martinez-Insua
University of Vigo
10 PUBLICATIONS 45 CITATIONS

Javier Perez Guerra
University of Vigo
38 PUBLICATIONS 56 CITATIONS

Some of the authors of this publication are also working on these related projects:

Microvariation in English View project

English language and its accreditation in HE View project
1. Introduction

In the literature the label ‘cleft’, as suggested in Jespersen (1909-49, 1937), is commonly used with structures such as (1) to (5), which, irrespective of their syntactic organisation, are so called “because of the apparent dismemberment of a single sentence entailed in their derivation” (Delahunty 1982: 5):

(1) What this paper describes is the cleft construction.
(2) The cleft construction is what this paper describes.
(3) It is the cleft construction that this paper describes.
(4) That is the cleft construction (that) this paper describes.
(5) They are real researchers that tackled this issue.

The construction in (1) is called basic pseudo-cleft or basic wh-cleft, (2) illustrates the reversed or inverted pseudo- or wh-cleft, the structure in (3) is known as cleft or it-cleft, the one in (4) is known as th-cleft, and the example in (5) illustrates a pronominal cleft. The treatment of the constructions in (1) to (5) as clefts is grounded not only on the formal similarities between the constructions but also on semantic, pragmatic and communicative considerations. To illustrate this, it is commonly assumed that it- and pseudo-clefts share the same truth-conditions or, as Prince (1978: 884) puts it, the “same objective information”, which implies that they are semantically identical. In this vein, Akmajian (1970: 149) maintains that it-clefts and pseudo-clefts “are synonymous, share the same presuppositions, answer the same questions, and in general they can be used interchangeably”. Declerck (1988: 209), however, claims that “there are numerous pragmatic factors that may induce the speaker to prefer one type of cleft to another in a particular context” [our italics] (see Traugott 2008: section 3 for semantic and informative differences between it- and pseudo-clefts).

It-clefts deserve a specific place in an account of English constructions because of at least two specific characteristics: the expletive nature of its introducer and the difficulty of accounting for the postverbal subordinate clause by resorting to grammatical concepts such as relativisation, detached postmodification, or even some kind of right dislocation. First, on strictly syntactic grounds, basic pseudo-clefts can be analysed as (unmarked) sentences in which a relative clause,³ be it headless or headed,
functions within the subject of the sentence. Second, with the same argument in mind, a reversed pseudo-cleft can be claimed to contain a relative clause fulfilling the syntactic role of subject predicative. Third, in *th-* and pronominal clefts the introductory (non-expletive) referential items together with the coreferring postverbal subject predicatives can constitute a complete clause, and the postverbal subordinate clauses can be said to be linked either to the subjects or to the subject predicative and thus constitute an example of postmodification (or even some sort of right dislocation).

This chapter focuses on the *it*-cleft construction and aims at, first, profiling it from the perspective adopted in Meta-Informative Centering Theory (hereafter, MIC) and, second, describing the major tendencies yielded by a diachronic corpus-based analysis. The English *it*-cleft will be presented here as a device of focalisation used for establishing a meta-informative contrast with the second part of the utterance, making it possible, in a Strict Word Order (SWO) language as English, to put this focalised constituent at the front of the utterance (thus contradicting the neutral order: given before-new). The study is organised as follows: in Section 2 we outline the barebones of MIC. Section 3 is devoted to the MIC-conformant meta-informative analysis of *it*-clefts (see Section 2 below for the characterisation of meta-information as the sequential ordering of information). In Section 4 we focus on the semantic and informative characteristics of *it*-clefts. In Section 5 we account for the data on which our corpus analysis is based and analyse the degree of variation undergone by the construction in Modern and Present-day English. Finally, Section 6 is devoted to the summary of the findings and the concluding remarks.

### 2. The theoretical framework: the MIC theory

Following the tradition of the Prague theory of Functional Sentence Perspective (Daneš 1974 or Firbas 1992, among others), the MIC theory contains also elements (anaphor vs. cataphor) which correspond to the backward and forward looking centres in the American Centering Theory (Grosz and Sidner 1986). Even if both theories were created separately, as Włodarczyk and Włodarczyk (personal communication) state, MIC easily integrates within the European framework a number of aspects and results of the American Centering Theory, there being several points of coincidence between both theories. The American framework may be understood as “an account of one aspect of discourse processing, local discourse structure, that makes specific claims about both processing complexity and discourse anaphora” (Walker, Joshi and Prince 1998: preface). Centering is “a model of the conversants’ center of attention in discourse that is concerned with the relationship of attentional state, inferential complexity, and the form of referring expressions” (Walker, Joshi and Prince 1998: 1). Such an approach to language aims to model discourse processing factors that might explain the differences in the degree(s) of coherence perceived by addressees when they receive and interpret messages.

The origins of the MIC theory go back to 1999, when Włodarczyk (see Włodarczyk and Włodarczyk 1999) used the concept ‘centre d’intérêt’ for the first time. Since then, this French term has been replaced by that of ‘centre of attention’ under the influence of the Grosz-Sidner centering-theoretical framework, on the one hand, and it underwent a few important revisions and reassessments in the subsequent...
papers published in English, on the other hand. In Grosz, Joshi and Weinstein (1995), centres of attention are defined at the text level: one constituent of an utterance is treated as a ‘forward’ or ‘backward looking centre’ in order to maintain the cohesive flow of information from one utterance to its successor. Forward and backward looking centres make it possible to give an account of the relations which bind utterances together into a coherent text. MIC assumes that, because of the linear order of speech sounds in human languages, no judgment may be uttered without selecting at least one centre of attention, and centering is regarded as a structuring operation at the levels of both text and utterance.

In the MIC model, the main focus is on ‘meta-information’, which is the label for the sequential ordering of ‘information’, that is, for the sequential ordering of the content of linguistic utterances. In this sense, the meta-informative level “is necessary in order to achieve the ordering of non-linear mental representations as texts (sequences of linguistic utterances)” (Włodarczyk and Włodarczyk, this volume). In order to communicate such mental non-linear situations, users of the language select the element(s) they will be treating as the centre(s) of attention in their messages and predicate something about it/them. Once the centres of attention (henceforth, CAs) have been selected from the referents or ‘anchors’ of the situation, the user establishes a hierarchy and turns the primary (most important) one into the subject of the utterance, while the secondary centre(s) is/are turned into the object(s). According to MIC, in English the primary CA is global and preverbal (represented in upper left nodes of phrase-markers), whereas the secondary ones are local (located in lower right nodes within the tree structure). The specific labels for one and the others are ‘Global CA’ and ‘Local CAs’, respectively.

Centering and predication are crucial notions within the MIC Theory. While centering involves selecting and highlighting one entity among others, predication implies saying something about the entity selected by the speaker as the global CA of the message. Thus, predication takes place when speakers produce linguistic expressions in which “some distinguished segments are highlighted as centres of attention” (Włodarczyk and Włodarczyk, this volume). From such a perspective, “no utterance can be formed in a natural language without choosing a CA and assigning to it a meta-informative Old or New status” (Włodarczyk and Włodarczyk, this volume). This implies that the meta-informative status of the CAs, as treated in this framework, concerns their oldness or newness. In other words, the meta-informative status of information concerns the way in which it is treated in the communicative event, the way speakers introduce it into their discourse. Ultimately, the old/new meta-informative status originates in the discourse strategies (partly fixed by syntactic rules) established and employed by the speaker (Włodarczyk and Włodarczyk, this volume).

MIC has already been presented elsewhere as an appropriate framework for the analysis and study of English thematically marked sentences, especially those where an Anonymous Subject (henceforth, AS) occupies the Global CA position that another semantically full constituent would occupy in the unmarked version of the sentence (see, Martínez-Insua 2011, Martínez-Insua and Pérez-Guerra 2012). This chapter aims at highlighting the appropriateness of this theoretical framework for the description of clefting as a focusing meta-informative strategy.

3. Description of it-clefts

This section tries to characterise it-clefts against the background of the MIC approach
to language, bearing in mind MIC’s distinction between Global and Local CAs. In Section 3.1 we describe the meta-informative organisation of the \emph{it}-cleft sentence and pay special attention to the distribution of the CAs. In Section 3.2 we describe the structural and grammatical features of the construction.

### 3.1. A MIC-compliant representation of \emph{it}-clefts

From a representative point of view, \emph{it}-clefts such as (3) above, repeated here for convenience, might be described in two ways:

(3) It is the cleft construction that this paper describes.

(i) Sentences where the expletive and semantically empty AS \emph{it} fills the Global CA slot, while a semantically full constituent functions as Local CA and is followed by a clause that constitutes its background. The meta-informative layout of the \emph{it}-cleft would then be described as in (6):

\begin{align*}
\text{(6)} & \quad \text{[}\text{It}_\text{AS}\text{]} \text{Global CA [be] [X] Local CA (FOCUS)} \text{ [introducer + [clause ...n.p.i...]]BACKGROUND}
\end{align*}

where the rightmost clause is the background contrasting with the focus expressed in the cleft part of the utterance. This clause contains a null pointer \([n.p.i.\text{]}\) left by the focalisation of one of its constituents \([X]\) which acts as the Local CA.

(ii) Sentences where the expletive and semantically empty AS \emph{it} fills the Global CA slot, while the semantically full Local CA is a clausal constituent whose unmarked meta-informative layout has been altered. As a result of such word-order alteration in the (main) clause, one of the components of the (subordinate) clause has been focalised and brought to initial position, leaving a null pointer behind. The meta-informative layout of the \emph{it}-cleft is then described as in (6\textsuperscript{'}):

\begin{align*}
\text{(6\textsuperscript{'})} & \quad \text{[}\text{It}_\text{AS}\text{]} \text{Global CA [be] [[X] [introducer + [clause ...n.p.i...]]] Local CA}
\end{align*}

where the Local CA is made up of a (rightmost) clause from which one of the constituents has been topicalised and placed before the introducer \([X]\), leaving a null pointer behind \([n.p.i.\text{]}\).

The first one of these two characterisations, where the rightmost clause is taken as the background of the utterance, seems to be the most appropriate one, as we see it. The role of the rightmost clause as a second-level CA is in line with the role of constituents such as \emph{a book}, in \emph{Mary gave John a book, to John in Mary gave a book to John, or with his key in John opened the door with his key}. Knowing that MIC does not allow for the presence of “neither double Local nor two-members single Local CAs” (Włodarczyk and Włodarczyk, personal communication), the rightmost clause has to be interpreted as a second-level CA (Włodarczyk and Włodarczyk 2006: 8). In other

\footnote{From a contemporary synchronic perspective, the expletive nature of the so-called ‘introducer’ \((\emph{it})\) is justified by the fact that agreement does not hold between the two main constituents of the copular construction \((\emph{it} as the grammatical subject and the focus or ‘X’), as sin \emph{It is my parents that I never trust} (see, in this respect, Pérez-Guerra 1999: Chapter 4, section 3.2). By using historical data, a number of scholars maintain that \emph{it} keeps referential (cataphoric) status (see Pérez-Guerra 1999: 170-171).}
words, according to the MIC theory, which bases its premises on actual surface structure, we argue that a link construction such as the *it*-cleft construction (see Section 4.1 in this respect) consists of the linking of two second-level (focus and background) CAs by means of a syntactic design governed by an almost semantically bleached linking verbal operator, namely the expletive *it*.5

3.2. A structural and grammatical description of *it*-clefts

For a construction to be considered as *it*-cleft, the following conditions should be met:6

(i) The Local CA ([X] in the representation in (6) above) and the null pointer are coreferring. For this condition to hold, X must materialise, in Givón’s (1984: 731) words, an entity; otherwise referentiality would be out of the question.7 The characterisation of X as an entity-denoting constituent excludes sentences like (7), adapted from Declerck (1988), from the class of clefts:

(7) It is always expensive what Cambridge University Press sells.

Example (7) illustrates the extraposition of the Subject *what Cambridge University Press sells* and the filling of the empty Global CA slot with the AS *it*.

(ii) The ‘introducer’ in the rightmost clause can be one of the following: Ø (Visser’s 1970: Chapter I apo koinou constructions), *that*, *who* or *which*. In the literature, other expressions – for instance, *what*, *when* (see Declerck 1997) or *where* – are frequently included in the set of possible introducers of the rightmost clauses of *it*-clefts. From a syntactic perspective, such (apparent) clefts with *wh*-forms are not syntactically different from extraposed headless relative clauses or pseudo-clefts, as shown in (8) and (9) vs (10) (similarly, Delahunty 1982: 268ff and Ball 1994a: 181):

(8) It is phrase-markers what I drew on the blackboard.
What I drew on the blackboard is phrase-markers.

(9) It was before the 1957-58 tour of South Africa, when Bagenal said half-jokingly before some of the team: ‘I wish I was going with you so-and-so’s.’ (Findlay: 210)

(10) It is phrase-markers that I drew on the blackboard.
*That I drew on the blackboard is phrase-markers.

(iii) The null pointer within the rightmost clause may fulfil a large array of functions: Global CA (in (11)), Local CA (in (3) above and in (12)), (non-sentence) adverbial (in (13) and (14)), adverbial complement or obligatory adverbial (in (15)), prepositional

---

5 This proposal is, in fact, in keeping with syntactic solutions which rely on the fact that the syntactic relation holding between the focus and the clause is subordination (see Section 3.2 in this respect).

6 Notice the markedly syntactic nature of the criteria and terminology employed in the following list.

7 Example (i) will only be acceptable if *a genius* is understood as an entity-denoting ‘nominal,’ that is, in an identifying (probably contrastive) way:
(i) It’s a genius that he is.
If *a genius* is interpreted predicatively, the sentence becomes unacceptable. Notice that only under very special circumstances can predicative nominals undergo clefting (see (iii) in the main text).
complement of a verb (as in (16)), of an adjective ((17)) or of a noun ((18)),
complement of a preposition (in (19)), predicative complement of the subject or of the
object in very special environments (examples (20) and (21)),9 taken from Declerck
1988), predicative10 (in (22)), or particle of a phrasal verb (in (23)).

(11) It is a gap that occurs in initial position.
(12) It is {#Ø11/ to} me that he dedicated the book.
(13) It was with much attention that I checked the last proofs of the article.
(14) It was only reluctantly that he agreed to help me with them at all.
(15) It is to Boston that she went.
(16) It is to my article that she was referring.
(17) It was about that Minister that the President was angry.
(18) It was of Syntactic Structures that he was the writer.
(19) That was the doctor I was speaking to.
(20) It’s pretty that my mother-in-law is, more than anything else.
(21) It’s a teacher that he is, not a butcher!
(22) ?It was teach English that he did at that time.
(23) It was (obviously) off that I turned the computer when I suspected it was
infected.

Disjuncts and conjuncts cannot occur in the Local CA position of clefts, as shown in
eamples (24) and (25), respectively:

(24) *It was frankly that he didn’t expect to discover the philosopher’s stone in his
study.
(25) *It is nonetheless that conjuncts can be cleft in languages other than English.

(iv) The null pointer/initiator belongs to one of the following categories:
- NP, as in (3) above,
- PP, as in (13), (15), (16), (17) or (18),
- Adverb Phrase, as in (14),
- Particle of a phrasal verb, as in (23) above.
- The Local CA in a cleft does not normally belong to the category clause,12 VP (see


9 In examples (20) and (21), pretty and a teacher are clearly contrastive and thus the meaning of the
construction is identifying, as required by condition (i) in the main text, otherwise the clefting of a
predicative complement would be disallowed.

10 Inflected verbs cannot be cleft (see (i) below) and (uninflected) VPs are accepted as Local CAs in clefts
only in informal Irish English (as in (22) in the main text; see Givón 1984: 731 for a detailed account):
(i) *It’s wore that John e a white suit at the lecture.

11 Huddleston (1984: 460) regards non-prepositional indirect objects in X-position of clefts like * It was
Liz that I bought the flowers as ungrammatical. Delahunty (1982: 87) only accepts them if they include
stranded prepositions, as in It was Liz that I bought the flowers for. This example should be included in
the category of complement of a preposition (see (iii) in the main text).

12 The acceptability of the clefts in (i), (ii) and (27) in the main text, including, respectively, an -ing, a
that- and an infinitive clause in X-position, indicates that the restriction on clausal clefting is maybe due
not to the theoretical organisation of the construction but to performance preferences:
(i) It was changing the word-processor that I most
hated.
footnote 10) or (non-contrastive) AP (in (26) below; see, however, (7) and (20) for acceptable APs in Local CA position):

(26) *It’s very unhappy that PhD students are.

(v) The rightmost clause can be either finite (supposedly a *that*-clause), as in (3) above, or nonfinite (-*ing* or infinitive clause), like the one in (27):

(27) [Within the United States,] it is Robinson to appear like a Jones. (Gibb:115)

(vi) The construction is introduced by an AS, materialised by dummy or expletive *it*. In our opinion, precisely because the anonymous nature of *it* is taken as one of the identifying features of the *it*-cleft construction,13 examples like (4) or (5) cannot be included in the label ‘cleft’. By contrast, they will be given a quasi-right-dislocation analysis, in which the rightmost constituent functions as a displaced postmodifier of the pronominal Global CA. This view receives support from the fact that, whereas the example in (4) is able to undergo reversion, which is a defining feature of identifying utterances, clefts cannot reverse their Global and Local CAs around *be* ((3’)):

(4’) The cleft construction (that) this paper describes is that.
(3’) *The cleft construction that this paper describes is it.

4. Informative aspects of *it*-cleft sentences

Despite their markedness, *it*-clefts, like other thematic systems (e.g. most *there*-sentences), normally have the truth-conditions of their unmarked counterparts, even though minor informative differences between both versions can be detected.14 In this section, devoted to the study of certain semantic and pragmatic aspects of the constructions under analysis, we pay attention, first, to the identifying nature of *it*-clefts (Section 4.1) and, second, to the consequences which the meta-informative layout of an *it*-cleft has for the informative design of the sentence (Section 4.2).

4.1. Clefts as identifying structures

Informatively speaking, and being a subtype of *be*-structures, clefts must be classified

(ii) ?It was that I preferred a four-section article that I told him. [special intonation is required]

Delahunty (1982, 1984) justifies the acceptability of *that*-clauses by contending that they are NPs. In fact, only *that*-clauses which are equivalent to entity-NPs, namely factive *that*-clauses, are eligible for clefting.

*It was that he was going to finish the book in two years that he thought is ungrammatical because it includes a non-factive cleft clause in X-position.

13 Visser (1970: 50) claims that examples of the type *they are my brothers that...* were clefts from the sixteenth to the nineteenth centuries.

14 Atlas and Levison (1981: 2), when they deal with the cleft in (i) and the noncleft in (ii), point out that “[(i)] exhibits presuppositional behaviour that [(ii)] does not, namely, the preservation of the inference to [(iii)] under denial and questioning of [(i)]” (see Halvorsen 1978: 18-19 for a similar account):

(i) It was John that Mary kissed.
(ii) Mary kissed John.
(iii) ‘Mary kissed someone’
in one of the two groups of link constructions, namely, identifying (or specifying\textsuperscript{15}) and attributive (or predicational) (see Enkvist 1979 or Declerck 1988 for the semantic and formal differences between both classes; see Halliday and Matthiessen 2004 or Thompson 2004 for the functional characterisation of both these processes). In a word, these types of link constructions can be described simply by saying that identifying \textit{be}-sentences specify a value for a variable\textsuperscript{16} (as suggested by Declerck, the semantic scheme of identifying sentences is ‘\(x=y\)’, interpretable as ‘assign the value \(y\) to \(x\)’, the value occupying the Local CA position and the variable occurring in the Global CA position), while attributive structures predicate a property of their Local CAs in a non-identifying way. Most clefts seem to illustrate the former type, whereas pseudo-cLEFTs may have both readings, that is, identifying and attributive. This leads us to disallow examples like (7) above, repeated here for convenience, as members of the class of cLEFTs:

(7) It is always expensive what Cambridge University Press sells.

It is obvious that (7) has not got an identifying meaning, since \textit{expensive} is not a value but a predication.

Let us now deal with the cLEFTs in (28) to (31), which are taken as attributive in the literature:

(28) It is a poor heart that never rejoices.

(29) It’s a wise child that knows its own father.

(30) It would be a brave man who marries her, won’t it? [taken from Declerck (1988)]

(31) Was it an interesting meeting you went to last night?

Regarding proverbial examples like (28) and (29), and less proverbial ones such as (30), whose reading is clearly attributive, we agree with Declerck that these examples are not cLEFTs but sentences involving some sort of relative clauses. Out of the several reasons which Declerck adduces in favour of his proposal, we will, just for the record, mention two. On the one hand, these examples cannot be uncLEFT ((28’) to (30’)); on the other, the selection of relative pronouns is closer to clear relative clauses than to the usual pattern of cLEFTs ((28") and (29") vs. (32)):

(28’) A poor heart never rejoices. [the meaning is different from that of (28)]

(29’) A wise child knows its own father. [the meaning is different from that of (29)]

(30’) A brave man would marry her. [the meaning is absolutely different from that of (30)]

(28") It is a poor heart that/which/who never rejoices.

(30") It would be a brave man that/who marries her, won’t it?

(32) It will be John that/#who will marry her, won’t he?

\textsuperscript{15} Declerck (1988) distinguishes between ‘specificationally-identifying’ and ‘descriptionally-identifying’ \textit{be}-structures. Since we are not aware of any instance of descriptionally-identifying cLEFT, we will not make use of Declerck’s distinction.

\textsuperscript{16} In Declerck’s (1988: 228) words, cLEFTs “suggest that an answer is being given to a question that is implicitly present in the mind of the hearer”.

As far as (31) above is concerned, Declerck concludes that, owing to its hybrid nature between the identifying and the attributive subclasses, this example illustrates a new type of *it-be* construction. Most of the features cited by Declerck evincing the attributive reading of the construction are semantic, whereas those signalling the identifying interpretation are grammatical or syntactic. From both a grammatical and a semantic point of view, the (supposedly) new class represented by (31) easily finds its way in the general class of (identifying) clefts. As we see it, Declerck’s judgements are influenced by the obvious attributive nature of the adjective *interesting* premodifying X in (31), which he extends up to the whole construction. Declercck mentions the fact that the Local CAs (X-segments) in examples like (31) can be quantified either by inserting *no* or by grading the adjective, and adds that this is a consequence of the attributive character of the sentence. Adjectives, both in attributive and in identifying contexts, can be expanded by constituents which are modifiers typical of attributive items and not of the contexts in which these adjectives are involved. In other words, the possibility of either inserting *no* before the adjective or of grading *interesting* in (31) simply demonstrates the attributive meaning of *interesting*, not of the whole example. Moreover, the supposedly ‘odd’ example (33) given by Declerck is perfectly acceptable in English, at least by the native speakers consulted:

(33) It is John that/who is going to lead and an interesting subject that is going to be discussed.

If our informants are correct in their predictions, the coordination of the clear cleft *It is John who is going to lead* and *(It is) an interesting subject that is going to be discussed* in (33) corroborates the non-syntactical semantic equivalence of the structure, or, in other words, the inclusion of (31) in the class of (identifying) clefts.

Declerck also compares (31) with other structures comprising verbs other than *be* which are clearly attributive:

(34) [A couple of weeks ago it seemed clear which candidate deserved voting for, but now many people think that] it has become a difficult choice they have to make.
(35) It used to be a really good painter who painted their portraits.

We agree with the attributive nature of these examples, which is demonstrated by their impossibility to be formulated as non-clefts, as shown in (34’) and (35’):

(34’) They have to make a difficult choice. [not equivalent to (34)]
(35’) A really good painter {used to paint / painted} their portraits. [not equivalent to (35)]

By contrast, ‘identifying’ (31) has a corresponding non-cleft:

(31’) Did you go to an interesting meeting last night?

Finally, Declerck claims that neither exhaustiveness nor contrast is involved in (31). In this connection, we concur that the justification of the acceptability of indefinite NPs in the Local CA of (identifying) clefts is the contrastive content of their modifiers. Put another way, *an interesting meeting*, even though formally indefinite, is specific at least from the speaker’s viewpoint. The contrast expressed by the example is confirmed by its intonation pattern, recognised as such by Declerck:
(31)” Was it an INTERESTING meeting that you went to last night? [–No, it was a BORING meeting...]

Once the contrastiveness reading of (31) has been accepted, the meaning of exhaustiveness can be justified. Finally, as we see it, tense restrictions affect not only attributive sentences, as Declerck tries to show in (36), but also identifying ones like (37):

(36) It would be/*is a more interesting subject that we would be discussing if John had his way.
(37) It would be/?? is Linguistics that we would be discussing if John had his way.

Summing up, we have tried to demonstrate that Declerck’s mixed type of identifying/attributive clefts belongs to the general one of (identifying) clefts. Apart from semantics, already discussed, all the grammatical factors (existence of noncleft counterparts, intonation and even pronominal connection) in Declerck’s study also point towards the aforementioned conclusion:

(38) Was it a MEETING that you went to last night?
(39) It was an excellent description of himself/*him that John wrote.

We have agreed with Declerck about the exclusion of certain it-be structures from the class of clefts. To attempt the analysis of these examples is beyond the aims of this chapter. It is their exclusion from the class of clefts that must remain uncontroversial for the purposes of the corpus study.

4.2. Meta-informative structure

From a communicative/pragmatic point of view, clefting rearranges the topic-comment structure of the sentence. On the one hand, the information in the rightmost that-clauses is presupposed (Prince’s 1978 or Declerck’s 1988 ‘informative-presuppositional’ clefts) and, on the other, the Local CA is focal (Declerck’s ‘stressed-focus’). This dual communicative nature of clefts is summarised in the following quotation by Enkvist (1979: 151): “clefted elements (...) express new information and evoke presuppositional sets” [our italics].

It is normally assumed that clefts are structures that determine pragmatic functions in a meta-informative way. More specifically, in adopting the meta-informative cleft pattern, one is formally recognising that the that-clause following the Local CA is pragmatically presupposed or given (Engelkamp and Zimmer 1983: 40, Brömser 1984: 330) from the speaker’s viewpoint, even though it is actually new for the hearer.17 However, as Lambrecht (1994: 70-71) points out, in a sentence like (40)

---

17 See Croft (1991: 119), who claims that “the information in the backgrounded clause refers back to an already open cognitive file, namely, a file established under ‘generally known facts’. Thus, even if the backgrounded clause is new information to the hearer, a cognitive file has been opened for it”. Similarly, this new/given duality is supported by Engelkamp and Zimmer (1983: 64) when they say that “cleft sentences are used in particular communicative situations, i.e. when the speaker fills an ‘active’ [author: givenness dimension] gap [author: newness dimension] in the hearer’s knowledge” [our italics].
based on his (2.21)], uttered out of the blue by a lecturer, the information conveyed by
that said that... does not have to be presupposed by the audience, which contradicts
both Hetzron’s (1975: 361) assertion that clefts cannot be “unprecedented”, and
Rochemont’s (1986) and Declerck’s (1988: 231) predictions that clefts cannot be used
out of the blue because they require the presupposed part to be in the hearer’s
consciousness:

(40) It was George Orwell that said that the best books are those which tell you what
you already know.

We contend that even though the content of the that-clause in (40) is clearly not shared
by the audience, it is – and here lies the justification of the cleft meta-informative
organisation– presented as such. Harold (1995: 158) claims that the information in the
that-clause of a cleft “is generally not in the listener’s mind. It may be known or
inferable or it may be totally new”. In this latter case, as pointed out before (see also
Declerck 1988), by means of the clefting device, the new information is presented as if
it were old. This leads us to conclude that the aim of English clefts is to trigger the
interpretation of the that-clause after the Local CA as given, be it a carrier of old
information or not.

Clefts are also claimed to bring forward to the attention of the hearer/reader a
certain (post-be) theme of the discourse as the focus of attention (the Local CA), which
makes clefts be considered meta-informative devices for focus-marking (Rochemont’s
1986 ‘constructional focus construction’). This ‘presentative’ function is outlined in,
for instance, Hetzron (1975: 362-364), Van Oosten (1986: 59) or Declerck (see the
section devoted to so-called ‘stressed-focus clefts’ in his 1988 study). Consequently,
the meta-informative layout of the cleft makes the hearer interpret George Orwell as
focal in, for example, (40) above. Once it has been characterised as such, the referent
of the Local CA may become the ensuing discourse theme.

The principle of end-focus (as part of the general one of given-before-new),
when applied to the it-cleft pattern, predicts that, on the one hand, the rightmost clause
of the it-cleft must contain information either unavailable or less predictable than the
information conveyed by other constituents preceding the final clause, and, on the
other, which is in fact a consequence of the previous one, the referentiality potential
of the segment in postverbal Local CA should be, communicatively speaking, more
referring than that of the final clause.

5. The data

This section accounts for the data employed in our analysis of it-cleft sentences in the
history of English. In Section 5.1 we account for the frequency of the construction from
Late Middle English to the present. Section 5.2 deals with the informative status of the
major constituents in the it-clefts in the database.

5.1. Frequency of it-clefts

Table 1 gives information about the database on which this investigation is based. The
Late Middle (LME) and Early Modern English (EModE) data have been retrieved from
The Helsinki Corpus of English Texts. The Late Modern English (LModE) examples
have been taken from A Representative Corpus of Historical English Registers
(ARCHER; see Biber et al. 1994) and the Corpus of English Dialogues (CED). Finally, the contemporary data have been selected from The Lancaster-Oslo/Bergen Corpus of British English (LOB, texts dated in 1961).

Table 1. The corpus (raw data and normalised frequencies [n.f.] per 100,000 words and 1,000 clauses)

<table>
<thead>
<tr>
<th>Period</th>
<th>words</th>
<th>clauses</th>
<th>n.f./100,000 words</th>
<th>n.f./1,000 clauses</th>
</tr>
</thead>
<tbody>
<tr>
<td>LME: 1420-1500</td>
<td>71,097</td>
<td>4,751</td>
<td>4.21</td>
<td>0.63</td>
</tr>
<tr>
<td>EModEI: 1500-1570</td>
<td>61,219</td>
<td>3,891</td>
<td>6.53</td>
<td>1.02</td>
</tr>
<tr>
<td>EModEII: 1570-1640</td>
<td>75,762</td>
<td>5,729</td>
<td>6.59</td>
<td>0.87</td>
</tr>
<tr>
<td>EModEIII: 1640-1710</td>
<td>62,940</td>
<td>4,360</td>
<td>17.47</td>
<td>2.52</td>
</tr>
<tr>
<td>LModE: 1710-1900</td>
<td>67,962</td>
<td>6,247</td>
<td>19.12</td>
<td>2.08</td>
</tr>
<tr>
<td>PDE: 1961</td>
<td>98,007</td>
<td>6,974</td>
<td>36.73</td>
<td>5.16</td>
</tr>
<tr>
<td></td>
<td>436,987</td>
<td>31,952</td>
<td>16.47</td>
<td>2.25</td>
</tr>
</tbody>
</table>

It-clefting has not been a productive thematic mechanism in the history of the English language. As Ball (1994b: 610) observes, a few “clefts and cleft-like constructions are attested in (...) OE and Early Middle English”. Whereas in LME the number of *it*-clefts is not significant at all in the corpus employed (0.63), at the end of the modern period, their frequency is slightly above two out of one thousand clauses. In PDE the proportion of clefts reaches and exceeds the figure of 5 out of one thousand clauses.

5.2. Information in *it*-clefts

In Section 4.1 we concluded that end-focus, when applied to the *it*-cleft pattern, predicts that the rightmost clause of the *it*-cleft conveys information which is not predictable and that the informative status of the Local CA should be more referring than that of the final clause. Tables 2 and 3 contain the data for the referentiality potential of, respectively, the rightmost clauses and the Local CAs of *it*-clefts. The basic taxonomy of informative content is as follows:

- referring (ref) constituents, either linguistically/textually or deictically, when the content of the linguistic expression has already been mentioned or alluded to in the discourse, belongs to universal knowledge, or is a current situational element. In keeping with Ariel (1996: 23ff), an expression will be regarded as linguistically/textually referring if it is not new in a span of seven sentences prior to its occurrence.

- low-referring (low-ref) constituents, when only non-head components of the expression

- non-referring (non-ref) constituents, when the referent can neither be recalled from the discourse domain linguistically – in the previous seven clauses –, situationally or permanently, nor derived from a previous referring expression.

Table 2. Referentiality of the sentence-final clause
According to the information displayed in Table 2, the informative content of the sentence-final clauses of the *it*-clefts in the database is referring in the majority of the cases. What is more, if the results of low-referring information, that is, of constituents whose head is not referring and whose modifiers/complements are referring, are computed together with the ones of referring information, the proportion of either completely or partially referring segments in final position is even higher. These data lead to the rejection of end-focus since such a pragmatic principle does not stand up well under the onslaught of the high proportion of available information conveyed by the rightmost clauses of the clefts under examination. In Atlas and Levinson’s (1981: 16) words, the *it*-cleft “contravenes the convention that old information precede new information”.

As regards the referentiality potential of the other major constituent of a cleft, namely, the Local CA, the scores in Table 3 evince a situation which is quite different from the one depicted in Table 2. Whereas most final clauses contain available referents, the Local CAs of the *it*-clefts analysed are normally non-referring, especially from EModE onwards. These results make it possible to corroborate the informative characterisation of *it*-clefts as focusing meta-informative strategies.

### 6. Summary and concluding remarks

In Section 1 we defined the concept ‘cleft sentence’, which was restricted to the meta-informative pattern known as ‘*it*-cleft’ since pseudo- or *wh*-clefts were regarded as identifying or attributive copulative constructions comprising sentence-initial *wh*-clauses realising the function Global CA, and pronominal and *th*-clefts were treated as examples of nonrestrictive relativisation or even (pseudo-) right dislocation. In Sections 1 and 2 we described the basic structure of an *it*-cleft as follows: Global CA in the form of the AS *it* followed by an link verb, then the Local CA and finally either a nonfinite or a finite non-meta-informatively-centered clause, the latter being preceded by one of the following introducers: Ø, *that*, *which* or *who*. Several conditions and remarks seem in order here concerning this structure, which were discussed in detail in the course of this paper:

(i) existence of a null pointer within the rightmost clause: One of the constituents of the
rightmost clause is promoted to Local CA position.

(ii) categories in Local CA position: Most categories can occur in Local CA position, non-contrastive adjective phrases (APs) and verbal constituents being excluded.

(iii) syntactic functions related to the Local CA segments: The Local CA segments can be associated with null pointers fulfilling most syntactic functions within the rightmost clause, with the exception of verb-based functional roles, disjuncts and conjuncts. The usual functions realised by the null pointers of the sentence-final clauses of the corpus were subject and adjunct, which was accounted for on the basis that the constituents fulfilling these functions cannot easily be focalised by means of other strategies.

(iv) information: The information conveyed by the Local CA segment is unavailable (either new or, as is claimed in the literature, ‘presented as new’ by the speaker) to the hearer in the majority of the cases. Such an assertion, which was corroborated by the data, characterised the cleft construction as a focusing meta-informative device, not conditioned by the given-before-new principle.

The mechanism of clefting has then become a thematising meta-informative strategy which has consolidated itself in the English language quite recently. By means of this system, the speaker focuses on the referent of the Local CA constituent that occupies the X-position and places heavy (normally available) information in final location within the background rightmost clause. The consolidation of the it-cleft construction has been corroborated both quantitatively and qualitatively. On the one hand, even though the strategy of clefting is not a frequent constructional option in the periods investigated, an important increase in the number of it-clefts could be noticed from EModE onwards. On the other hand, the qualitative spread of the meta-informative pattern was corroborated by means of the growing number of both syntactic functions and categories associated with the Local CA segment. In fact, with a few exceptions, almost every category can occupy the position reserved for the Local CA segments, and the Local CAs constituents can be related to null pointers realising almost every syntactic role in the rightmost clauses (subject, object, predicative, prepositional complement, adverbial complement, adjunct).

References

Sources

ARCHER = A Representative Corpus of Historical English Registers 1990-1993/2000. Compiled under the supervision of Douglas Biber and Edward Finegan at Northern Arizona University, University of Southern California, University of Helsinki and Uppsala University.


of information to accompany the Lancaster-Oslo/Bergen Corpus of British English, for use with digital computers. Oslo: University of Oslo (Department of English).

Secondary references


