Development of dependent structures: an instance from the history of English

Cristina Suárez
Universidad Illes Balears

"such clauses are (almost) always at the margins
of the main clause, (almost) never flanked
by material from the main clause"
(O'Neil 1976: 199)

Abstract

Evidence has been provided for a diachronic development of dependent structures from independent ones and this process has recently been described as a process of inter-clausal grammaticalization (Hopper and Traugott 2003: chapter 7). Studies of clause linkage generally assume that dependent structures derive from originally independent ones. The aim of this paper is to illustrate the development of a particular example of dependent structures, particularly the case of relative clauses in the history of the English language.

This paper is structured as follows. First, the theoretical framework is presented. Next, I move to the analysis of relative clauses. I will present a brief description of the corpus and will pay more attention to the corpus analysis. Finally, some conclusions from the analysis are drawn.

1. Clause linkage

Clauses in contact in discourse inevitably acquire relations with one another. Such relations established in clause linkage must be considered as existing at two levels: a syntactic one and a semantic (or semantico-pragmatic) one, both in mutual dependence, as clearly put by Givón (1979: 208): "Syntax is not an independent, autonomous level of sentence organization, but rather a dependent, functionally motivated entity, whose formal properties reflect (…) the properties of the explanatory parameters which motivate its rise." Although the semantic nature of the relation between two interclausal clauses has traditionally been used as the primary argument in the classification of relations, it is not possible to establish a one-to-one correspondence between a given semantic relation and a given syntactic structure, as is going to be demonstrated. This presentation is concerned with the different structural manifestations of one functional unit, that is, relative clauses.

In traditional grammars, only two analytic tools are available for the description of syntactic relations holding between adjoined clauses, namely, co-ordination and subordination. Towards the end of the 20th century, the studies on the syntactic types of clause linkage have become more explicit in their recognition of the lack of clearly defined boundaries between the different structural representations of interclausal relations. The prevailing idea is now the existence of a cline ranging from minimal to maximal interdependence of the clauses combined, what Givón has defined as a "multi-

point graduated scale of clause integration" (1990: 826). Actually, the syntax of clause linkage is basically conceived of as offering a wide range of structural possibilities between two extreme options which indicate, respectively, the lowest and highest degrees of integration between clauses.

Foley and Van Valin (1984) distinguish different types of clause linkage relations relying on two fundamental parameters: dependence and embedding. They identify tree basic types of relation through the interplay of embedding and dependence across clauses:
- co-ordination, which involves neither dependence nor embedding;
- subordination, which involves both embedding and dependence;
- co-subordination, which involves dependence but not embedding.

And are illustrated as follows:

Co-ordination:
[- embedding, -
dependence]

Subordination:
[+ embedding, +
dependence]

Co-subordination:
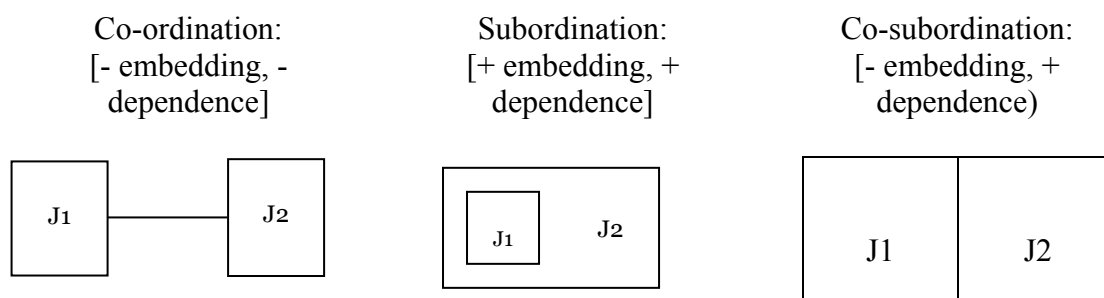[- embedding, +
dependence)

Figure 1: Three nexus types (Foley and Van Valin 1984: 242)

Although Foley and Van Valin qualitatively distinguish three basic types of clause linkage, these are prototypes, and their study implies the recognition of a gradient from least to most integration between the linked clauses.

More recently, Hopper and Traugott (2003: chapter 7) devote one chapter of their study on grammaticalization to the study of inter-clausal relations. Grammaticalization theory is of great relevance for the present study, as it is primarily interested in how syntax arises out of the use of language (discourse) in communicative settings (as opposed to generative syntax, which sees syntax as a basic schema of sentence formation) (Fischer 2007: 211). A similar view had already been held by Matthiessen & Thompson (1988), who hypothesize that clause linkage "in grammar has evolved as a grammaticalization of the rhetorical units in discourse defined by rhetorical relations" (1988: 301). Actually, since this study it has been widely assumed that discourse structure and complex sentence structure have much in common and that the latter is a more grammaticalized way of representing relationships between states of affairs than the former. This argument has been further –and more recently- developed by Hopper and Traugott, who consider that

> "from the point of view of language change, the initial formation of a complex clause involves the combining into one integrated structure of two separate and autonomous nuclei that are mutually relevant. The act of combining the clauses and signalling this combination linguistically is grounded in rhetorical production strategies. The new single structure becomes more complex in the process, because it now consists of two subparts. " (Hopper and Traugott 2003: 169)

As was the case of Foley and Van Valin, Hopper and Traugott also distinguish tree cluster points in the cline of clause combining, which, from left to right, indicate a

greater degree of integration between the clauses linked. The cline they establish can be roughly expressed as:

parataxis > hypotaxis > subordination

And each of these cluster points can be more thoroughly defined if we take into account two sets of features which had already been used in combination by Foley and Van Valin, namely dependence and embedding:
- Parataxis: the two (or more) clauses stand in nuclear position, which involves syntactic (not semantic or pragmatic) independence of the clauses linked [-dependent, - embedded];
- Hypotaxis: one of the clauses is a nucleus and the other(s) function(s) as syntactic margins, being relatively independent on the nucleus but not included properly in any constituent of the latter. This involves a relation of interdependence [+ dependent, - embedded];
- Subordination or, "in its extreme manifestation" embedding: one of the clauses is a nucleus and the other is a margin functioning as, or within, a constituent of the nucleus and begin, thus, an exemplification of complete dependence [+ dependent, + embedded].

In addition to these studies, Christian Lehmann (1988: 183) produced a typology of clause linkage which might be applied at different historical stages of a language. He intends to define the most important parameters to be analyzed in the formation of complex sentences. Such parameters are presented in the form of basically independent individual continua which correlate with each other only to certain extent and are by no means parallel in development. The interrelation of these continua yields a more general continuum ranging from maximal elaboration to minimal compression. The increasing degree of compression of the complex construction as we proceed rightwards in the cline is marked in the morphosyntax of the clause, as suggested by each of the individual continua. He proposed the following six relevant parameters for the analysis of clause linkage strategies:
- Hierarchical downgrading of the subordinate clause: parataxis > embedding.
- Syntactic level of the subordinate clause within the main clause: sentence > word.
- Desententialization of the subordinate clause: loss of clausal properties, such as illocutionally force, tense, aspect, mood, actants and circumstants. Not relevant to us here.
- Grammaticalization of the main verb: this parameter is not relevant to us here since it would lead us to the domain of simple clauses.
- Interlacing of the two clauses.
- Explicitness of the linking: syndesis > asyndesis

After revising these approaches, in my own view the syntax of interclausal relations is best regarded as involving a continuum of structural integration. Hopper and Traugott's view will be basically followed, and will be combined with Lehman's typology of clause linkage. We will recognize the e existence of a series of continua containing different structural possibilities defined on the basis of variable degrees of accomplishment of the relevant features. The continua we are referring to are potentially the same as those in Hopper and Traugott:

[- embedding] ……………………………………………………………………… [+ embedding]

[- dependence] …………………………………………………………………. [+ dependence]

Additionally, Lehmann's criterion on the degree of explicitness of the linking will be added in the codification of the relevant syntactic structures. According to this criterion, tighter cohesion is achieved by the introduction of more explicit grammatical marking, but also by the deletion of redundant elements, such as lexical repetitions and resumptive pronouns.

[- embedding] ……………………………………………………………… [+ embedding][1]

[- dependence] …………………………………………………………………. [+ dependence][2]

[+ explicitness of linking]…………………………………………. [- explicitness of linking]

These three continua function as diachronic tools, although it is relevant to emphasize that they might be (and some of them are) synchronically present.

2. Relative clauses

Although traditionally relative clauses are associated with subordination, this is far from being true. For the present research, I will follow Keenan and Comrie's functional definition of the concept 'relative clause':

> Any syntactic object to be a relative clause if it specifies a set of objects (perhaps a one member set) in two steps: a larger set is specified, called the *domain* of relativization*,* and then restricted to some subset of which a certain sentence, the *restricting* sentence, is true. The domain of relativization is expressed in surface by the *head NP* and the restricting sentence by the *restricting clause*, which may look more or less like a surface sentence depending on the language. [Original emphasis] (Keenan and Comrie 1977: 63-64)

The history of the English language has shown relativization by means of paratactic [-dependent, -embedded] structures and also hypotactic [+ dependent, + embedded] structures, in agreement with Matthiessen and Thompson, who reinforce Keenan and Comrie's definition in their consideration that "the function of a clause is not a necessary consequence of a particular syntactic pattern" (Matthiessen and Thompson 1988: 225).

---

[1] It roughly corresponds to Lehmann's 'hierarchical downgrading': parataxis > embedding.

[2] It roughly corresponds to Lehmann's 'syntactic level of the subordinate clause within the main clause': sentence > word. He conceives various levels: the lower the level of subordination, the more tightly the subordinate clause is integrated into the main clause.

3. Corpus description

The data for the present study have been extracted from *The Helsinki Corpus of English Texts: Diachronic and Dialectal*. Prose texts from late Old English (950-1050) and early Middle English (1150-1250) have been included in the analysis. Poetry was excluded, as it was thought that it would be affected by constraints of rhythm (e.g. alliteration, stress, etc.) and meter, and the results obtained from poetical texts were likely to be biased. Table 1 below sets out the number of words and tokens found per chronological subperiod.

Periods	Nr of words	Nr of tokens
O3 (950-1050)	67,390	354
M1 (1150-1250)	49,110	164
TOTAL	116,500	518

Table 1: Description of the corpus.

As can be seen from Table 1, the corpus consists of 116,500 words and includes 518 examples.

4. Analysis

In the analysis different criteria relevant for the degree of clause linkage of relative clauses have been applied.

5.1. Embedding

By embedding, we understand a process whereby a clause comes to be included within another clause. A clause could be [- embedded] and [+ embedded]. Following Lehmann (1988) and Hopper and Traugott (2003), this continnum ranges from paratactic structures to embedded ones (parataxis > embedding). In the present research, the criterion that we used to determine the degree of embedding of the relative clause is the position it occupies with respect to the main clause. In the periods of the English language under analysis, there are two main types of relative clauses:
- [- embedded] clauses: within [- embedded] clauses we have distinguished two syntactic structures:

   o extraposed relative clauses in which the relative clause is separated from the antecedent it resumes and moved to the end of the main clause:

(1) Ælce dæg we syngiað and ælce   dæg we sceolon urne Hælend$_i$ gladian mid sumre
    every day we sing   and every day we must   our Saviour gratify with certain
    godnysse, [se$_i$ ðe æfre wile   us mannum mildsian]
    goodness   rel   ever wants   us to men   show mercy
    'We sing every day and every day we must gratify our Saviour with certain goodness, who always wants us to show mercy to men.'

<div align="right">[Q O3 IR HOM AELFR15:47]</div>

o so-called left-dislocated relative clauses, through which the relative clause is displaced to the front of the main clause, together with the antecedent it resumes:

(2) Ac se$_i$ [þe (i) god onginneþ and on þon   þurhwunaþ oþ ende of his lifes], se
but he   rel   good begins and in this way continues until end of his life this
bið hal   geworden.
is   hale been
'But he who begins good and continues therein until the end of his life shall be saved.'

[Q O2/3 IR HOM BLICK2: 139]

- [+ embedded] relative clauses: in which the relative clause and the antecedent follow each other, and the relative clause is embedded within the main clause:

(3) Eower Fæder$_i$ [se$_i$ on heofenum is], wat   hwæs eow   þearf   biþ
Your Father rel in heaven is knows what for you necessary is
'Your Father who is in heaven knows what is needful for you.'

[Q O2/3 IR HOM BLICK2: 103]

These examples differ gradually on the parameter of [± embedding]. At the starting end of the pole, we come across examples of extraposition. The relative clause is on a very high syntactic level (example 1) and works as an apposition to the main clause. At a lower syntactic level, we come across examples of what has been labelled left-dislocation (example 2). We have a main clause, one element of which (normally the subject) is left dislocated in order to introduce the relative clause and not to embed it within the main clause. The lowest instantiation of syntactic embedding is illustrated by intraposed relative clauses, or relative clauses which are inserted within the main clause and are dependent on the head noun (normally a subject) (example 3). If we rank order these syntactic positions on the cline, we would come across the following:

[- embedding] …………………………………………………………………… [+ embedding]
Extraposition………………….………………LD[3]
……………………………………….Intraposition

The distribution in my corpus is shown in the following figure:

---

3 Although it appears in the middle, I consider that it is between the weaker and stronger extremes of embedding, not at any point in particular.
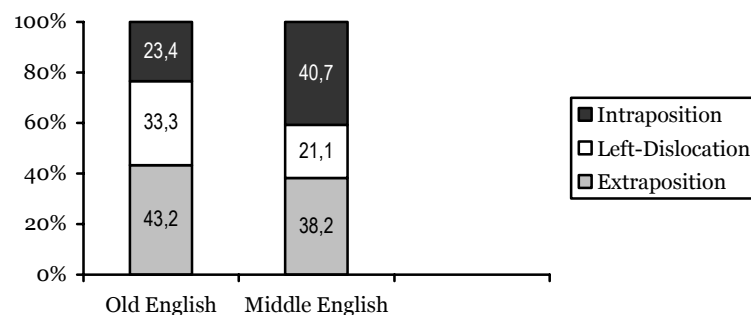
Figure 2: Degree of embedding in Old and Middle English

The results from my study reflect that this continuum works diachronically. There is an important increase in the degree of embedding of relative clauses through the axis of time. In Old English, [- embedded] structures are the most frequent option. In Middle English, [+ embedded] structures importantly increase.[4]

5.2. Degree of explicitness of the linking

We base this criterion on Lehmann (1988: 213). He identifies seven different types of grammatical marking of linkage between clauses, arranged from greater to lesser explicitness (syndesis > asyndesis):
1. anaphoric subordinate clause;
2. gerundial verb
3. prepositional phrase
4. connective adverb
5. specific conjunction;
6. universal subordinator;
7. non-finite verb form

Although Lehmann (1988) identifies 7 different types of linking, I will restrict this analysis to only three, which correlate with the three main types of relativizers used in the history of the English language. They have been arranged
in a continuum of decreasing explicitness:
- Anaphoric subordinate clause: deictic pronoun which refers back to an element of the preceding clause. This is illustrated by so-called relative pronouns *se/seo/þæt*, also used as demonstrative pronouns in Old English.[5] These relative pronouns agree in number and gender with the antecedent (and in case with the syntactic function of the relative clause they introduce):

(4) ðonne is þær on neaweste sum   swiþe   mære burh betwih þære sæ$_i$ on þæm
    then   is there   in nearness certain exceedingly famous city   towards the sea in the
    munte   Garganus geseted, [se$_i$ is haten Sepontus].

---

[4] This increase takes place at the expense of left-dislocated constructions which in Middle English suffer an important decrease (see Carkeet 1976).

[5] Actually in some cases it is not easy to decide whether it is a demonstrative pronoun or a relativizar (see Suárez Gómez 2005).

mountain Garganus set   rel is called Sepontus
'There is a very famous city nearby towards the sea [...] set in the hill Garganus, called Sepontus'

<div align="right">[Q O2/3 IR HOM BLICK 17: 18]</div>

- Specific conjunction: it is illustrated by the compound relativizer (*seþe*), composed of a highly deictic element (*se*) and the universal, unmarked subordinator *þe* (see example (1), and they always function as relativizers;

- Universal subordinator: a less explicit subordinator which introduces a relative clause but establishes no reference with the antecedent it resumes. This is illustrated by means of the so-called invariable relativizer *þe* and also *þat*:

(5) Forþon þære burge$_i$ nama [þe   is nemmed Gerusalem] is gereht sibbe
        gesyhþ,
    Because the   city   name rel   is called   Jerusalem means   of
peacesight
    forþon þe   halige saula þær   restæþ
    because the holy   souls there rest
    'For the name of the city which is called Jerusalem signifies 'sight of peace', because the holy souls   rest there'

<div align="right">[Q O2/3 IR HOM COBLICK 6: 25.1003]</div>

If we rank order these relativizers on a cline of [+ explicit] to [- explicit], we would come across the following:

[+ explicit] ………………………………………………………………………………………
[- explicit]
se……………....……………….…………………*Seþe*
……………………………………….*þe/þat*

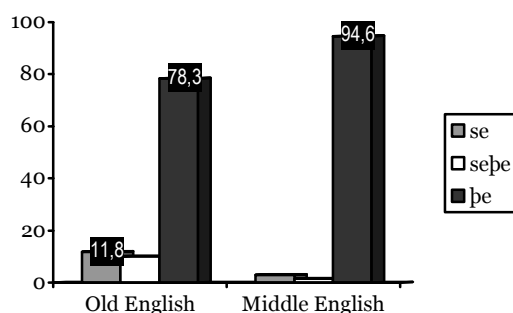The distribution in my corpus is shown in the following figure:

Figure 3: Degree of explictness in Old and Middle English

The results show that the variation in terms of the degree of explicitness of the linking is not very relevant in the two periods under analysis. Nevertheless, if we compare the two periods there is an important diachronic change. Although both Old and Middle English favour the use of [- explicit] linking devices, in Old English [+ explicit] devices are more frequently used than in Middle English.

5.3.   Dependence

By dependence we understand the degree of autonomy that the members of the clause complex have. Following Longracre's terminology (1985), we will refer to the 'nucleus' ([- dependent]) and 'margin' ([+ dependent]) of a clause complex. The possible combinations of clauses are the following:

- Nucleus + nucleus
- Nucleus + margin(/satellite)

At the starting pole of the continuum [-dependence], we come across examples which show a very high syntactic level between the two clauses forming the complex sentence. We have semantic/discourse/pragmatic dependence, but no syntactic dependence. Basically what we have is the combination of two nucleus or main clauses (example (4)). At the end of the pole, we come across examples which show a very low syntactic level, the relative clause being a margin, that is, subordinate and being downgraded to a particular, well-defined constituent within the main clause (example (5)).

Although the continua of dependence and embedding seem to run parallel, this is not necessarily so. We may come across [- dependent] structures, which might be embedded, as examples (3) illustrates.

In order to analyze this criterion, the previous two criteria were analyzed in combination. It is assumed that example (4), extraposed and introduced by relativizer *se* is [- dependent] than example (3), introduced by the same relativizer *se* but in an intraposed relative construction. This criterion might be enriched using further criteria (length, word-order, type of relative clause, etc.).

With the aim of determining the degree of dependence of the examples under analysis, we created an index combining the two criterion and possibilities within each where assigned a value:

- Degree of embedding:
  - Extraposition (1)
  - Left-Dislocation (0.5)
  - Intraposition: (0)
- Degree of explicitness:
  - *Se* (1)
  - *Seþe* (0.5)
  - *Þe* (0)

As a consequence, taking into account the existing possibilities, they were ordered in the continuum as follows:

[- dependency] ……………………………………………………………………………[- dependency]

Extraposed *se* (2)	Extraposed *seþe* /LD *se* (1.5)	Extraposed *se*/LD *seþe*/Intraposed *se* (1)	LD *þe*/ Intraposed *seþe* (0.5)	Intraposed *þe* (0)

Extraposed relative clauses introduced by *se* will illustrate the lowest degree of dependency and intraposed relative constructions introduced by *þe* the highest degree of

dependency. As Figures 4 and 5 illustrate, [- dependency] is higher in Old English and [+ dependency] in Middle English.
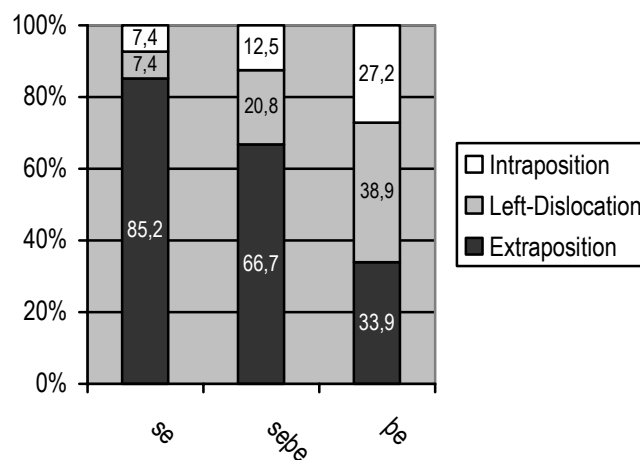

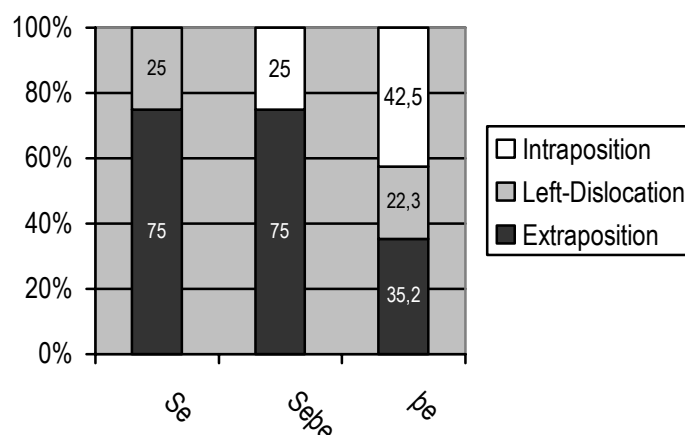Figure 4: Degree of dependence in Old English

Figure 5: Degree of dependence in Middle English

The comparison of Figures 4 and 5 shows us the diachronic evolution of relative clauses in the history of English. We observe a decrease of extraposed constructions introduced by *se* from Old to Middle English, but, what is more relevant, a very important increase of intraposed constructions introduced by *þe*. As clearly shown in Figure 3, invariable relativizer *þe* is the only productive relativizer in both periods and the one which provides more trustworthy information on the change we are dealing with. In Old English, intraposition was clearly disfavoured. In Middle English, by contrast, intraposition becomes the favoured option with this relativizer in particular.

5. Concluding remarks

How do these continua correlate? There is a very clear correlation between hierarchical downgrading and the syntactic level of the subordinate clause: the higher the degree of

dependence, the higher the degree of embedding, although, as has been demonstrated, it is not a one-to-one correlation. Explicitness of the linking also plays a role in the other continua. The lesser explicitness of the linking (*þe*), the higher the degree of dependence. Therefore, at the starting pole of the continuum we would come across [-dependent, - embedded, + explicit] constructions, as that illustrated by example (4), that is, by a combination in which two clauses and the linkage are maximally elaborated (independent and syntactically equal full clauses). At the opposite pole of the continuum we would come across [+ dependent, + embedded, - explicit] constructions, as that illustrated by example (7), that is, a main clause and a subordinate clause embedded on an element of the main clause on a low syntactic level. In the middle of these two poles we come across a variety of examples, such as (2) and (3) which are non-extreme in some (or all) of the parameters, result of the fact that the parameters do not run parallel. What is their effect on the diachronic evolution? The results from Figures 2 and 3 show that the evolution in the continua show an important diachronic evolution of relative constructions towards [+ embedding] and [- explicit] constructions. This is further reinforced by figures 4 and 5.

References

Carkeet. David 1976. Old English correlatives. An exercise in internal reconstruction, *Glossa* 10(1): 44-63.

Fischer, Olga 2007. *Morphosyntactic Change. Functional and Formal Perspectives*. Oxford: Oxford University Press.

Foley, William A. and Valin, Robert D. van 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

Givón, Talmy 1979. *On Understanding Grammar*. New York: Academic Press.

Givón, Talmy 1990. *Syntax. A Functional-Typological Approach*. Amsterdam and Philadelphia: John Benjamins.

Hopper, Paul and Elizabeth Closs Traugott 2003. *Grammaticalization*. Cambridge: Cambridge University Press.

Keenan, Edward L. and Comrie, Bernard 1977. Noun Phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8: 63-99.

Lehmann, Christian 1988. Towards a typology of clause linkage. In Haiman, John / Thompson, Sandra (eds) *Clause Combining in Grammar and Discourse*. Amsterdam and Philadelphia: John Benjamins, 181-225.

Longacre, Robert E. 1985. Sentences as combinations of clauses. In Shopen, Timothy (ed.) *Language Typology and Syntactic Description*. Complex Constructions. Cambridge: Cambridge University Press, 234–286.

Matthiessen, Christian and Thompson, Sandra A. 1988. The structure of discourse and subordination. In Haiman, John / Thompson, Sandra A. (eds.) *Clause Combining in Grammar and Discourse*. Amsterdam and Philadelphia: John Benjamins, 275-329.

Suárez Gómez, Cristina 2002. *Se*: determiner or relativizer. In Palacios, Ignacio et al. (eds.) *Fifty years of English studies in Spain (1952-2002). A commemorative volume*. Santiago de Compostela: Universidade de Santiago de Compostela, 329-337.