

‘These hands, they are apt enough to dislocate and tear thy flesh’: On Left Dislocation in the Recent History of the English Language*

Javier Pérez Guerra & David Tizón-Couto
University of Vigo
jperez@uvigo.es

Abstract

As part of a major project on the syntactic organisation of written discourse in the recent history of the English language, this paper tackles the distribution of sentences comprising left-dislocated constituents in a corpus of texts from late Middle English onwards. Once the phenomenon of left dislocation has been properly defined, this investigation will concentrate on the analysis of the corpus in the following directions: (i) statistical evolution of left dislocation in the recent history of the English language; (ii) the influence of orality and genre on left dislocation; (iii) information conveyed by the left-dislocated material, that is, the discourse-based referentiality potential of the left-dislocated constituents in terms of recoverability, and its association with end-focus; and (iv) grammatical complexity of the left-dislocated material and its association with end-weight.

1 Introduction

This paper deals with the linguistic phenomenon of ‘left dislocation’¹ (henceforth, LD), exemplified in (1):

- (1) This paper, I have not had the opportunity of reading it at the workshop.

In LD constructions, a sentence-initial constituent is suprasegmentally detached from the (rest of the) sentence and is prototypically resumed by a pronominal copy in the sentence.

An issue to which much attention has been devoted in the literature on LD has been whether the strategy of LD must be dealt with as either a discourse- or simply as a sentence-based phenomenon. In this respect, the syntactic

* The research which is here reported has been funded by the Spanish Ministry of Science and Technology, grant number BFF2001-3505, and Xunta de Galicia, grant number PGIDT01PXI20404PR. Both grants are hereby gratefully acknowledged.

¹ The label *left dislocation* was first suggested, to our knowledge, by Ross (1967: 232). English LD belongs to the subtype of ‘hanging topic left dislocation’, as reported by, for example, Van Riemsdijk (1997) and Vat (1997).

nature of LD has been called into question on many occasions (Keenan 1976: 253; Keenan & Schieffelin 1976: 241; Vat 1997: 95-100; Ziv 1994: 631f).² In this paper, on a par with Pérez-Guerra (1999: Chapter 6), we shall contend that LD is relevant to syntax since left-dislocated constituents must meet two (syntactic) conditions: first, they must not be insertable directly in the clause, and, second, they must be related to the material after the comma or pause semantically and, sometimes, syntactically. As far as the first condition is concerned, the integration of, for example, *this paper* in (1) above within the clause *I have not had the opportunity of reading it at the workshop* would lead to (syntactic) redundancy, since the object slot of the predicate *reading* would be filled by both the proform *it* and the left-dislocated constituent *this paper*. As regards the second condition, the connection between *this paper* and *it* in (1) is both semantic and syntactic (Geluykens' 1992 'average' LD), since, on the one hand, both constituents are coreferring (semantic link) and, on the other, *this paper* is capable of functioning as the object of *reading* in the absence of *it* (syntactic link). An example of non-syntactic semantic connection between the LD constituent and the sentence (Gundel's 1988: 244 'double-subject construction' or Lambrecht's 1994: 193 'unlinked topic construction') is (2), in which *this paper*, on the one hand, cannot be integrated into the sentence (**I'm going crazy this paper*) and, on the other, is not resumed by any constituent in the ensuing clause.

(2) This paper, I'm going crazy.

This paper in (2) is, however, semantically related to the proposition *I'm going crazy* since it constitutes the 'aboutness' of the speaker's craziness. Such semantic connection between the left-dislocated expression and the sentence is defined in more detail in (3), which constitutes the definition of LD which we shall stick to in this paper:

(3) LD REQUIREMENTS:

A sentence is said to contain a left-dislocated constituent if it satisfies (a) and either (b) or (c):

- a) A non-vocative (see Van Oosten 1986: 32) constituent other than the unmarked theme is in sentence-initial position, and a pause (comma, in writing) is 'felt' to occur between that segment and the rest of the clause. This sentence-initial constituent cannot fulfil a function in the sentence which it introduces. In other words, it cannot be inserted directly in the syntactic structure of the ensuing clause.

² The discourse nature of LD has been accepted even in 'syntactocentric' proposals couched in the generative framework, such as Emonds (2004: 106ff), who maintains that the maximal projection of which the left-dislocated material is an immediate constituent is a so-called 'Discourse Shell'.

- b) A syntactic relation holds between the preposed segment and another element fulfilling a basic function in the sentence.³ In other words, a copy (or a referent, in those cases showing backward pronominalisation) of the preposed segment occurs in the sentence.^{4,5}
- c) A semantic relation⁶ holds between the preposed segment and another element fulfilling a basic function in the sentence, in such a way that both the preposed and its related elements share semantic features ('coreference' in Gundel's 1988:223 terminology). Alternatively, the element in the sentence may be, semantically speaking, 'part' of the preposed one, which thus functions, informatively speaking, as setting.⁷

³ By 'basic function' we understand a complement position. In actual fact, as pointed out by Rodman (1974), adjuncts, which are modifiers, cannot be left-dislocated:

- (i) *In Rutgers, they hold a wonderful stock of electronic references on Optimality Theory.

⁴ The necessity for a relation of coreference to hold between the left-dislocated constituent and a copy or referent in the clause requires that the left-dislocated expression is capable of invoking an extralinguistic entity. In other words, adjectives, adverbs, prepositions, etc. are hard to use as left-dislocated constituents since they do not refer to an entity in the extralinguistic world. Nominal expressions, which by definition are referential, are thus the prototypical categories that will most likely occur in the slot reserved for left-dislocated segments. Clauses, as instantiations of facts, are also possible left-dislocated expressions. Just for the record, (i) and (ii) illustrate, respectively, a small clause (*all Nonsense*) and a *that*-clause (*the Man has something of a Notion at Dress*) in LD position, these both being resumed by the proform *it* in the clause.

- (i) All Nonsense, I know it (H. Brooke, *The Fool of Quality*: 75)
- (ii) the Man has something of a Notion at Dress, I confess it –
(F. Coventry, *Pompey the Little*, Book 1: 26)

⁵ Backward pronominalisation is possible in LD examples, as shown in (i), taken from Haaften *et al* 1983:

- (i) [The first of his_i papers]_j, I think [every linguist]_i would qualify it_j as a failure.

The referential link marked by means of subscript 'i' illustrates backward pronominalisation between the proform *his* and the referent *every linguist*. Subscript 'j', in contrast, evinces forward pronominalisation between *the first of his papers* and the proform *it*.

⁶ The referential relation between the left-dislocated segment and the constituent in the clause is ruled by general pragmatic principles such as the 'Parallel Function Strategy', as reported in, for example, Cowan 1995:34. In this respect, Gundel (1975:88) suggests that such semantic connection is governed by a maxim stating that '[i]n order for a comment, C, to be successfully predicated of a topic, T must be of a type or category such that it is logically possible for C to be true or false of T', where T stands for the preposed segment and C for the sentence. Dik (1998: 394) claims that '[f]or any pair of Theme T [JPG/DTC: LD] and clause C to make sense, it must be *relevant* to pronounce C with respect to T' [our italics].

⁷ The so-called setting-subtype is here illustrated in (i), in which the left-dislocated constituent *the best, and most beloved of wives, of mothers, of mistresses* is resumed by *her*,

Examples such as (4) below are not accounted for by the definition in (3) and thus, *contra* Geluykens (1992: 20ff), among others, are not regarded as examples of LD:⁸

- (4) As for LD, the syntactic integration of the dislocated material in the matrix sentence is not possible.

We will not consider *as for LD* in (4) as a left-dislocated constituent since it fulfils a function in the sentence, namely, topical or aspectual modifier, which contradicts condition (a) in (3), that is, the impossibility for left-dislocated constituents to be inserted in the clause. (4) will be classified as an example of topicalisation, a syntactic strategy closely related to LD since both constructions share the feature of semantic affinity between the fronted constituent and the clause.⁹ LD and topicalisation differ precisely in their compliance with condition (a): whereas left-dislocated constituents cannot be integrated in the syntactic structure of the clause, topicalised segments fulfil major syntactic functions in the sentence. The degree of the syntactic integration of a topicalised expression in the sentence depends on its syntactic function. To give an example, the connection between topicalised *as for LD* and the clause in (4) above is weaker than that holding between *LD* and the clause in, for example, (5) below:

- (5) LD material we cannot integrate in the matrix clause.

In example (5) *LD material* is a topicalised object, that is, ‘moved’ from its unmarked postverbal position as the object of *integrate*. Whereas *as for LD* in (4) is a modifier, *LD material* in (5) is an (internal) argument. The omission of *as for LD* in (4) has no consequences for the interpretation of the sentence, whilst *LD material* in (5) cannot be deleted. Be that as it may, neither (4) nor (5) can be taken as examples of LD since, to a larger or a lesser extent, *as for LD* and *LD material* realise syntactic functions in their sentences or, in other words, are syntactically integrated in the clauses.

which functions as the genitive determiner of *domestick character*, and not by a proform fulfilling a major function.

- (i) The best, and most beloved of wives, of mothers, of mistresses, her domestick character is most lovely;
(F.M. Brooke, *The History of Lady Julia Mandeville*, Vol. I: 11)

⁸ Most of the corpus examples initiated by topic adjuncts contain the introducers *as to* and *as for*, illustrated here in (i) and (ii), respectively:

- (i) As to its being the language in Paradise, this is not very probable,
(T. Amory, *John Buncl. Vol. I: 38:40*)
(ii) As for you my beloved Son, you are now turn'd of fourteen,
(M. Davys, *The Accomplish'd Rake: 3*)

⁹ See Pérez-Guerra 1999: 200-202 for syntactic differences between LD and topicalization.

The definition in (3) prevents vocatives and similar exclamative constructions from being regarded as examples of LD. We justify this constraint by assuming that vocatives constitute speech units by themselves and are not analysed as part of the clauses containing the so-called copies. To give an example, (6) is claimed to consist of two speech units, namely *unhappy woman!* and *I can only regard her as an object of pity!* and thus is not eligible as an illustration of LD:

- (6) Unhappy woman! I can only regard her as an object of pity!
(F. Burney, *Evelina*: 5)

Vocatives must not be confused with appositive constructions even though they share structural similarities. In this vein, our corpus contains examples of appositions of the sort in (7), with the first term acting as a premodifier of the head term, which resemble the instances of vocatives already discussed.

- (7) Your most humble Servant, cry'd Sir John, I find then you are going to compleat your happy Circumstances in that mighty Blessing call'd a Wife, (M. Davys, *The Accomplish'd Rake*: 15)

(7) cannot be taken as an example either of a (second-person)¹⁰ vocative since *your most humble Servant* corefers with first-person *I* or of LD, since the initial constituent is integrated into the syntactic structure of the clause and fulfils the syntactic function of premodifier within the subject *Your most humble Servant (...) I*.

The distinction between a vocative and a premodifying appositive term is blurred in examples such as (8):

- (8) How, my dear creature, have I used you inhumanly? (S. Fielding, *The Cry*: 97)

My dear creature can be analysed as either a vocative or the premodifier of *you*. Be that as it may, (8) is not an example of LD since sentence-initial *my dear creature* either constitutes a specific speech unit by itself or realises the function of premodifier within the object *my dear creature (...) you*.

A further condition on LD which is specified, in passing, in the definition in (3) is the necessity for the left-dislocated constituent and its 'copy' to occur

¹⁰ Exclamative expressions which do not convey second-person reference will not be treated as vocatives and can, in consequence, be taken as examples of LD. To give an example, we do not object to the analysis of *we!* in (i) as a left-dislocated constituent, resumed by the second occurrence of *we* – repetition, as a subtype of LD, is common to specifically spoken language:

- (i) We! [my Lord!] [cried they with one voice], we would not go up to the gallery for your Highness's revenue. (H. Walpole, *The Castle of Otranto*: 40)

in the same clause. This requirement prevents examples such as (9) and (11) from being regarded as instances of LD:

- (9) Lord Viscount Fondville, he would not have you omit Viscount for the world, left us this morning (F.M. Brooke, *The History of Lady Julia Mandeville*: 73)
- (10) ‘Kissing!’ said the Lady, ‘do you call that no Crime?’ (H. Fielding, *Joseph Andrews*: 41)
- (11) On this she asked me if I knew Polly Philips. Undoubtedly, says I, the fair girl which was so tender of me when I was sick, (J. Cleland, *Memoirs of a Woman of Pleasure, Vol. I*: 75)

On the one hand, in (9) above, *he would not have you omit Viscount for the world* is a comment clause which is inserted with the matrix clause *Lord Viscount Fondville (...) left us this morning*. Put differently, *Lord Viscount Fondville* cannot be said to be a left-dislocated constituent resumed by means of *he* in the parenthetical clause since the former functions as the subject of *left*. Example (10) is, in contrast, an example of LD since *kissing* and *that* belong to the same sentence *Kissing (...) do you call that no Crime?*. This is an example of a direct reported discourse interrupted by the clause including the verb of speech *said*.

On the other hand, the referential link holding between *Polly Philips* and *the fair girl which was so tender of me when I was sick* in example (11) above is not an instantiation of LD, since the two constituents are uttered by different speakers – *she* (or the speaker) and *I*. This obviously shows that these segments cannot be constituents of the same clause. It seems in order here to point out that our (‘same-sentence’) constraint is not necessarily in keeping with the literature on LD. In this respect, Geluykens (1992: 23-24) accepts dialogic LD since, in his opinion, LD is an interactive strategy whose definition must be stated in pragmatic rather than in syntactic terms. To our knowledge, if examples such as (11) above are treated as instances of LD, then there are no syntactic differences between LD and simple reference, which is the semantic phenomenon that accounts for the connection between *Polly Philips* and *the fair girl which was so tender of me when I was sick*.

The ‘same-sentence’ constraint, as already defined, leads to the exclusion of (12) from the class of LD constructions:

- (12) ’tis that fine Piece of his, where – Yes, yes, I have read it very often; I remember it perfectly well
(F. Coventry, *Pompey the Little, Book 1*: 63-64)

Since *that fine Piece of his, where (...)* and *it* belong to different clauses, example (12) will simply constitute an illustration of the pronominalisation of *that fine Piece of his, where (...)* by the personal pronoun *it*. This example has

be, syntactically speaking, identical to (14) above and therefore would be not eligible as a left-dislocated construction.¹¹

2 The Corpus

As part of a major project on the thematic organisation of the clause in the recent history of English, the aim of this paper is to analyse LD once word order has been fixed in the language. To that end, we base our investigation on data taken from three computerised corpora containing material from late Middle (lME), early Modern (eModE), late Modern (lModE) and Present-day English (PDE). The data for the lME and the eModE periods have been retrieved from *The Helsinki Corpus of English Texts* (Kytö 1996). The literary works by eighteen authors in the *Chadwyck-Healey* collection (*Eighteenth Century Fiction*) have served as the basis for the eighteenth century. The *Lancaster-Oslo/Bergen Corpus of British English* (Johansson 1978) has offered the material for PDE. Table 1 gives the word totals for each period:

¹¹ When the proform or the resumptive expression holds a semantic relation of multiple coreference with the constituent occurring before the break, the analysis of the construction as LD is out of the question. Multiple coreference is illustrated in (i), in which *they* refers to *Don Medenta* plus *Charlotta*:

- (i) Don Medenta leading Charlotta, they wandered to a Place where they saw some
Trees growing very close together, (P. Aubin, *Charlotta Du Pont*: 50)

The 'same-sentence' constraint predicts that example (i) would be eligible as LD if the left-dislocated segment and the copy occur in the same sentence. This would imply that *Don Medenta leading Charlotta* constitutes not an independent clause but a non-clausal category. If *leading Charlotta* were a postmodifier of *Don Medenta*, then *Don Medenta leading Charlotta* would be a noun phrase which could be claimed to be left-dislocated. Since the supposed copy proform *they* resumes not the singular referent materialised by *Don Medenta* plus its nonfinite postmodifier but *Don Medenta* plus *Charlotta*, *Don Medenta leading Charlotta* does not convey unique reference and thus cannot be analysed as a (left-dislocated) noun phrase. The constituent occurring before the comma is, in consequence, a subordinate clause functioning as a (time/reason/manner) adjunct.

Table 1: The corpus

Period	Word totals
IME	71,097
eModE	199,921
lModE	311,566
PDE	98,007
Totals	680,591

3 A Statistical Overview of LD

The distribution of LD in the periods investigated is shown in Table 2. So that the raw results (Roman font) can be compared among periods, we have normalised the results per 1,000 words (italics):¹²

Table 2: Distribution of LD

Type	IME	eModE	lModE	PDE	Total
LD proper	85/ <i>1.19</i>	78/ <i>0.39</i>	46/ <i>0.14</i>	10/ <i>0.1</i>	219/ <i>0.32</i>
<i>wh</i> -LD	13/ <i>0.18</i>	18/ <i>0.09</i>	1/ <i>0.003</i>	0	32/ <i>0.04</i>
Totals	98/ <i>1.37</i>	96/ <i>0.48</i>	47/ <i>0.15</i>	10/ <i>0.1</i>	251/ <i>0.36</i>

Two types of LD have been identified in Table 2, namely LD proper and *wh*-LD, in which a *wh*-constituent appears in a position reserved for the left-dislocated material. An example of *wh*-LD is (16):

- (16) what þou fyndis þer-in, do it of clene
(The 'Liber De Diversis Medicinis' in the Thornton Manuscript:10)

The referent of the unbound relative clause *what þou fyndis þer-in* in (16) is resumed by the proform *it* in the ensuing kernel clause. Since the location of *wh* left-dislocated constituents in sentence-initial position is not due to the fact that they are *wh*-segments, as confirmed by (17), we shall consider *wh*-LD as a subtype of LD proper.

- (17) do what þou fyndis þer-in of clene

The figures in Table 2 show that *wh*-LD is highly stigmatised in the language, at least from lModE onward, since only one example complying with the *wh* subtype has been recorded in lModE and none in the PDE period. Even though the lack of examples of *wh*-LD could be claimed to be a consequence of the literary nature of the texts, the PDE subcorpus, which consists of different text types, points towards the progressive avoidance of *wh*-LD. The only example which has been classified as *wh*-LD in lModE is (18) below:

¹² Normalisation per number of clauses has been carried out in, for example, Pérez-Guerra, forthcoming and has not modified the results significantly.

- (18) whatever enemies I find without, I will always endeavour not to cherish one in my own bosom.

(S. Fielding, *The Cry*, Vol. I: 38)

One is a partitive proform which resumes the relative clause *whatever enemies I find without*. The example thus conforms to condition (c) in (3) above since, semantically speaking, *one* conveys a ‘part’ of the preposed *wh*-clause, the latter functioning, informatively speaking, as setting.

When applied to the distribution in Table 2, chi-square demonstrates that the rate of variation undergone by LD – both LD proper and *wh*-LD – from IME onwards is statistically significant.¹³ In fact, the proportion of LD from IModE onwards amounts to approximately 1 out of 10,000 words, which places the strategy of LD in a marginal position as far as word order is concerned. Such a conclusion is in keeping with the process of syntactisation (see Pérez-Guerra, forthcoming), in progress in these periods, according to which peripheral constituents which cannot be integrated in the syntactic structure of the clause are avoided in at least written (planned) linguistic production. In the fight towards full syntactisation, LD is especially vulnerable since it involves elements which, by definition, cannot be accommodated within the syntactic structure of the clause. In order to confirm whether the decrease undergone by left-dislocated material across time can also be applied to other fronting strategies or not, in Table 3 we include information about the distribution of topicalisation¹⁴ (see Section 1 above) and adverbial fronting (sentence-initial adjuncts, disjuncts and conjuncts) in the same periods – unfortunately, no data can be offered for the strategy of topicalisation in the IModE period. The normalised figures are given in italics:

Table 3: *Distribution of other fronting strategies*

Type	IME	eModE	PDE	Total
topicalisation	1,225/17.22	2,676/13.38	1,171/11.94	5,072/13.74
adverbial fronting	54/0.75	667/3.33	336/3.42	1,057/2.86
Totals	1,279/17.98	3,343/16.72	1,507/15.37	6,129/16.6

¹³ For $p \leq 0.001$, $\chi^2 = 26.4245$. The distribution is significant.

¹⁴ Sentence-initial adverbial complements, illustrated in (i), are included in the data for topicalisation since its location in initial position is less natural, and thus more marked, than in postverbal position. Adjuncts, in contrast, occur naturally in sentence-initial position and are, in consequence, included in the counts for adverbial fronting:

- (i) [But in this commendation of musike I wold nat be thought to allure noble men to haue so moche delectation therin, that,] in playinge and singynge only, they shulde put their holle studie and felicitie (T. Elyot, *The Boke Named The Governour*: 26)

The information in Table 3 suggests that the decline of LD in written speech is in keeping with the general decrease of topicalisation.¹⁵ The figures also corroborate that adverbial fronting does not run parallel to topicalisation and LD, since the presence of an adjunct, disjunct or conjunct in sentence-initial position has minor consequences for the syntactic organisation of a sentence.¹⁶

4 Suprasyntactic Variables Affecting LD

4.1 Style and Genre

Table 4 displays the distribution of LD in the corpus according to the style/genre of the texts, namely, expository, instructive and statutory.¹⁷

Table 4: Style and LD

Type	IME	eModE	IModE	PDE
expository	0	13	-	0
instructive	71	24	-	2
narrative	11	28	47	8
statutory	2	1	-	0
others	14	30	-	0
Totals	98	96	47	10

Since all the texts in the eighteenth-century subcorpus belong to the narrative subcategory, only the IME, eModE and PDE data allow for further

¹⁵ Kohonen (1978: 174) claims that the number of clauses with fronted constituents decreases whilst the number of sentences conforming to SVO word order increases. As far as the proportion of adjuncts specifically is concerned, corpus-based Breivik & Swan 1994: 28 leads to the conclusion that '[t]he relative frequency of sentences with initial adverbial has apparently decreased since the Old English period'. Pérez-Guerra (1999: 220) shows that the number of sentence-initial conjuncts decreases across time. In contrast, disjuncts are surprisingly more frequent in PDE; such a fact reinforces the claim that at least contemporary English is more involved than it used to be in older periods (see Traugott 1989, 1995: 44-49; Finegan 1995:8-10; Bækken 1998:7; González Álvarez 2002: 287, 289; or, on a more theoretical basis, Langacker 1990).

¹⁶ For $p \leq 0.025$, $\chi^2 = 8.4679$. The distribution is significant.

¹⁷ The ascription of the IME and eModE texts to the categories expository, instructive, narrative and statutory has been done by following the classification developed by the compilers of *The Helsinki Corpus*. In that corpus, the tag <Z> labels the different texts are EXP(ository), INS(tractive), NARR(ative) or STAT(utory). We have applied the same categorisation to the PDE texts. As pointed out in the main text, all the samples in the IModE subcorpus belong to the narrative subcategory. A final remark seems in order here since the statistical validation of the distribution in Table 4 is severely conditioned by the fact that all the data in the statutory and the expository rows belong to solely a few textual samples.

examination as far as the style of the texts is concerned. Even though the figures in Table 4 have not undergone any kind of statistical normalisation, one easily reaches the conclusion that LD cannot be characterised as a strategy applying to exclusively informal language since most of the corpus examples are found in instructive and narrative text types, which clearly instantiate formal English.

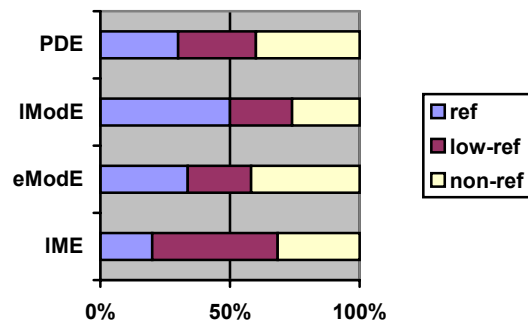
In the same vein, Pérez-Guerra (1999: 223-24) points out that the proportion of LD is considerably lower in the LME, eModE and PDE spoken texts than in the written – and even the written-to-be-spoken – passages, and that the number of sentences with left-dislocated material decreases across time independently of the written or the spoken character of the texts. This fact contradicts Geluykens' (1992: 34) claim that LD is more frequent in the spoken medium.

4.2 Information

This paper will not deal with the functions of LD in the discourse – expressing contrast, emotion, setting the scene, identifying a part of previous discourse, self-correction or hesitation, etc. – and will simply examine the informative potential of the left-dislocated material in the corpus in order to either corroborate or refute the informative characterisation of LD as is maintained in the relevant literature. To that end, we have designed a basic taxonomy of informative content which classifies left-dislocated constituents as follows:

- 'referring', either linguistically/textually or deictically, when the content of the linguistic expression has already been mentioned or alluded to in the discourse, belongs to the 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 previous to its occurrence.
- 'low-referring', when only non-head components of the expression – i.e. its complements or modifiers – are informationally 'available', when the entity denoted by the expression is 'derivable' from the linguistic context, or when its referent has been alluded to in the previous discourse in a span of more than seven clauses, and
- 'nonreferring', 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.

Graphic 1 displays the proportions of the informative category just described in the corpus:



Graphic 1: LD and information

The results in Graphic 1 do not warrant any conclusions as far as the historical characterisation of LD in terms of information is concerned, since the proportions of recoverable ('ref'), partially recoverable ('low-ref') and irrecoverable ('non-ref') information are practically identical in IME and PDE. This graphic also reveals that less than 40 percent of the left-dislocated segments are absolutely non-referring, which is not in keeping with the common belief that left-dislocated segments are highly irrecoverable (Geluykens 1992:33; Keenan & Schieffelin 1976; Prince 1997:124), that is, either absolutely new or re-introduced.¹⁸ The proportions in Graphic 1 show that the majority of the left-dislocated material has been verbalised in the immediately previous discourse.

4.3 Length

Table 5 outlines the average length – number of words – of the constituents occupying a left-dislocated position in the periods under examination:

¹⁸ Such informative characterisation of LD as a strategy which allows the insertion of irrecoverable referents in initial position leads to the distinction between left-dislocated material and average subjects, usually sentence-initial, since the latter normally convey recoverable information. Our own data do not, however, contradict the previous claim since what Graphic 1 portrays is the informative heterogeneity of left-dislocated segments, which contrasts with the high proportion of absolutely recoverable subjects. Whereas the unmarked organisation of the clause, with sentence-initial subjects, is in keeping with the principle of given-before-new (Quirk *et al* 1985: 1357), the occurrence of left-dislocated material in initial position does not corroborate such a principle, since dislocated themes are not clearly characterised as either recoverable or irrecoverable.

Table 5: Length

Period	Length
IME	8.9
eModE	8.8
IModE	11.4
PDE	7.4

Table 5 shows that LD must be characterised as a highly marked organisation strategy with respect to the length of the sentence-initial material since in all the periods the average number of words in the left-dislocated constituents is significantly higher than the length of unmarked subjects – 2.08 words in Pérez-Guerra 1999: 56. This fact indicates that the occurrence of a constituent in the position reserved for left-dislocated segments does not comply with the principle of end-weight (Biber *et al* 1999:898; Quirk *et al* 1985: 1361-62).

(19) exemplifies the resumption of a complex left-dislocated constituent by means of the proform *he* in the main clause:

- (19) Valerius, though a little Opposite at first, yet, upon his Mother's pressing, and repeating how far my Happiness was the Object, if not the whole End of the Undertaking, he at last consented,
(J. Barker, *J. Exilius*, Vol. I: 46)

5 Concluding Remarks

This paper has undertaken the characterisation of LD as a (semantically-constrained) syntactic strategy which alters the unmarked organisation of the clause. Once LD was defined and distinguished from other competing constructions – topicalisation, subordination, etc. –, special attention was devoted to the evolution of LD in the recent history of the English language, namely from 1420 onwards. In this respect, we have examined the diachronic frequency of the construction, its informative potential, the connection between LD and written/spoken and/or formal/informal language, as well as the compliance of LD with the principle of end-weight.

The main conclusions are as follows. First, the data show, on the one hand, the radical decrease of *wh*-LD and, on the other, the diminution of LD proper, in accordance with the general decrease of other fronting strategies such as topicalisation. Second, LD is more productive in written and in formal texts, which contradicts other assumptions found in the relevant literature. Third, the informative nature of left-dislocated constituents is considerably heterogeneous since they accommodate both recoverable and irrecoverable referents. Finally, left-dislocated constituents are very long, which places LD

in a highly marked position with respect to (lack of) compliance with end-weight.

Summing up, LD, both in written/spoken and formal/informal language, has been a marked syntactic strategy in the recent history of the language and, as such, is not subject to the principles which rule the organisation of the clause, namely, given-before-new, end-weight and the necessity of integration in the syntactic skeleton of the clause.

References

- Anagnostopoulou, Elena, Henk Van Riemsdijk and Frans Zwarts. eds. 1997. *Materials on Left Dislocation*. John Benjamins, Amsterdam.
- Ariel, Mira. 1996. Referring Expressions and the +/- Coreference Distinction. In Thorstein Fretheim and Jeanette K. Gundel, eds., *Reference and Referent Accessibility*. John Benjamins, Amsterdam, pp. 13-35.
- Bækken, Bjørg. 1998. *Word Order Patterns in Early Modern English with Special Reference to the Position of the Subject and the Finite Verb (Studia Anglistica Norvegica 9)*. Novus Press, Oslo.
- Biber, Douglas, Stig Johansson, Geoffrey Leech, Susan Conrad and Edward Finegan. 1999. *Longman Grammar of Spoken and Written English*. Longman, Harlow.
- Breivik, Leiv Egil and Toril Swan. 1994. Initial Adverbials and Word Order in English with Special Reference to the Early Modern English Period. In Dieter Kastovsky, ed., *Studies in Early Modern English*. Mouton de Gruyter, Berlin, pp. 11-43.
- Cowan, Ron. 1995. What are Discourse Principles Made of? In Pamela Downing and Michael Noonan, eds., *Word Order in Discourse*. John Benjamins, Amsterdam, pp. 29-49.
- Dik, Simon C. 1998. *The Theory of Functional Grammar. Part 2: Complex and Derived Constructions*. Mouton de Gruyter, Berlin (Ed. Kees Hengeveld).
- Eighteenth-century Fiction*. CD-ROM produced by Chadwyck-Healey.
- Emonds, Joseph. 2004. Unspecified Categories as the Key to Root Constructions. In David Adger, Cécile de Cat and George Tsoulas, eds., *Peripheries: Syntactic Edges and their Effects*. Kluwer, Dordrecht, pp. 75-120.
- Finegan, Edward. 1995. Subjectivity and Subjectivisation: An Introduction. In D. Stein and S. Wright, eds., pp. 1-15.
- Geluykens, Ronald. 1992. *From Discourse Process to Grammatical Construction: On Left-Dislocation in English*. John Benjamins, Amsterdam.
- González Álvarez, Dolores. 2002. *Disjunct Adverbs in Early Modern English: A Corpus-Based Study*. University of Vigo, Vigo.
- Gundel, Jeanette K. 1975. Left Dislocation and the Role of Topic-Comment Structure in Linguistic Theory. *Working Papers in Linguistics* 18. [Ohio State University]
- Gundel, Jeanette K. 1988. Universals of Topic-Comment Structure. In Michael Hammond, Edith A. Moravcsik and Jessica Wirth, eds., *Studies in Syntactic Typology*. John Benjamins, Amsterdam, pp. 209-42.
- Haaften, Ton V., Rik Smits and Jan Vat. 1983. Left Dislocation, Connectedness and Reconstruction. In K. Ehlich and Henk C. Van Riemsdijk, eds., *Connectedness in Sentence*. Tilburg University, Tilburg, pp. 43-69.
- (The) *Helsinki Corpus of English Texts*. Compiled by the University of Helsinki (see Kytö 1996).

- Johansson, Stig. 1978. *Manual of Information to Accompany the Lancaster-Oslo/Bergen Corpus of British English, for Use with Digital Computers*. University of Oslo (Department of English), Oslo.
- Keenan, Edward L. 1976. Towards a Universal Definition of 'Subject'. In Charles Li, ed., pp. 303-33.
- Keenan, Elinor Ochs and Bambi B. Schieffelin. 1976. Topic as a Discourse Notion: A Study of Topic in the Conversations of Children and Adults. In Charles Li, ed., pp. 335-84.
- Kohonen, Vilko. 1978. *On the Development of English Word Order in Religious Prose around 1000 and 1200 A.D.* Research Institute of the Åbo Akademi Foundation, Åbo.
- Kytö, Merja. comp. 1996. *Manual to the Diachronic Part of the Helsinki Corpus of English Texts. Coding Conventions and Lists of Source Texts*. University of Helsinki (Department of English), Helsinki.
- (The) *Lancaster-Oslo/Bergen Corpus of British English*. Compiled by the Universities of Lancaster and Oslo and the Norwegian Centre for the Humanities at Bergen (see Johansson 1978).
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge University Press, Cambridge.
- Langacker, Ronald W. 1990. Subjectification. *Cognitive Linguistics* 1/1: 5-38.
- Li, Charles N. ed. 1976. *Subject and Topic*. Academic Press, New York.
- Pérez-Guerra, Javier. 1999. *Historical English Syntax. A Statistical Corpus-based Study on the Organisation of Early Modern English Sentences*. Lincom Europa, München.
- Pérez-Guerra, Javier. Forthcoming. Word Order after the Loss of the Verb-second Constraint or the Importance of Early Modern English in the Fixation of Syntactic and Informative (Un-) Markedness.
- Prince, Ellen F. 1997. On the Functions of Left-Dislocation in English Discourse. In Akio Kamio, ed., *Directions in Functional Linguistics*. John Benjamins, Amsterdam, pp. 117-43.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech, and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. Longman, London.
- Rodman, Robert. 1974. On Left-Dislocation. *Papers in Linguistics* 7: 437-66.
- Ross, John Robert. 1967. *Constraints on Variables in Syntax*. Indiana University Linguistics Club, Bloomington.
- Traugott, Elizabeth Closs. 1989. On the Rise of Epistemic Meaning in English: An Example of Subjectification in Semantic Change. *Language* 65/1: 31-55.
- Traugott, Elizabeth Closs. 1995. Subjectification in Grammaticalization. In D. Stein and S. Wright, eds., pp. 31-54.
- Stein, Dieter and Susan Wright. eds. 1995. *Subjectivity and Subjectivisation: Linguistic Perspectives*. Cambridge University Press, Cambridge.
- Van Oosten, Jeanne. 1986. *The Nature of Subjects, Topics and Agents. A Cognitive Explanation*. Indiana University Linguistics Club, Bloomington.
- Van Riemsdijk, Henk C. 1997. Left Dislocation. in E. Anagnostopoulou *et al*, eds., pp. 1-10.
- Vat, Jan. 1997. Left Dislocation, Connectedness and Reconstruction. E. Anagnostopoulou *et al*, eds., pp. 67-92.
- Ziv, Yael. 1994. Left and Right Dislocations: Discourse Functions and Anaphora. *Journal of Pragmatics* 22: 629-45.