Revista
Signos
Estudios de Lingüística

SciELO *Chile*
Scientific Electronic Library Online

SCOPUS
Clarivate
Analytics

WEB OF SCIENCE™

Anaphoric Encapsulation, Text Information Structure and Discourse Topicality

Encapsulación anafórica, estructura informativa textual y topicalidad discursiva

Margarita Borreguero Zuloaga
UNIVERSIDAD COMPLUTENSE DE MADRID
ESPAÑA
mbzuloag@filol.ucm.es
https://orcid.org/0000-0003-1992-0766

Abstract

The contribution of anaphoric encapsulators to the information structure of the utterance has often been restricted to observations regarding their position as sentence topics (theme) or as part of the comment (rheme). However, they may appear in other positions that can be better delimited with the help of a model of discourse segmentation such as the Basel Model. I explore the different discourse functions of anaphoric encapsulators in the three main units delineated in this model: the Nucleus or propositional core of the utterance, the Frame or the information chunk preceding the Nucleus, and the Appendix or slot for background information. Besides, I claim that the possibility for an anaphoric encapsulator to express the discourse topic of (part of) a text is intimately related to its position in the Utterance and propose a tentative scale of access to discourse topicality for this type of anaphors.

Keywords: anaphora, encapsulation, information structure, discourse topic, Basel Model.

Resumen

El análisis del papel que juegan los encapsuladores anafóricos en la estructura informativa de los enunciados se ha limitado por lo general a observaciones relativas a la posición que ocupan como temas o bien como parte del rema de los enunciados. Sin embargo, estas anáforas aparecen en otras posiciones enunciativas que pueden delimitarse con la ayuda de un modelo de segmentación del discurso como el Modelo de Basilea. En este trabajo se exploran las diversas funciones discursivas de los encapsuladores anafóricos en las tres unidades principales establecidas por este modelo: el Núcleo o relación proposicional central del enunciado, el Marco o segmento informativo que precede al Núcleo, y el Apéndice o hueco estructural para la información secundaria. Además, se defiende que la posibilidad de que un encapsulador

anafórico exprese el tópico discursivo de (parte de) un texto está íntimamente ligado a su posición en el Enunciado y se propone una escala tentativa de acceso a la topicalidad discursiva para este tipo de anáforas.

Palabras clave: anáfora, encapsulación, estructura informativa, tópico discursivo, Modelo de Basilea.

INTRODUCTION

No NP is an anaphoric encapsulator on its own. Encapsulation is not a formal property of NPs—although some formal characteristics are compulsory to speak of encapsulation—but a discursive function fulfilled by particular NPs in particular texts. The possibility of delimiting a discourse segment (antecedent), reifying it as a process, event or speech act, and presenting it as a conceptual object and a new referent makes this type of NPs not only one of the fundamental mechanisms to build the network of textual semantic isotopies by guaranteeing referential continuity, but also a text-structuring device at the level of text information structure and discourse topicality.

Anaphoric encapsulators (henceforth AEs)[1] have been defined as NPs with a fairly rigid structure:[2]

- a determiner which establishes an anaphoric relationship to previous information in the text (definite article, demonstrative or third-person possessive)[3];
- a noun which belongs mainly to one of these four types:
  a) a deverbal noun formed from the lexical root of a verb whose valency structure is eventually transformed in a series of nominal complementizers; the NP with this type of nominal core is called nominalization and has the capability of transforming events in individual entities (Ferrari, 2002; Méndez García de Paredes, 2003; Azpiazu, 2004);
  b) a general noun or hypernym, such as 'fact', 'thing', 'event', with a high degree of desemantization and as a consequence the possibility to refer to complex information (Halliday & Hasan, 1976; Pelo, 1986; Vignuzzi, 1986; Swales, 2001);
  c) a noun designating an abstract, non-physical entity, such as 'fight', 'deceit', 'rise';
  d) any noun that usually refers to first-order entities but can be metaphorically interpreted to refer to processes, events, or actions (including speech acts; Llamas Saíz, 2010b; López Samaniego, 2014; Pecorari, 2021), such as 'door', 'tunnel', 'path', 'wink';
- optionally, adjectival and prepositional modifiers.

Their antecedents[4] in discourse, codified through at least one predicative relationship in a NP (Mortara Garavelli, 1971; González, 2008), an utterance or a

MARGARITA BORREGUERO ZULOAGA

sequence of utterances («proposition-like pieces of information», Schmid, 1998: 4),[5] are not individual or first-order entities, but what Lyons (1977) calls higher-order entities: either (a) processes, situations, events or acts (second-order entities), or (b) concepts, abstract entities, speech acts[6] (third-order entities) (Conte, 1998; Lala, 2010b; Izquierdo Alegría & González, 2013b; Pecorari, 2015c).

This means that the antecedent is not a referring expression (López Samaniego, 2013; Pecorari, 2015c)[7], but a discourse entity, i.e., a mental representation elaborated through the cognitive anchoring of a process in spatio-temporal circumstances, and therefore there is no co-reference between the AE and its antecedent. In fact, AE are typically non-co-referential anaphors.

The specificity of this type of NP is its multifunctionality as a cohesive device. On the one hand, their anaphoric nature allows them to link an utterance to a previous one (or to a sequence of utterances) by fulfilling a summarising function of part or the totality of its content; on the other hand, their usual thematic position contributes to the textual dynamism by offering an anchorage for new predications and fostering the shaping of discourse topics further developed in the text (García Negroni, Hall & Marín, 2005; López Samaniego, 2010; Izquierdo Alegría & González, 2013b). Due to this double look at what has previously been said and to what comes next, the discourse function of AEs has metaphorically been conceived as a text-structuring mechanism acting as a door hinge (Borreguero Zuloaga, 2006).

What I would like to question in this paper is a simplistic view on the contribution of AEs to text information structure, usually conceived within the limits of thematic-rhematic progression scheme. I agree that AEs play a key role in the information structure of discourse, but I think their contribution to discourse topic organization needs to be carefully analysed. My hypothesis is that the function of AEs at the information structure level is not always to constitute a new theme to anchor the development of new information (Borreguero Zuloaga, 2006, 2018) or to be part of the rheme of the main proposition of the utterance to convey an explicit evaluation (López Samaniego, 2014; González & Izquierdo Alegría, 2020). As a matter of fact, AEs are not always part of a thematic-rhematic progression scheme (Daneš, 1974), but may fulfil other functions according to their position in the utterance and in the text. Moreover, it must be taken into account that information functions are strictly linked to the organization of discourse topicality.

Notwithstanding the importance of AEs at the information level, this aspect has been somewhat neglected in previous studies (with remarkable exceptions such as Conte, 1998; López Samaniego, 2014; Pecorari, 2015a, 2015c; González & Izquierdo Alegría, 2020). Most of the extensive research conducted about this mechanism has prioritised its function at the argumentation construction of texts, particularly the possibility of conveying axiological content in the nominal core and/or in the

adjectival modifiers (González, 2008; Lala, 2010a; Izquierdo Alegría & González, 2013a, 2013b; Llamas Saíz, 2010a, 2010b) and therefore the reconceptualization of the antecedent to create a particular persuasive effect.[8]

The role played by AEs in text structure strives on three fundamental factors: a) the reification of the propositional content of a previous utterance (or utterances); b) the resulting establishment of a new discourse referent, as AE is not a co-referential anaphor, this operation has been called 'hypostasis' by Conte ([1996]1999) or 'hypostatization' by Lyons (1977);[9] c) its position in textually relevant places to contribute to the semantic hierarchy of discourse topics, typically the first utterance of a paragraph:

> «Le incapsulazioni anaforiche possono quindi comparire in punti nodali dell'archittetura testuale e contribuire in modo significativo all'organizzazione e alla strutturazione del testo» (Conte, 1998: 158).[10]

In (1) the series of processes presented in the first paragraph is reificated as 'a scenario' in the second paragraph. A new referent (*este panorama* 'this scenario') is then constructed and introduced by the writer. Its position at the beginning of the following paragraph is not casual: it offers a resumption of what has been previously said and gives way to a new discourse topic, which is better interpreted in the frame established by the AE.

(1) <u>El recurso al ahorro se irá agotando, los tipos de interés encarecerán las hipotecas y restarán renta a millones de hogares, las medidas del Gobierno son temporales y antes o después se irán retirando; la estabilización del empleo sigue teniendo recorrido, pero es necesariamente finita</u>, etcétera.
Pues bien, ante **este panorama**, Comisiones Obreras considera prioritario para la clase trabajadora y para el país retomar un acuerdo ambicioso y valiente (Sordo, Unai, "Propuesta para alcanzar un acuerdo salarial", *El País*, January 18th, 2023, p. 21).
'The recourse to savings will be exhausted, interest rates will make mortgages more expensive and will reduce the income of millions of households, the Government's measures are temporary and sooner or later will be withdrawn; the stabilisation of employment still has a chance, but it is necessarily finite, etcetera.
Faced with **this scenario**, Comisiones Obreras considers it a priority for the working class and for the country to take up again an ambitious and courageous agreement.'

In the following, in order to explore my hypothesis, I first propose to correlate the functions of AEs at the information level to their utterance positions, and then to describe their contribution to the organization of discourse topics and subtopics. All my examples are taken from national Spanish journalistic texts published in the years

2000-2023. They have been manually selected by the author, except when otherwise indicated (see section 2.2.).

Utterance positions are accurately described when utterances are segmented into discourse units that maintain hierarchical relationships among them. In the field of Romance linguistics, several models of discourse segmentation have been proposed since the 1990s [for an overview of models of discourse segmentation applied to Spanish oral and written texts, Pons & Salameh (2021); Borreguero Zuloaga (2021), respectively; for a broader view on models applied to other Romance languages, Pons (2014); Borreguero Zuloaga, Atayan & Grosse (2018)]. In this paper, the utterance segmentation will be conducted in the frame of the so-called Basel model (Ferrari, Cignetti, De Cesare, Lala, Mandelli, Ricci & Roggia, 2008; Ferrari & Borreguero Zuloaga, 2015), that will be presented in section 1, a model which has already been applied to the study of Italian AEs (Lala, 2010b; Pecorari, 2015a, 2015c).

In section 2, I will examine the information functions of AEs when they occupy the three utterance positions described in the Basel model as Nucleus, Frame, and Appendix. These functions are described taking into account the contribution of AEs to the organisation of discourse topicality (Brown & Yule, 1983, 1993), i.e., the way in which AEs are part of the strategies regarding discourse topics such as introduction, continuation, closure, and framing (Charolles, 2020). In section 3, the focus is on AEs functioning as discourse topics (DT) and how their position in the utterance may facilitate the projection of an AE into a higher level of discourse organization. Two fundamental dimensions of textual building are related here: the network of semantic isotopies and the assemblage of different information levels that constitutes the utterance. The last section is the conclusion, which presents a summary of the factors that influence the processing of AEs according to previous studies and suggests that utterance position and access to the level of discourse topicality could be added to the list of those factors.

1. Model of discourse segmentation for written texts: the basel model

The Basel model focuses on the segmentation of discourse units and their hierarchical relations and it is inspired by previous proposals to segment spoken discourse into prosodic units separated by prosodic breaks that convey a complete message, called utterances [for French, Blanche-Benveniste (1990); for Italian, Cresti (2000)]. According to Blanche-Benveniste (1990), each utterance is articulated in several prosodic units, but only one is obligatory, the so-called nucleus (*noyau*); other units may precede (*préfixe*) or follow (*suffixe, postfixe*) the nucleus. In written texts, prosodic segmentation has been replaced by other criteria such as punctuation, syntactic relations, and illocutionary force (Ferrari et al., 2008; Ferrari, 2014).

The Basel model proposes a top-down segmentation process: complex units are divided into simpler ones according to their contribution to the information structure. The Paragraph[11] serves as a starting point, since it usually consists of a sequence of Utterances linked by coherence and cohesion relations, shows thematic unity in terms of the discourse topic, and is often delimited by formal opening and closing devices, such as typographical marks (indentation, spacing before and after the paragraph) and linguistic elements (discourse markers indicating a change of topic or subtopic, AEs and textual anaphors referring to what was said in previous paragraphs). Paragraphs are then divided into Communicative Units defined by semantic and pragmatic criteria. These units have an illocutionary function—typically assertive in argumentative and expository writing—and a textual function that defines their role in the text: explanation, exemplification, concession, topic elaboration, etc. (Ferrari et al., 2008). Communicative Units are interrelated in three independent dimensions:

- the thematic or referential dimension: how the text evokes the extralinguistic world it refers to, i.e., the cognitive activation of discourse referents along the text and the relations between them;
- the logical dimension: logical and argumentative relations between propositions and speech acts, usually (but not always) marked by connectives;
- the polyphonic dimension: different points of view are indicated by a number of linguistic elements (e.g., first-person deictics (verbal forms, pronouns, possessives) vs. third-person verbal forms and NPs indicate the shift from the voices of the participants to the voice of the narrator).

The meaning of Communicative Units is partly linguistically codified (in the Utterance) and partly implicit. The relationship between codified and implicit meaning, based on common ground (information from the communicative situation and shared knowledge of the participants), sets in motion the inferential processes that allow the reader to achieve a complete and deep understanding of the author's ultimate communicative intention. The Utterance (U) has no predefined linguistic form: it can be a simple or complex sentence, a noun, as in vocatives, or any other lexical word.[12] The main criteria for its delimitation are the linear arrangement of linguistic elements between strong punctuation marks (period, colon, semicolon), the dependent syntactic relations between its components, and a homogeneous illocutionary force.

In order to segment a Paragraph into Utterances (marked by a double slash and numbered: U1, U2), different linguistic criteria are considered. In (2) three of them are at work: the strong punctuation marks, which is the most decisive criterion and most times a necessary and sufficient one; the cohesion relations created by the coreferential elements (the NP *las series juveniles* 'youth series' in U1 and the tacit pronoun—personal pronoun in the English translation—in U2); the change in the

temporal frame opening of each Utterance (Charolles, Le Draoulec, Péry-Woodley & Sarda, 2005): *durante años* 'for years' vs. *de repente* 'suddenly'.

(2) // Durante años, **las series juveniles**ᵢ, con sus historias de instituto, fueron una constante en la televisión española. //ᵤ₁ // De repente, **[Ø]**ᵢ desaparecieron de las cadenas en abierto. //ᵤ₂ (N. Marcos, "Adolescentes más allá del sexo", *El País*, August 28th, 2023).

'//For years, youth seriesᵢ, with their high school stories, were a constant feature of Spanish television // // Suddenly, theyᵢ disappeared from free-to-air channels //'

Utterances can be analysed from different perspectives: a) the cognitive activation of referents, b) the topic-comment structure, and c) the information structure. On the latter level, the Basel model distinguishes between a main and a secondary level, i.e., a foreground and a background level. The result of segmenting an Utterance into Information Units is shown in Table 2.

Table 1. Discourse units in the basel model (Borreguero Zuloaga, 2014: 351)

Paragraph				
Communicative Units				
Codified meaning → **Utterance**				Implicit meaning
Main level		Secondary level (Parenthetical comments)		
Information Units				
Frame	**Nucleus**		**Appendix**	
(Topic)	Topic	Comment		

The Utterance can be divided into three Informational Units, but only the Nucleus, as already mentioned, is obligatory. The Nucleus defines the illocutionary force, i.e., the type of communicative action that motivated the illocutionary act, in addition to the textual function of the whole Utterance, i.e., its specific contribution to the textual composition (Ferrari et al., 2008). In (3) the whole U1 is a Nucleus:

(3) // /Cristina Navajas fue secuestrada por policías vestidos de civil en julio de 1976./ᴺᵘᶜˡᵉᵘˢ // (C. Lambertucci, "«Es una derrota de la dictadura militar»", *El País*, July 29th, 2023)

'Cristina Navajas was abducted by plain-clothes police officers in July 1976.'

The Nucleus, if it takes the form of a proposition, can be articulated in a Topic-Comment information structure; otherwise, it can take the form of a completely new information structure (for example, in the case of existential or presentative structures: *Hay demasiada gente en esta habitación* 'There are too many people in this room'). The Topic is defined in this model as a conceptual object that denotes a discourse referent and what the sentence is about.[13] Conversely, the Comment is an informational

element that expresses a predicative relation attributed to the Topic (Ferrari et al., 2008). The Nucleus in (3) is articulated in Topic and Comment as follows:

(3') // /[Cristina Navajas]_Topic [fue secuestrada por policías vestidos de civil en julio de 1976]_Comment./_Nucleus //

In complex Utterances it is possible to have more than one Nucleus, typically in the case of juxtaposed and coordinated clauses like in (4):

(4) // /Dos ejemplares de pigargo, el águila más grande de Europa, realizan un vuelo majestuoso sobre un río./_Nucleus //_U1 /Se persiguen,/_Nucleus1 /juegan en el aire/_Nucleus2 //_U2. Uno se posa, grácil, en la rama de un árbol y el otro le sigue al instante. (M. A. Medina, "Águilas gigantes en un limbo de Asturias", *El País*, June 7th, 2023)
'Two white-tailed eagles, the largest eagle in Europe, perform a majestic flight over a river. They chase each other, they play in the air. One perches gracefully on the branch of a tree and the other follows in an instant.'

The Nucleus may be preceded or followed by other units. The Frame Unit always precedes the Nucleus and its function is to delimit the semantic-pragmatic domain that defines the relevance of a Nucleus (Zampese, 2005; Ferrari et al., 2008); it can extend over several Nuclei. Frame units can refer to spatio-temporal, causal or final circumstances (cf. temporal information in the Utterances in (2): *durante años* 'for years' and *de repente* 'suddenly' are Frames of their respective Utterances), sources of information, or any other epistemic or evidential information. They can also contain connectives that link an Utterance to previous ones, or hanging Topics that are not related to their Comments (Borreguero Zuloaga, 2014). The most common punctuation mark separating a Frame from the Nucleus (or Appendix) is the comma.

Both the Nucleus and the Frame may be followed by an Appendix Unit, which enriches and modulates the meaning of the Units, but whose contribution to the overall meaning of the text is not as relevant, because the information conveyed only affects the previous Information Unit. Appendices usually contain modal elements related to the subjectivity of the author, evaluations or personal opinions, or clarifications, such as descriptions of the referents mentioned in the Nucleus or Frame. The Appendix can be intensive, i.e., inserted within a Nucleus, or extensive, i.e., inserted after the Nucleus, a position suitable for locating complex information. They can also follow a Frame or another Appendix. The text in (5) offers an example of two types of Appendixes: Frame's Appendix and Nucleus' Appendix:

(5) // /Lejos de las cámaras,/^Frame /donde impera ese mensaje optimista y de llamada a la unidad,/^Appendix1 /el presidente,/^Nucleus- /a solas con sus ministros,/^Appendix2 /fue bastante claro:/^-Nucleus //_U1 (C. E. Cué, "Sánchez se

rodea de empresarios para pedir unidad y presionar al PP", *El País*, August 28th, 2020).

'Away from the cameras, where the message of optimism and a call for unity prevails, the president, alone with his ministers, was quite clear:'

Following the Frame *Lejos de las cámaras* 'Away from the cameras', a relative proposition conveys background information already known to the reader (previously in the text the focus was previously on this optimistic message). This relative proposition constitutes an Appendix of the Frame (Appendix 1) because it just recalls previous information about the scenario "in front of the cameras". On the other hand, the Nucleus is interrupted by a second Appendix (Appendix 2), offering information about the circumstances in which the action takes place (alone with his ministers). This is an intensive Appendix of the Nucleus.

In the following section, we will see that AEs may occupy different positions in the Utterance corresponding to these three Units (Nucleus, Frame, and Appendix) and that their position is a key factor to describe their contribution to the Utterance information structure.

2. Anaphoric encapsulators and utterance position

Following the segmentation of the Utterance in Information Units proposed by the Basel model, I will present in this section how the presence of AEs in the three main Units (Nucleus, Frame, and Appendix) contributes to the text information structure. Two characteristics of these NPs guarantee their role as text-structuring mechanisms at the referential dimension: on the one hand, the reification and hypostasis operations carried out by AEs bring about new referents (emerging from given information); on the other hand, the capacity to encompass extensive text passages favours these referents to become discourse topics, as we will see in section 3.

2.1. Anaphoric encapsulators in the Nucleus Unit

We may say that the prototypical position of AEs is as Topic in a Nucleus Unit with propositional structure.[14] Their encompassing nature, i.e., their capacity to summarize complex propositional content in just one NP, and the constitution of a new conceptual object make them prime candidates to occupy the Topic position (Ferrari, 2002).[15] When fulfilling a Topic function, they usually occupy the initial position of the Nucleus (which does not necessarily coincide with the initial position of the Utterance) and establish a relationship of aboutness with the Comment, a predicative structure that conveys information about the Topic (Ferrari et al., 2008). The Comment will then express «information which is relevant to and which increases the addressee's knowledge of this referent» (Lambrecht, 1994).

The new referent is semantically enriched by the co-text to enhance its communicative relevance. In (6) the propositional content of U1 describes a series of acts (clapping, whistling, booing) performed by the audience at a tennis match; these acts are summarised in the NP *la historia*,[16] which occupies the Topic position in the Nucleus of the Utterance immediately following U1, and is further developed in the Comment: it is added that this behaviour occurs each time that a French player is at the tennis court.

(6) // <u>Se aplauden los errores del contrario, se silba y se abuchea cualquier mínimo gesto de árbitros o jugadores que ponga en entredicho o anule un punto de la estrella local</u>.//$_{U1}$ // /[**La historia**]$_{Topic}$ se repite cada vez que entra en pista un o una tenista francesa./$_{Nucleus}$ //$_{U2}$ Una suerte de nacionalismo con tintes *hooliganescos* se impone cada vez en el deporte de la raqueta, famoso hasta no hace mucho per el *fair play* y la buena educación (L. Galán, "Los forofos invaden Roland Garros", *El País,* June 7th, 2023)

'The opponent's mistakes are applauded, and any gesture by the umpires or players which calls into question or cancels out a point by the local star is whistled and booed. History repeats itself every time a French tennis player enters the court. A kind of nationalism with hooliganism overtones is imposed each time in the sport of racquetball, famous until not so long ago for fair play and good manners.'

By recategorizing the antecedent as a new referent and placing it at the Topic position, the mental representation of the discourse entity remains active and available to the reader (López Samaniego, 2013), and the cognitive effort to link the AEs to the antecedent is minimised, notwithstanding the fact that the selection of the nominal core may have no lexical or semantic relationship with the information in U1. When the AE encapsulates a whole proposition or a whole Utterance (like in 6), we are faced with a typical schema of global thematic progression[17] (Topic progression in the terms of the Basel model, Ferrari et al., 2008). Although in this scheme the previous Utterance contains usually a predicative relationship between a Topic and a Comment that is to become the Topic of the next Utterance, it is also possible (like in 6) that Utterances that are not articulate in Topic-Comment, the so-called all-new or thetic Utterances, are resumed by the Topic of the following Utterance. In any case, global thematic progression is a mechanism that guarantees referential continuity (Ferrari, 2002).

The AE in Topic position may enter later a different type of progression, the so called 'constant Topic progression' by which a Topic establishes a coreferential relationship to a previous Topic. Its Utterance-initial position and the coincidence of Topic and syntactic subject function makes the AE particularly prominent and explains why it is so easily linked to a new anaphor: «an antecedent in grammatical subject position in a canonical active sentence would be very prominent and thus in

MARGARITA BORREGUERO ZULOAGA

(psychological) focus» (Cowles & Garnham, 2005: 735). The continuation of the text in (6') is as follows:

(6') "Divertirse", "pasarlo bien", son eufemismos para denominar la conducta desconsiderada y, a menudo, intimidatoria, de amplios sectores de las gradas. // /[**El fenómeno**]_{**Topic**} no es nuevo./_{Nucleus} //_U // /Hace tiempo que [Ø]_{**Topic**} viene ocurriendo/_{Nucleus} //.

"Having fun", "having a good time" are euphemisms for the inconsiderate and often intimidating behaviour of large sections of the stands. The phenomenon is not new. It has been going on for a long time.'

La historia is now substituted by the co-referential NP *el fenómeno* 'the phenomenon' and by a tacit pronoun in two subsequent Utterances whose Comments further develop the same Topic: 'it is not new, it has been going on for a long time'.

As has been observed in previous studies (Borreguero Zuloaga, 2006; López Samaniego, 2014; González & Izquierdo Alegría, 2020), the Topic position presupposes information that is cognitively more accessible than the one in the Comment, either because it has been previously introduced in the co-text, it is part of the visual input in a communication or belongs to the shared encyclopaedia, which is often the case in journalistic texts. This favours the presence of axionyms and evaluative modifiers in AEs occupying this position, because the evaluative content goes more easily unnoticed in a position usually reserved for given or old information (Pecorari, 2015b). In fact, the relationship between the AE and its antecedent is not direct, as in the case of coreferential NPs, but mediated by inferential processes that lie on the semantic information retrieved from the co-text and the encyclopaedic knowledge.

Furthermore, the syntactic form of the AE, in particular the type of determiner, also plays a role as it has been hypothesized by Ariel (1988) and adapted to Spanish by Figueras (2002): demonstrative premodifiers guarantee a more straightforward relationship with the antecedent and, especially in the case of axiological and metaphorical AEs, are almost compulsory (Dam, 2014; Pecorari, 2016). Let's compare examples (7) and (8).

(7) Como tenía por costumbre, el policía Ignacio Pérez Álvarez había acudido el 30 de enero de 1990 al bar Los Claveles, de Galdako (Bizkaia, 29.285 habitantes) a tomar el aperitivo. <u>A la salida, pocos minutos antes de las tres de la tarde, ETA lo estaba esperando para matarlo.</u> Tenía 39 años, esposa y tres hijos de 14, 10 y 5 años.

// /[**Ese crimen**]_{**Topic**}/^{Nucleus} – en el que la banda usó una bicicleta bomba por primera vez – /ha engrosado el listado de atentados sin resolver durante más de tres décadas./^{Nucleus} //_U (J. J. Gálvez, "Condenada la exetarra Guisasola por el asesinato de un policía en 1990", El País, June 6th, 2023)

'As was his custom, on 30 January 1990, policeman Ignacio Pérez Álvarez had gone to the Los Claveles bar in Galdako (Bizkaia, population 29,285) for an aperitif. When he left, a few minutes before three o'clock in the afternoon, ETA [terrorist group in the Basque country] was waiting to kill him. He was 39 years old, had a wife and three children aged 14, 10 and 5.

That crime—in which the gang used a bicycle bomb for the first time—has swelled the list of unsolved attacks for more than three decades.'

(8) "Si al final de mi mandato todos los brasileños tuvieran la posibilidad de de desayunar, almorzar y cenar, habré cumplido la misión de mi vida. (…) Mitigaremos el hambre, crearemos empleo, atacaremos el crimen, combatiremos la corrupción y crearemos mejores condiciones de educación para educación de bajos ingresos desde el comienzo mismo de mi Administración. (…) Los tiempos de prueba que enfrenta Brasil requieren austeridad en el uso del dineo público y una lucha implacable contra la corrupción".[18]

// /[Esa utopía factible]$_{Topic}$ cuatro años después,[sic] aporta un resultado agridulce:/^{Nucleus} //$_U$ (J. Estefanía, "Brasil, ¿qué ha sido de aquel sueño?", *El País*, September 25th, 2006)

' "If at the end of my mandate all Brazilians have the possibility to have breakfast, lunch and dinner, I will have accomplished my life's mission (...) We will mitigate hunger, create jobs, attack crime, fight corruption and create better education conditions for low-income education from the very beginning of my administration (...) The testing times Brazil is facing require austerity in the use of public money and a relentless fight against corruption".

That feasible utopia four years later brings a bittersweet result:'

In (7) the AE is constituted by a second-order demonstrative *ese* (indicating an intermediate grade of distance with respect to the deictic centre) and a non-deverbal eventive noun (*crimen* 'crime'). The presence of the verb *matar* 'to kill' in the antecedent, belonging to the same semantic field, allows for an easy inferential process to link the AE to its antecedent. Conversely in (8), the AE contains an abstract noun (*utopía* 'utopia') accompanied by a paradoxical adjectival modifier (*factible* 'feasible'), because utopias are by definition no feasible states of affairs. In this case, the terms used to encapsulate the Brazilian prime minister's discourse demand a higher cognitive effort (and a wider lexical knowledge) to link the AE to its antecedent. González and Izquierdo Alegría (2020) claim that discursive labels (highly informative AEs), notwithstanding their new way of recategorizing the referent, are not an obstacle to retrieve the antecedent. Nonetheless, this is not always the case, as it can be observed in (8), especially if the interpretation of the AE implies the activation of encyclopaedic knowledge which is not always available for a part of the readership. The presence of

MARGARITA BORREGUERO ZULOAGA

the demonstrative is one of the factors that allows for the correct interpretation. The other one is the Utterance position.

In fact, their place in the Utterance's information structure eases the interpretation and helps the achievement of the correct inferential process. Both AEs occupy exactly the same position: they are Topics of their respective Nuclei; they occupy the first position in the Nucleus (and in the Utterance) and also the first position in a new Paragraph (I will come back to this later on).

As we will see in section 3, this is the most determinant position for discourse topics (DTs). However, the degree of contribution to the global structure of discourse topicality decreases when the AE is part of the Topic but does not exactly coincide with the Topic. A lower degree corresponds to AEs in the Comment (see 9), a position from which it is difficult to become a DT and with which I will not deal here, although it is a very interesting position for the text argumentative structure because evaluative content is particularly transparent in this position (López Samaniego, 2014; González & Izquierdo Alegría, 2020).[19]

(9) En una conferencia de prensa, <u>el líder ruso afirmó de nuevo que no rechaza conversaciones de paz sobre Ucrania</u>. "No las rechazamos; para que ese proceso comience es necesario que haya acuerdo por ambas partes", afirmó en la noche del sábado. // /Ya el viernes,/$_{Frame}$ /[Ø]$_{Topic}$ [había reiterado **esa predisposición al diálogo** ante los líderes africanos]$_{Comment}$/^{Nucleus1} y /había culpado a Kiev de la imposibilidad de llevarlo a cabo/^{Nucleus2} // ("Rusia reforzará su flota con 30 nuevos buques este año", *El País*, July 31st, 2023)
'At a press conference, <u>the Russian leader again affirmed that he does not reject peace talks on Ukraine</u>. "We do not reject them; for such a process to begin there needs to be agreement on both sides," he said on Saturday night. Already on Friday, he had reiterated **this willingness to engage in dialogue** with African leaders and blamed Kiev for the impossibility of holding such talks.'

To sum up, when the AE introduces a new referent as Topic, it contributes to a more economical interpretation and enhances the text coherence in an optimal way. As a matter of fact, the Nucleus' Topic is the only Topic in the text that actively contributes to the thematic progression (Pecorari, 2015c), unlike Topics that are found in Frame and Appendix Units.

2.2. Anaphoric encapsulators in the Frame Unit

Elements placed in the Frame position play a fundamental procedural and cognitive role: a) they regulate the processes of knowledge mobilization required for interpretation of relations between propositions and b) distribute propositional contents into homogeneous blocks or chunks (Charolles et al., 2005). For this reason,

connective (anaphoric and non-anaphoric) devices are usually allocated in the Utterance's Frame.

Given the cohesive nature of AE, it is to be expected that this type of NPs frequently appears in this position with the general function of delimiting the validity of the Nucleus content (González & Izquierdo Alegría, 2020).[20] Unlike the Nucleus, which has mainly two internal structures, either propositional (Topic-Comment structure) or all-new information (as in presentative sentences), the Frame Unit displays a vast array of possibilities:

a) a parenthetical connective such as *sin embargo* 'however', *en cualquier caso* 'anyway', etc.;
b) a Topic-Comment structure in temporal, causal and final subordinate clauses with finite verbal form;
c) an implicit subordinate clause with non-finite verbal form in the Comment and a zero anaphor as Topic;
d) a Prepositional Phrase with different degrees of grammaticalization and conveying a variety of meanings, from evidentiality (*al parecer* 'apparently') to spatial and temporal conditions (*en 1981* 'in 1981', *en el vecino parque* 'in the nearby park').

In the following, I will analyse the position of AEs in the Frame Unit in just one type of structure: as part of a prepositional phrase.

2.1.1. Prepositional Phrases in the Frame Unit

When they occur in the Frame Unit, AEs are often part of a Prepositional Phrase (PP), either simple (Preposition + AE) or complex (Preposition + Noun + Preposition + AE). The instruction conveyed by the preposition in the case of simple PP or by the noun in the case of complex PP is similar to the one conveyed by adverbial connectives: they express causality, consequence, finality or adversativity. In this position they develop a stronger cohesive function that can be assimilated to that of a connective (in fact, many connectives grammaticalized from constructions containing anaphors, e.g., *de ahí que*). Prandi (2006) speaks of relational AE, because through AEs in this position the antecedent establishes a logical relationship to the core information of the new utterance, indicating the cause, the consequence or a hindering for the state-of-affairs reported in the Nucleus (Montolío, 2013):

> «One of the main textual functions of the Frame Unit is to guarantee the referential continuity of the text, through the use of anaphora; when the Frame Unit is filled by an adverb-like expression containing a relational encapsulator […], referential continuity is thus associated and strictly intertwined with logical coherence. The Frame Unit gives more prominence to the logical relation itself, objectified by the anaphor, and

clarifies the semantic connection holding between the information core of the utterance and the co-text on the left» (Pecorari, 2015a: 305).

In the case of simple PP, in Spanish the most usual prepositions prefacing AEs are *ante* 'faced with', *desde* 'from', *con* 'with', *para* 'for', *por* 'for', *tras* 'after'. I have conducted a corpus-based research on a small sample of journalistic texts (the subset of newspapers and journals published in Spain in the years 2000-2004 in the CREA corpus; 8,474,325 tokens) with the query "preposition + demonstrative". I have limited the query to the singular forms of demonstratives (except for the distal or third-order demonstrative which was discarded because of its low frequency in contemporary Spanish and scarce use in the construction of AEs): *este/a, ese/a, tal*.[21] Besides the five simple prepositions, the query has been extended to four complex PP: *a causa de* 'because of', *a consecuencia de* 'as a result of', *a partir de* 'as of', *a pesar de* 'despite of'.[22] Table 1 shows the quantitative results of the query. To give a more accurate information about the type of AEs occurring in this structure, the nominal cores are reported, but adjectival and prepositional modifiers have been eliminated.

Table 2. AEs in Frame Unit in Spanish journalistic texts in the CREA corpus (2000-2004)

PP	Demonstrative	Nouns in AE	Tokens	Types
A causa de	este / esta	siniestro	1	1
	ese / esa / tal	-	-	-
A consecuencia de	este / esta	accidente, victoria	2	2
	ese / esa / tal		-	
A partir de	este	momento[23] (12), punto (3), planteamiento (2), contexto, ejemplo, dato	20	6
	esta	idea (2), posición (2), muestra, premisa, realidad, observación, situación, definición, oposición, perspectiva, confluencia	13	11
	ese / esa / tal	momento (42), instante (2), punto (2), éxito, esquema, premisa	49	6
A pesar de	este	dato, inconveniente, equilibrio, volumen, acuerdo, fracaso, arrebato, dominio, conservadurismo, plante	10	10
	esta	situación (2), barrera, derrota, victoria, amenaza, subida, información, derivación, actividad, opinión	11	10
	ese / esa / tal	esperanza, cerrazón	2	2
Ante	este	panorama (8), dato (2), acontecimiento, hecho, hallazgo, maremágnum terminológico, regalo, vacío, riesgo, rumor, aumento, cúmulo de agresiones, conjunto de peticiones	21	13
	esta	situación (39), realidad (5), circunstancia (3), visión, disyuntiva, incapacidad, posibilidad, petición, oleada, suma, comparación, actitud, sospecha, medida, queja, agresión, tesitura, concepción, diversidad, coyuntura	65	20
	ese / esa	escenario, cifra, muestra, respuesta, situación	5	5
	tal	situación (6), desafío, circunstancia	8	3
Desde	este	punto de vista (16), planteamiento (4), momento (4), enfoque, aspecto, prisma	27	6
	esta	perspectiva (25), postura (2), óptica (2), visión, convicción, actitud, premisa, posición, autoexigencia	35	9
	ese	momento (13), punto de vista, punto, enfoque, criterio	17	5
	esa / tal	perspectiva (2), posición, óptica, formación, intención	6	5
Con	este	fin (12), objetivo (8), propósito (6), panorama (6), sistema (5), resultado (4), proyecto (4)	121	66[24]
	esta	medida (9), frase (8), perspectiva (6), iniciativa (5), decisión (5), información (4), victoria (4), idea (4), pregunta (4)	113	59
	ese	objetivo (8), punto de partida (2), superdesarrollo, diagnóstico, espíritu, planteamiento, informe, ritual, inicio, cometido, afán	19	12
	esa	financiación, receta, deuda, convicción, mezcla, premisa, formación, actividad, tesis, sensación, palabra, actitud, información, manta liada a la cabeza	14	14
	tal	motivo (2), medida, fin	4	3
Para	este	proceso (2), tipo de situaciones (2), proyecto, propósito, uso	7	5
	esta	cita (2), cuestión, actuación, iniciativa	5	4
	ese / esa / tal	fin (2), propósito, caso	4	3
Por	este	motivo (115), esquema, tipo de fórmula, hecho, acuerdo	119	5
	esta	razón (66), secuencia, situación, aportación, vía	70	5
	ese	motivo (13), camino	14	2
	esa	razón (8), regla de tres (2)	10	2
	tal	motivo (6), razón (2)	8	2
Tras	este	proceso (3), resultado (7), encuentro (5), acto (3)	65	44
	esta	operación (8), victoria (4), aventura (3), decisión (3)	79	58
	ese	éxito, impulso, empeoramiento, récord, golpe de efecto, parón, despliegue, enunciado, trámite, balance, fulgor	11	11
	esa	aclaración, entrega, medida, escena, declaración, argumentación, reunión, pausa, imagen, parábola	10	10
Total			**965**	**419**

As it can be observed, AEs in PP in Frame Units are a relatively frequent cohesive mechanism and offer a variety of nominal cores. The results point to a higher productivity in the case of simple PP and proximity demonstratives (Figure 1). On the other hand, the number of types outnumbering the 10 occurrences is reduced to 9 nouns, that constitute the 52.6% of the occurrences in the corpus.[25] In the case of *por este motivo* (14.4%) and *por esta razón* (8.3%) we may well speak of grammaticalization process at stake (Pecorari, 2014b, for similar observations regarding it, *per questo motivo* 'for this reason').
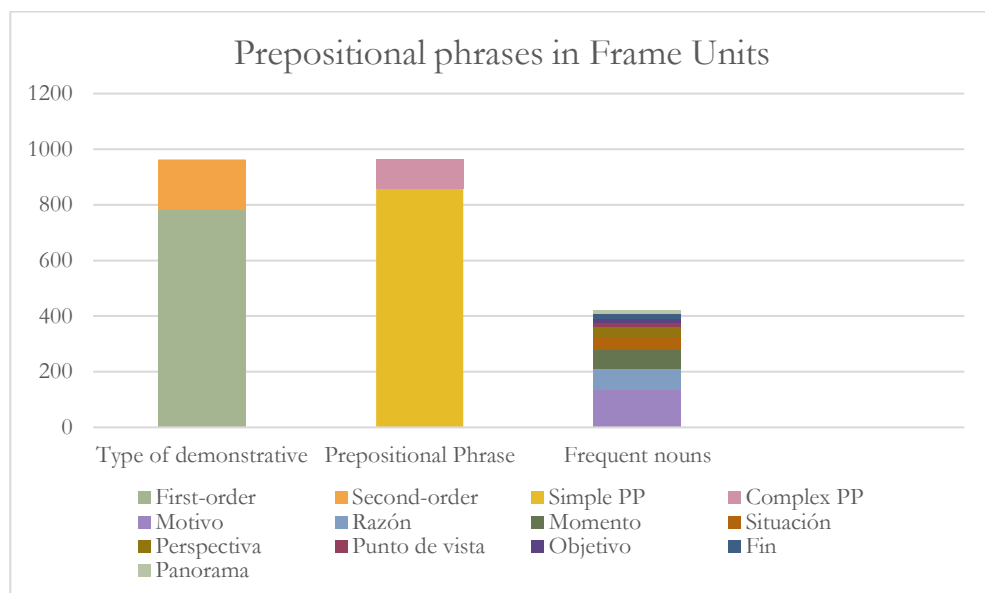
Figure 1. Types of demonstratives, prepositional phrases and frequent nouns in the CREA subcorpus of Spanish journalistic texts (2000-2004)

AEs inside a PP in the Frame Unit have mainly two functions at the level of discourse topicality: a) the establishment of a background in relation to which the information conveyed by the new Nucleus is relevant (as we have seen in 1); b) the closure of a previous topic to give way to a new discourse topic. This is the case in (10) where the AE 'this setback' closes a discourse topic (the rejection of application for human trial of a new medicine) and the text continues talking about the medicine's owners moving to Belgium and other countries.

(10) Eso es lo que intentó la familia. Incluso recurrió a una compañía farmacéutica nacional, Laboratorios Rovi, para que le prestara asesoramiento técnico a la hora de elaborar la documentación para tramitar los permisos pertinentes. Pero la solicitud de ensayo en humanos fue rechazada por la Dirección General de Farmacia y Productos Sanitarios del Ministerio de Sanidad (entonces aún no existía la Agencia Española del Medicamento).

El comité evaluador adujo los siguientes motivos: "dudas acerca de la base científica que soporta la eficacia terapéutica del producto", que el mecanismo de acción era desconocido y que antes de estudiarlo en voluntarios sanos debían despejarse incógnitas sobre su caracterización, cuantificando los componentes e identificando las proteínas activas.

// /Tras **este varapalo**,/[Frame] Rovi retiró su apoyo y los Chacón optaron por seguir investigando en el extranjero.// Y acudieron a países como Bélgica, cuya legislación en ensayos clínicos es más laxa y no requiere aprobación ministerial, a Alemania y a la república ex soviética de Georgia.

'That is what the family tried to do. They even turned to a national pharmaceutical company, Laboratorios Rovi, for technical advice in preparing the documentation to process the relevant permits. But the application for human trial was rejected by the General Directorate of Pharmacy and Health Products of the Ministry of Health (the Spanish Medicines Agency did not yet exist).

The evaluation committee gave the following reasons: "doubts about the scientific basis that supports the therapeutic efficacy of the product", that the mechanism of action was unknown and that before studying it in healthy volunteers, questions about its characterisation had to be cleared up, quantifying the components and identifying the active proteins.

After this setback, Rovi withdrew its support and the Chacóns opted to continue their research abroad and turned to countries such as Belgium, whose legislation on clinical trials is more laxer and does not require ministerial approval, Germany and the former Soviet republic of Georgia.'

Two final observations: from the analysis of this small corpus and although the types of AEs are not the focus of this research (Borreguero Zuloaga, 2018), I may say that AEs in Frame Units are mainly non-evaluative, they are rarely formed on the basis of axyonyms, nouns conveying an evaluation of the encapsulated information, nor are they accompanied by axiological adjectives. This is only a first impression and will require further research in the future, but it is consistent with González and Izquierdo Alegría (2020)'s claim that evaluative AEs are found in thematic (here Topic) and especially rhematic (here Comment) positions in the main proposition (here the Nucleus).

On the other hand, AEs in Frame Units appear frequently at the beginning of Paragraph and this is strictly linked to the management of DTs, as we will see in section 3, because the Frame Unit helps to delimit different DTs or to signal the continuity of the same DT.

2.3. Anaphoric encapsulators in the Appendix Unit

This is by long the least common position for AEs, maybe due to its low degree of cognitive salience. In fact, information in the Appendix is by definition a non-focalised information, a content that is deemed to be marginal, secondary, superfluous.

(11) España prepara un plan integral contra la trata de seres humanos destinados a la explotación sexual. // /La embajadora especial para Derechos Humanos,/Nucleus- /Silvia Escobar,/Appendix dio a conocer las directrices ayer en Viena en el marco de la Organización para la Seguridad y la Cooperación en Europa (OSCE),/-Nucleus /impulsora de **esta lucha** que requiere tanto de medidas nacionales como de estrecha colaboración transfronteriza/Appendix (El País, September 12th, 2007; from Llamas Saíz, 2010b: 110)
'Spain is preparing a comprehensive plan against trafficking in human beings for sexual exploitation. Special Ambassador for Human Rights Silvia Escobar unveiled the guidelines yesterday in Vienna within the framework of the Organisation for Security and Cooperation in Europe (OSCE), the driving force behind this fight, which requires both national measures and close cross-border cooperation.'

This position in an Information Unit subordinated to the Nucleus does not allow for a straightforward interpretation of AE (without underestimating the lexico-semantic difficulties to link 'a comprehensive plan' with 'this fight'). In fact, in (11) it is not easy to identify the antecedent. *La lucha* (the fight against trafficking in human beings for sexual exploitation) is a metaphorical NP to encompass Spain's project to minimize this type of traffic but also other international projects in the same direction.

Interpretation conditions are eased when the AE contains a nominalization (*la condena* 'the sentence'), as in (12). Nonetheless, here again the AE does not express a DT because the text goes on talking about internal problems in the Constitutional Court and not about the initial sentence.

(12) Fragoso fue condenado a una multa de 1260 euros, sustituible por una pena privativa de libertad en caso de impago.
// /El sindicalista recurrió a la Audiencia de A Coruña,/Nucleus1 /que confirmó **la condena**,/Appendix y /posteriormente acudió en solicitud de amparo al Constitucional,/Nucleus2 entonces con mayoría conservadora, que se lo denegó.// Hubo, sin embargo, cinco votos particulares de los magistrados… (J. M. Brunet, "Estrasburgo condena a España por vulnera la libertad de expresión de un sindicalista", *El País*, June 10th, 2023).
'Fragoso was sentenced to a fine of 1260 euros, replaceable by a custodial sentence in case of non-payment.

The trade unionist appealed to the Court of A Coruña, which upheld the sentence, and subsequently applied to the Constitutional Court, then with a conservative majority, which refused to grant him amparo.'

Therefore, AEs in the Appendix Unit are rarely referred to in the following text and have very low probabilities of constituting discourse topics.

3. Discourse topics and anaphoric encapsulators

The concept of discourse topic (DT) is an elusive one. Speakers have a clear intuition about texts dealing with different topics and the role these topics play in the internal organization of texts in paragraphs, sections, and chapters. However, among linguists there is no consensus about what defines a DT, rather there is a fuzzy conviction that any formal attempt to define DT is doomed to failure (Brown & Yule, 1983,1993). Leaving aside approaches in the field of formal semantics such as van Kuppevelt (1995) and Asher (2004), which do not consider real texts, the most successful definition of DT was the one proposed by van Dijk (1977) as a complex proposition logically entailed by the set of propositions expressed by the sequence of utterances. Nevertheless, this definition must face different problems, one of which is the length and the complexity of such proposition in the case of long texts and another the fact that it does not take into account the implicit meaning activated during text interpretation but only formal semantic representations of propositions.

I will thus adopt here the definition by Brown and Yule (1983, 1993): a DT emerges at the intersection of the activated knowledge through linguistic elements and the shared knowledge at a certain point in discourse. The DT has a dynamic nature and may evolve through the text, but it allows to determine what is relevant (in the sense of Grice's maxim of relation or relevance) and what is not for that specific text imposing an aboutness relationship to every content that is to be perceived as belonging to that DT. It may be verbalised by a NP, a proposition or a sequence of propositions, but what is crucial is that text coherence is grounded in topic continuity. Besides, to the extent that a text may be constituted by more than one DT, it is assumed that DTs can be hierarchically ordered. Downing, Neff, Carretero, Martínez-Caro, Pérez de Ayala, Marín and Simón (1998: 268) distinguish between global topics and local topics. The former are «sequentially organized and represent a wide concern which may subsume the conceptual content of a wide stretch of discourse», while the latter «are hierarchically structured under the 'umbrella' of the [discourse]-topic which unifies them».

A large number of studies has been devoted to the analysis of formal markers delimiting DTs, whether they are discourse markers (Fraser, 2009; Downing et al., 1998; Charolles, 2020), adverbials (Charolles et al., 2005; Ho-Dac & Péry-Woodley, 2009) or orthographic and typographic marks (paragraph segmentation, Martínez-Caro, 2014). Other authors have been more concerned with the internal ordering of

DTs, especially in the field of pyscholinguistics (Kintsch & van Dijk, 1978).[26] Here I will limit my observations to the role played by AEs in the management of discourse topicality.

Bánréti (1981) was the first to point out that anaphoric encapsulators make the topic of the preceding text explicit. According to him, text organisation can be conceived as a successive organization of contexts by which: «the *preceding text* becomes an *element* of the *context* as the statements that have *already* been made become parts of the speaker's inventory of actual knowledge and beliefs» (Bánréti, 1981: 43) and «by the means of these anaphoric relations the speaker himself communicates what he regards as the topic of his text» (Bánréti, 1981: 48). He defines DT as «a probable consequence of (some segment of) the text and of the context» (Bánréti, 1981: 49), i.e., the DT cannot be established on the basis of either of them alone; the DT should follow from the whole sequence of propositions in the preceding text, and the context reconstructed by the interlocutor on the background assumptions of the speaker must be part of that set of (implicit) propositions.

In Spanish linguistics the relation between AEs and DTs has been mentioned in some studies but has not received sufficient attention. González and Izquierdo Alegría (2020: 783) claim that in newspaper editorials «el referente del encapsulador suele coincidir o estar íntimamente vinculado con el tópico general del discurso»[27], but they don't explore further this connection. Ribera (2016), in his study on Catalan AE in parliamentary discourse, considers that one clear indicator that a referential expression is a DT is the number of anaphorical elements that refer to it. According to this author, AE occupy an intermediate area between sentence topic and DT, but not further clarification on this subject is offered.

In this section, I claim that the possibility for an AE to be part of the text semantic organization at the level of discourse topicality is linked mostly to its position in the Utterance. AEs fulfilling the function of Topics in a Nucleus Unit are more easily raised to the status of DTs. From a cognitive perspective, it has been said that an entity mentioned by a nominal expression in a prominent syntactic argument position (here a prominent position in information structure) is in focus (Zulaica-Hernández, 2009) and focal emplacements are optimal locations for lexical material verbalising a DT.

Space limitations make not possible to reproduce whole texts where this phenomenon is evident, but I will reproduce some extracts from a couple of texts where conversion of Topics into DTs can be observed. In (13) the AE *el apagón* 'the blackout' encapsulates the first Utterance of the text and becomes the Topic of the Nucleus of the next Utterance. This is a cognitively salient position, as we have seen in section 2.1, that eases the process of becoming part of the Frame Unit of the successive Paragraph in order to constitute the background information with respect

to which the new information (minimum services, the company's reactions, consequences, etc.) are introduced. The text is closed with a new occurrence of the AE in the Topic position of the Nucleus. It is interesting to remark that all the occurrences are in the first Utterance and the first position of the respective Paragraphs. Further evidence about this AE constituting a discourse topic is the presence of the same noun in the title.

(13) [1st Paragraph] <u>Toda la isla de La Gomera (unos 21.800 habitantes) se quedó sin suministro eléctrico a las tres de la madrugada de ayer</u>. // **/[El apagón]**_{Topic} se debió a un incendio que se produjo en la parte de las instalaciones de la central térmica de El Palmar […]/^{Nucleus} //_{U2}

[2nd Paragraph] // /Durante **el apagón**,/^{Frame} /los servicios esenciales de la isla funcionaron con normalidad […]/^{Nucleus} //_U

[8th and last Paragraph] // /[**El apagón**]_{Topic} también ha tenido consecuencias en el suministro de agua […]/^{Nucleus} // (EFE, "La Gomera sufre un apagón general tras un incendio en su central térmica", *El País*, July, 31st, 2023)

[1st Paragraph] The entire island of La Gomera (some 21,800 inhabitants) was without electricity supply at three o'clock yesterday morning. The blackout was caused by a fire that broke out in the El Palmar part of the thermal power station facilities [...].

[2nd Paragraph] During the blackout, the island's essential services functioned normally [...]

[8th Paragraph] The blackout also had an impact on the water supply [...]

In the second example, the role of the AE in the organization of discourse topics is not so obvious because the different coreferential NPs are subject to *variatio*. Once again, due to space limitations, I am constrained to present just some excerpts from the text.

(14) [1st Paragraph] <u>Caixabank ganó entre enero y junio 2137 millones de euros, un 35,8% más que en el mismo periodo de 2022</u> […]. <u>La entidad catalana consigue mejorar de forma notable sus resultados</u> […]

[3rd Paragraph] // /[**El avance** en el semestre]_{Topic} ha sido notable incluso tras el pago del impuesto temporal a la banca/^{Nucleus} // […] <u>las ganancias se habrían disparado en el primer semestre casi un 60%.</u>

[4th Paragraph] // /[**La mejora**]_{Topic} se explica, principalmente, por el impulso del negocio […]/^{Nucleus} //

[5th Paragraph] // /[**El acelerón**]_{Topic} se explica por el impacto del nuevo precio del dinero […]/^{Nucleus} // // /[**Este impulso]**_{Topic} está cerca de llegar a su término,/^{Nucleus} al menos con la velocidad de los últimos trimestres […]// (H. Gutiérrez, "Caixabank dispara sus ganancias hasta junio por la mejora de sus márgenes", *El País,* July 29th, 2023)

'[1st Paragraph] Caixabank earned between January and June 2,137 million euros, 35.8% more than in the same period of 2022 [...]. The Catalan entity has managed to improve its results notably [...]

[3rd Paragraph] The progress in the semester has been remarkable even after the payment of the temporary tax to the banks [...] the profits would have shot up in the first half of the year by almost 60%.

[4th Paragraph] The improvement is mainly explained by the momentum of the business [...].

[5th Paragraph] The acceleration is explained by the impact of the new price of money [...] This momentum is nearing its end, at least at the speed of the last few quarters [...].

The first AE encapsulating the content of Paragraph 1, *el avance* 'the progress' in the 3rd Paragraph, is taken up at the beginning of the 4th Paragraph by *la mejora* 'the improvement' and then at the beginning of the 5th Paragraph with the synonym *el acelerón* 'the acceleration' and in the middle of that long Paragraph with *el impulso* 'the momentum'. The presence of coreferential AEs in the 3rd, 4th, and 5th Paragraph that refer back to the information in the 1st Paragraph shows how an AE in Topic position has been promoted to discourse topic and establishes a semantic isotopy running through half of the text (which actually has eight Paragraphs).

Regarding the Frame Unit, it may have scope over just one Utterance and in this case the AE serves as a linking device with the previous Utterance but is not discursively developed in the text. In (15), the fact that Argentina must pay 3 billion dollars of its debt next week is encapsulated with the—certainly ironic—AE *el bache* 'the gap', but this is a side topic in the text, not further developed. In fact, the following Paragraph returns to the main topic, opening the first Paragraph (the new agreement between Argentina and the IMF).

(15) El Gobierno argentino celebró que el nuevo acuerdo quita la discusión sobre la deuda con el FMI de la campaña electoral. Los tiempos, sin embargo, no están con <u>Argentina que tiene vencimientos por 3.000 millones de dólares la semana próxima</u>. Para cubrir **el bache**, la Casa Rosada ha pedido créditos puente a otros organismos internacionales, como el CAF-Banco de Desarrollo de América Latina.
La revisión del acuerdo ha tenido en vilo al país. (J. P. Criales, "El FMI y Argentina cierran otro acuerdo por 6800 millones", *El País*, July 29th, 2023)
'The Argentine government welcomed the fact that the new agreement takes the discussion about the IMF debt out of the election campaign. The times, however, are not with Argentina, which has maturities of 3 billion dollars next week. To cover the gap, the Casa Rosada has asked for bridge loans from other international organisations, such as the CAF-Development Bank of Latin America.

The revision of the agreement has kept the country on tenterhooks.'

But an AE in a Frame Unit may well convey the background information for a sequence of Utterances or for a whole Paragraph, signalling the discourse topic, as in (13) above. As a matter of fact, Charolles et al. (2005: 124) highlight that:

> «the readers apply a default principle whereby they tend to attach an incoming utterance to the ongoing frame, and as a consequence expect the writer to explicitly signal a change of frame.»

In Conte (1996, 1999: 111-112)'s words:

> «anaphoric encapsulation quite often occurs in the initial point of a paragraph [either as a Nucleus' Topic or as part of the Frame Unit] and thus functions as an organizing principle in discourse structure. As a starting point of a new paragraph, anaphoric encapsulation is the shortest imaginable summary of the preceding discourse portion. »

Conversely, the rare AEs in the Appendix Unit are never promoted to the level of discourse topics.

What is also remarkable from a textual point of view is that the discourse referent, playing the role of DT, is not introduced immediately by a referring expression, but through a fully-fledged utterance, later hypostatized by an AE. There is a division of labour between strategies used for the introduction of DT and strategies used for its establishment as a discourse referent.[28]

CONCLUSIONS

Assessing the interplay between the presence of AEs and the information structure of text will also help us to have a better understanding of the difficulty to process them (García et al., 2005). There are several factors that contribute to this difficulty: the abstract character of the nominal core, the informational density, the absence of verbal arguments (or their reduction to noun complementizers), their polyphonic nature (with the voice of the author of the text often concealing the voice of the participants in the reported events), and the presuppositional nature of their semantic content.

Distance and syntactic complexity are two factors that directly influence the salience of the antecedent, but also the contextual and encyclopaedic knowledge that the reader is assumed to retrieve during the interpretation process.[29] To avoid mistakes in assessing the accessibility of the anaphorical expression (Figueras, 2002), i.e., in linking an anaphor to its antecedent, AEs are preferred over textual (pronominal) anaphors.

According to previous studies, the ability of the reader to link the AE to the previous text is based on different factors. Here is a short list of some of the most fundamental ones:

a) the syntactic and semantic complexity of the antecedent (González, 2008; López Samaniego, 2013); this has a direct consequence on the reader's capability of identifying and delimiting the antecedent due to the fact that, as cognitive approaches have claimed, the antecedent is not only a set of lexical expressions but also the reconstruction of a state of affairs based on the reader's context and encyclopaedia knowledge, a knowledge that is activated by a discourse segment in the co-text preceding the AE (Loureda et al., this volume);[30]

b) the type of encapsulator (López Samaniego, 2015; Borreguero Zuloaga, 2018) and thus the complexity of the inferential process necessary to retrieve the antecedent: metaphorical encapsulators are harder to process than 'neutral' ones, because they presuppose both an encyclopaedic knowledge shared with the writer (Conte, 1998) that must be activated and the capacity to recognize the conceptual metaphors underlying their use (Llamas Saíz, 2010b; Pecorari, 2021); the global meaning is thus the result of an inferential process (D'Addio, 1988), ensuing from the contextual enrichment of a sequence of lexical elements; conversely, nominalizations are easier to process because they have a common lexeme with the main verb of the predication in the antecedent; however, a semantically weak form (general nouns), not providing any significant information, hardly contributes to the identification of the referent (Lala, 2010b; López Samaniego, 2010), and this explains that it could be harder to process than NPs (Loureda et al., this volume);[31]

c) the type of anaphoric element (definite article, demonstrative): the instructions about the accessibility of the antecedent vary according to the type of anaphoric element pre-facing the AE (Ariel, 1998; Figueras, 2002);[32] most authors claim that demonstratives are the prototypical pre-modifier in AEs and almost compulsory in the case of metaphorical ones (Conte 1996, 1999), as we have seen in section 2.1.; in other words, «the less cognitively accessible the referent, the greater the tendency to use the demonstrative pronoun» (Dam, 2014);

d) the distance between the anaphoric encapsulator and the encapsulated text, because referential distance contributes to promoting a discourse entity to a certain cognitive status (Cowles & Garnham, 2005; Zulaica-Hernández, 2009); the distance should be measured horizontally (the number of clauses between the AE and the antecedent) and vertically (the number of syntactic levels, i.e., embedded clauses, between them); this is strictly related to the number of

entities introduced between the antecedent and the AE that could compete as antecedents of this last NP.

To these factors I will add the utterance position of the AEs and the possibility to become a discourse topic. Salient positions such as Topic in a Nucleus Unit or Frame Unit at the beginning of a Paragraph ease the reader's task of retrieving the antecedent of the AE, while other less prominent positions like Comment in a Nucleus Unit or Appendix may present more difficulties, a claim that awaits empirical confirmation similar to the one provided by Loureda et al. (this volume) regarding the processing of different types of encapsulation mechanisms.

On the other hand, not every AE fulfils textual functions to the same extent (Pecorari, 2015c). In a similar way to Ariel's accessibility scale, we can propose a scale regarding the possibility of an AE to become the expression of a discourse topic according to its Utterance position, from the most prominent to the less prominent one: Topic in Nucleus > part of the Topic in Nucleus > PP in Frame > Comment in Nucleus > Appendix.[33]

My claim is that AEs play a fundamental role not only in the information structure at the Utterance level but also at the higher level of discourse structure and that these two levels are intimately linked. The informative contribution of the AE is more variegated than acting as a theme (Nucleus Topic) or as part of the rheme (Nucleus Comment). It may set a framework for successive utterances or close a previous DT (Frame Unit) but also recall a background information to characterize a referent with its scope limited to part of the Utterance (Appendix Unit). The articulation of the Utterance in different information levels explains why placing an AE in one or the other determines its contribution to the text global coherence.

REFERENCES

Álvarez de Mon y Rego, I. (2001). Encapsulation and Prospection in Written Scientific English. *Estudios Ingleses de la Universidad Complutense*, 9, 81-101.

Ariel, M. (1988). Referring and Accessibility. *Journal of Linguistics,* 24, 65-87.

Asher, N. (2004). Discourse Topics. *Theoretical Linguistics*, 30, 163-201.

Azpiazu, S. (2004). *Las estrategias de nominalización. Estudio contrastivo del estudio nominal.* Frankfurt: Peter Lang (Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation, 15).

Bánréti, Z. (1981). The Topic of Texts and the Interpretation of Texts. In J. S. Petöfi (Ed.), *Text vs. Sentence Continued* (pp. 43-57). Hamburg: Buske (Papiere zur Textlinguistik, 29).

Bertucci, R. A. (2006). Anáforas encapsuladoras: uma análise em textos de opinião. *Revista Letras*, 70, 207-221.

Blanche-Benveniste, C. (1990). *Le français parlé, études grammaticales*. Paris : Éditions du CNRS.

Borreguero Zuloaga, M. (2006). Naturaleza y función de los encapsuladores en los textos informativamente densos (la noticia periodística). *Cuadernos de Filología Italiana*, 13, 73-95.

Borreguero Zuloaga, M. (2014). Left Periphery in Discourse. Frame Unit and Discourse Markers. In A. Dufter & Á. Octavio de Toledo (Eds.), *Left Peripheries in Spanish. Diachronic, Variationist and Comparative Perspectives* (pp. 45-382). Amsterdam: Joh Benjamins.

Borreguero Zuloaga, M. (2018). Los encapsuladores anafóricos: Una propuesta de clasificación. *Caplletra,* 64, 179-203.

Borreguero Zuloaga, M. (2021). Las unidades del discurso y el texto escrito. In Ó. Loureda & A. Schrott (Eds.), *Manual de Lingüística del Hablar* (pp. 373-391). Berlin: De Gruyter.

Borreguero Zuloaga, M. & Octavio de Toledo, Á. S. (2007). Presencia y función de los encapsuladores en las crónicas periodísticas del s. XVIII. *Philologia Hispalensis,* 21, 119-153.

Borreguero Zuloaga, M., Atayan, V. & Grosse, S. (Eds.) (2018). *Models of Discourse Units in Romance Languages. Revue Romane,* 53, 1.

Brown, G. & Yule. G. ([1983] 1993). *Discourse Analysis.* Cambridge: Cambridge University Press. [*Análisis del discurso*, trad. esp. de S. Iglesias, Madrid: Visor, 1993].

Casado Velarde, M. (2008). Algunas estrategias discursivas en el lenguaje periodístico de hoy. *Boletín Hispánico Helvético,* 12, 71-97.

Charolles, M. (2020). Discourse Topics and Digressive Markers. *Journal of Pragmatics* 161, 57-77.

Charolles, M., Le Draoulec, A., Pery-Woodley, M.-P. & Sarda, L. (2005). Temporal and Spatial Dimensions of Discourse Organisation. *French Language Studies,* 15, 115-130.

Conte, M. E. ([1988]1999). *Condizioni di coerenza. Ricerche di linguistica testuale.* Firenze: La Nuova Italia. [Nueva edizione con l'aggiunta di due saggi a cura di B. Mortara Garavelli. Alessandria: Edizioni dell'Orso, 1999].

Conte, M. E. ([1996]1999). Anaphoric Encapsulation. In W. De Mulder & L. Tasmowski (Eds.), *Coherence and Anaphora. Belgian Journal of Linguistics* (pp. 1-10). [*Condizioni di coerenza* (pp. 107-114). Alessandria: Edizioni dell'Orso, 1999].

Conte, M. E. (1998). Il ruolo dei termini astratti nei testi. In G. Bernini, P. Cuzzolin & P. Molinelli (Eds.), *Ars linguistica. Studi oferti da colleghi ed allievi a Paolo Ramat in occasione del suo 60° compleanno* (pp. 151-160). Roma: Bulzoni.

Cowles, H. W. & Garnham, A. (2005). Antecedent Focus and Conceptual Distance Effects in Category Noun-Phrase Anaphora. *Language and Cognitive Processes, 20*(6), 725-750.

Cresti, E. (2000). *Corpus di italiano parlato*. Firenze: Accademia della Crusca.

D'Addio, W. (1988). Nominali anaforici incapsulatori: un aspetto della coesione lessicale. In T. De Mauro, S. Gensini & M. E. Piemontese (Eds.), *Dalla parte del ricevente: percezione, comprensione, interpretazione. Atti del XXVI Convegno Internazionale di Studi della Società Linguistica Italiana* (pp. 143-151). Roma: Bulzoni.

Dam, L. (2014). The Interpretation of Encapsulating Anaphors in Spanish and their Functions. *Folia Linguistica, 48*(1), 37-60.

Daneš, F. (1974). Functional Sentence Perspective and the Organization of the Text. In F. Daneš (Ed.), *Papers on Functional Sentence Perspective* (pp. 106-128). Prague/The Hague: Publishing House of the Czechoslovak Academy of Sciences/Mouton.

Downing, A., Neff, J., Carretero, M., Martínez-Caro, E., Pérez de Ayala, S., Marín, J. & Simón, J. (1998). Structuring and Signalling Topic Management. *Lacus Forum*, 29, 267-279.

Ferrari, A. (2002). Aspetti semantici e informativi della nominalizzazione sintagmatica. In C. Marello & G. L. Beccaria (Eds.), *La parola e il testo. In onore di Bice Mortara Garavelli* (pp. 179-204). Alessandria: Edizioni dell'Orso.

Ferrari, A. (2014). The Basel Model for Paragraph Segmentation. The Construction Units, their Relationships and Linguistic Indication. In S. Pons (Ed.), *Models of Discourse Segmentation. Explorations across Romance languages* (pp. 23-53). Amsterdam/Philadelphia: John Benjamins.

Ferrari, A., Cignetti, L., De Cesare, A.-M., Lala, L., Mandelli, M., Ricci, C. & Roggia, C. E. (2008). *L'interfaccia lingua-testo*. Alessandria: Edizioni dell'Orso.

Ferrari, A. & Borreguero Zuloaga, M. (2015). *La interfaz lengua-texto. Un modelo de estructura informativa.* Madrid: Biblioteca Nueva.

Figueras, C. (2002). La jerarquía de la accesibilidad de las expresiones referenciales en español. *Revista Española de Lingüística, 32*(1), 53-96.

Flowerdew, J. (2015). Revisiting Metadiscourse: Conceptual and Methodological Issues Concerning Signalling Nouns. *Iberica,* 15, 15-34.

Francis, G. (1994). Labelling Discourse: An Aspect for Nominal-Group Lexical Cohesion. In M. Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 83-101). London: Routledge.

Fraser, B. (2009). Topic Orientation Markers. *Journal of Pragmatics,* 41, 892-898.

García Negroni, M. M., Hall, B. & Martín, M. (2005). Ambigüedad, abstracción y polifonía del discurso académico: interpretación de las nominalizaciones. *Revista Signos. Estudios de Lingüística, 38*(57), 49-60.

González, R. (2008). Las nominalizaciones como estrategia de manipulación informativa en la noticia periodística: El caso de la anáfora conceptual. In I. Olza Moreno, M. Casado Velarde & R. González (Eds.), *Actas del XXXVII Simposio Internacional de la Sociedad Española de Lingüística* (pp. 247-259). Pamplona: Universidad de Navarra.

González, R. (2010). Gramática y discurso: Nominalización y construcción discursiva en las noticias periodísticas. In C. Martínez Pasamar (Ed.), *Estrategias Argumentativas en el Discurso Periodístico* (pp. 119-146). Frankfurt: Peter Lang.

González, R. & Izquierdo Alegría, D. (2020). Encapsulación y estructura informativa: Análisis cuantitativo y cualitativo de las variables influyentes en la selección de patrones temático-remáticos en el editorial periodístico. *Zeitschrift für romanische Philologie*, 136, 749-788.

Halliday, M.A.K. & Hasan, R. (1976). *Cohesion in English.* London: Longman.

Ho-Dac, L. M. & Péry-Woodley, M.-P. (2009). A Data-Driven Study of Temporal Adverbials as Discourse Segmentation Markers. *Discours* 4. *Linearizaiont and Segmentation in Discourse* (special issue) [online]

Izquierdo Alegría, D. & González, R. (2013a). Encapsulación y etiquetas discursivas en el discurso parlamentario: Función argumentativa a partir de un corpus paralelo. *Oralia*, 16, 186-221.

Izquierdo Alegría, D. & González Ruiz, R. (2013b). La evaluación a través de un mecanismo cohesivo: Las etiquetas discursivas en un corpus de editoriales. In C. Llamas Saiz, C. Martínez Pasamar & M. Casado Velarde (Eds.), *Léxico y Argumentación en el Discurso Público Actual* (pp. 149-186). Franfkurt: Peter Lang.

Kintsch, W. & van Dijk, T. A. (1978). Towards a Model of Text Comprehension and Production. *Psychological Revue, 85*(5), 363-394.

Korzen, I. (2016). Come riassumere un messaggio politico? Strategie e strutture di incapsulazione anaforica nei discorsi del parlamento europeo. In G. Ruffino & M. Castiglione (Eds.), *La lingua variabile nei testi. Atti del XV Congresso della Società Internazionale di Linguistica e Filologia Italiana* (pp. 443-459). Firenze: Franco Cesati.

Lala, L. (2010a). Incapsulatori. In R. Simone (Ed.), *Enciclopedia dell'Italiano.* Roma: Treccani [online]. Retrieved from: <www.treccani.it>.

Lala, L. (2010b). L'incapsulatore anaforico *cosa* nell'orale e nello scritto. In A. Ferrari & A.-M. De Cesare (Eds.), *Il parlato nella scrittura italiana odierna: riflessioni in prospettiva testuale* (pp. 57-78). Bern: Peter Lang.

Lambrecht, K. (1994). *Information Structure and Sentence form. Topic, Focus and the Mental Representations of Discourse Referents.* Cambridge: Cambridge University Press.

Llamas Saíz, C. (2010a). Argumentación en la noticia periodística: El caso de la anáfora conceptual metafórica. En C. Martínez Pasamar (Ed.), *Estrategias Argumentativas en el Discurso Periodístico* (pp. 147-170). Frankfurt: Peter Lang.

Llamas Saíz, C. (2010b). Interpretación del discurso ajeno: La anáfora conceptual metafórica en la noticia periodística. *Revista de Investigación Lingüística,* 13, 107-126.

López Samaniego, A. (2010). Documentos profesionales con destinatarios no expertos. El empleo de los mecanismos referenciales en la sentencia del 11M. *Revista Signos. Estudios de Lingüística, 43*(72), 99-123.

López Samaniego, A. (2013). Las etiquetas discursivas: Del mantenimiento a la construcción del referente. *ELUA,* 27, 167-197.

López Samaniego, A. (2014). *Las etiquetas discursivas: Cohesión anafórica y categorización de entidades del discurso.* Pamplona: Eunsa.

López Samaniego, A. (2015). Etiquetas discursivas, hiperónimos y encapsuladores: Una propuesta de clasificación de las relaciones de cohesión referencial. *RILCE, 31*(2), 435-462.

Lundquist, L. (2005). Noms, verbes et anaphores (in)fidèles. Pourquoi les Danois sont plus fidèles que les Français. *Langue Française*, 145, 73-91.

Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.

Magris, M. (2000). I nominali incapsulatori nel diritto. Contributo alla competenza testuale e lessicale del traduttore. In D. Veronesi (Ed.), *Linguistica giuridica italiana e tedesca / Rechtslinguistik des Deutschen und Italienischen* (pp. 459-468). Padua: Unipress.

Martínez-Caro, E. (2014). El párrafo como unidad discursiva: Consideraciones de forma y contenido relativas a su demarcación y estructuración. *Estudios de Lingüística del Español*, 35, 197-221.

Mattheoudakis, M. & Hatzitheodorou, A.-M. (2011). Discourse Organising Nouns in American and Greek University Students' Argumentative Essays: A Contrastive Study. In E. Kitis, N. Lavidas, Nina Topintzi & T. Tsangalidis (Eds.), *Selected papers from the 19th International Symposium on Theoretical and Applied Linguistics (ISTAL, 9th April 2009)* (pp. 329-333). Thessaloniki: Monochromia.

Méndez García de Paredes, E. (2003). Nominalización y tipo de texto. In J. L. Girón, S. Iglesias, F. J. Herrero Ruiz de Loizaga, A. Narbona & J. J. de Bustos Tovar (Eds.), Estudios ofrecidos al profesor José Jesús de Bustos Tovar (pp. 1015-1032). Madrid: Editorial Complutense.

Moirand, S. (1973). Le rôle anaphorique de la nominalisation dans la presse écrite. *Langue Française,* 28, 60-78.

Montolío, E. (2013). Construcciones conectivas que encapsulan. [A pesar de + SN] y la escritura experta. *Cuadernos Aispi,* 2, 115-132.

Moreno, A. I. (2004). Retrospective Labelling in Premise–Conclusion Metatext: An English–Spanish Contrastive Study of Research Articles on Business and Economics. *Journal of English for Academic Purposes*, *3*(4), 321-339.

Mortara Garavelli, B. (1971) Fra norma e invenzione: lo stile nominale. *Studi di grammatica italiana,* 1, 271-315.

Pecorari, F. (2014a). Ai confini dell'incapsulazione anaforica: strategie incapsulative non prototipiche. In E. Pîrvu (Ed.), *Discorso e cultura nella lingua e nella letteratura italiana. Atti del V Convegno internazionale di italianistica dell'Università di Craiova, 20-21 settembre 2013* (pp. 257-269). Firenze: Franco Cesati.

Pecorari, F. (2014b). L'incapsulazione zero: aspetti semantici, informativi e testuali. In F.R. Macaluso (Ed.), *La lingua variabile nei testi letterari, artistici e funzionali contemporanei. Analisi, interpretazione, traduzione. Testi presentati al XIII Congresso della SILFI (Società Internazionale di Linguistica e Filologia Italiana). Palermo, 22-24 settembre 2014*. Palermo: Centro de studi filologici e linguistici siciliani.

Pecorari, F. (2015a). Where Textual Cohesion Meets the Logical Dimension: Relational Anaphoric Encapsulation. *Estudos Linguísticos*, 10, 297-312.

Pecorari, F. (2015b). Incapsulatori anaforici e discorso riportato nell'italiano giornalistico: sfruttamenti coesivi della dimensione enunciativa. In J. Miecznikowski, M. Casoni, S. Christopher, A. Kamber, E. M. Pandolfi & A. Rocci (Eds.), *Norme linguistiche in contesto / Sprachnormen in Kontext / Normes langagières en contexte / Language Norms in Context. Actes du colloque VALS-ASLA 2014 (Lugano, 12-14 février 2014). Bulletin Suisse de Linguistique Appliquée*, numéro spécial, I, 227-244.

Pecorari, F. (2015c). La coesione testuale dei lanci di agenzia: uno studio delle anafore di ordine superiore. *Revue Romane*, *50*(2), 222-278.

Pecorari, F. (2016). L'incapsulazione anaforica nell'italiano contemporaneo: analisi di un corpus giornalistico. *Bulletin suisse de linguistique appliquée*, 104, 27-50.

Pecorari, F. (2021). Cohesive Figures: Metaphorical and Metonymic Anaphoric Encapsulation. In I. Rizzato, F. Strik Lievers & E. Zurro (Eds.), *Variations on Metaphor* (pp. 28-44). Newcastle upon Tyne: Cambridge Scholars.

Pelo, A. (1986). I nomi generali nella lingua dei giornali italiani. In K. Lichem, E. Maray & S. Knaller (Eds.), *Parallela 2. Aspetti della sintassi dell'italiano contemporaneo. Atti del 3° incontro italo-austriaco di linguisti a Graz, 28-31 maggio 1984* (pp. 205-214). Tübingen: Gunter Narr.

Peña, G. (2004). La anáfora lexical y el discurso de divulgación científica: Aproximación contrastiva. In R. López Carrillo & J. Suso López (Eds.), *Le français face aux défis actuels: histoire, langue et culture* (pp. 783-794). Granada: Universidad de Granada.

Peña, G. & Olivares, Mª A. (2009). La anáfora conceptual en el discurso de divulgación científica en francés. *RSEL, 39*(2), 177-194.

Pons, S. (Ed.) (2014). *Discourse Segmentation in Romance Languages*. Amsterdam: John Benjamins.

Pons, S. & Salameh, S. (2021). Las unidades del hablar: La oralidad. In Ó. Loureda & A. Schrott (Eds.), *Manual de lingüística del hablar* (pp. 393-417). Berlin: De Gruyter.

Prandi, M. (2006). *Le regole e le scelte. Introduzione alla grammatica italiana.* Torino: UTET.

Ribera, J. E. (2016). Encapsulació i estructura informativa en el debat parlamentari. Una anàlisi contrastiva (català – español – anglès). *Studia Romanica Posnaniensia*, *43*(2), 29-50.

Schmid, H.-J. (1998). Constant and Ephemeral Hypostatization: Thing, Problem and Other Shell Nouns. In B. Caron (Ed.), *Proceedings of the 16th International Congress of Linguists (Paris, July 20-25, 1997).* Amsterdam: Pergamon/Elsevier.

Sinclair, J. (1983). Plans of Discourse. *The Twofold Voice. Essays in Honour of Ramesh Mohan.* Salzburg: Universität Salzburg.

Swales, J. M. (2001). Metatalk in American Academic Talk: The Cases of Point and Thing. *Journal of English Linguistics*, *29*(1), 34-54.

Van Dijk, T. A. (1977). *Text and Context.* London: Longmann.

Van Kuppevelt, J. (1995). Main Structure and Side Structure in Discourse. *Linguistics,* 33, 809-833.

Vignuzzi, U. (1986). I nomi generali in italiano: alcuni problemi tra sintassi e semantica. In K. Lichem, E. Maray & S. Knaller (Eds.), *Parallela 2. Aspetti della sintassi dell'italiano contemporaneo. Atti del 3° incontro italo-austriaco di linguisti a Graz, 28-31 maggio 1984* (pp. 317-328). Tübingen: Gunter Narr.

Zampese, L. (2005). La struttura informativa degli articoli di cronaca: natura e funzioni dell'Unità di Quadro. In A. Ferrari (Ed.), *Rilievi. Le gerarchie semantico-pragmatiche di alcuni tipi di testo* (pp. 173-216). Firenze: Cesati.

Zulaica-Hernández, I. (2009). On the Cognitive Status of Antecedents in Spanish Discourse Anaphora. In P. Cantos Gómez & A. Sánchez Pérez (Eds.), *A survey of corpus-based research* (pp. 646-663). Murcia, Asociación Española de Lingüística de Corpus (AELINCO) [online]. Retrieved from: www.um.es/lacell/aelinco.

NOTES

[1] As it is frequent with linguistic phenomena, this cohesive mechanism is known by different terms, referring specifically to the nominal core ('shell nouns', 'carrier nouns') or to the whole NP ('discourse labels'), often specifying their anaphoric nature ('retrospective labels', 'conceptual anaphors', 'pragmatic anaphors'), including less felicitous ones such as 'grammatical metaphors' or 'anaphoric nouns'. I will use the term 'anaphoric encapsulators' (Conte, 1996, 1999) for its transparency in describing the discourse function of these NPs. The term seems to have been introduced by Sinclair (1983), although the first description of this

textual phenomenon goes back to at least Porzig's *Die Leistung der Abstrakta in der Sprache* (1930-31). However, it was Halliday and Hasan's (1976) description of general nouns as cohesive devices that started the series of countless studies about this type of anaphors at our disposal today. In the field of text linguistics some pioneer studies are Moirand (1973), Mortara Garavelli (1971), Bánréti (1981), Conte (1988, 1999), D'Addio (1988). However, the bulk of books and papers on this subject in different languages and considering different text types has been published from 1990's onwards. López Samaniego (2014) for a detailed account of the terminological question, and Conte (1998) for a brief historiographical notice.

2 I adopt a classical restricted definition of AE excluding pronouns that may also encapsulate a segment of text such as *eso* or *esto* 'that, this', which I prefer to call 'textual anaphors' (although 'metatextual anaphors' is probably a less ambiguous term), as well as elliptic subjects, which are possible in pro-drop languages and have been called zero incapsulators (e.g, *Un hombre ha asesinado a su mujer y a sus dos hijos y se ha suicidado. Ø Ha sucedido en Torrevieja.* 'A man killed his wife and two sons and committed suicide. It happened in Torrevieja', Pecorari, 2014b). For a larger conception of AE, Lala (2010); Pecorari (2014a, 2014b, 2015a, 2015c). However, I do not restrict the nominal core of the encapsulator to the so-called *discursive labels* (López Samaniego, 2014, 2015), i.e. AEs whose nominal core does not share the lexical root with any verbal predicate in the antecedent. Mine is a functional definition of AE, based on operations such as reification, hypostasis and contribution to information structure; the morphological nature of the nominal core is not, in my view, a central characteristic.

3 Some pre-modifiers such as the deverbal adjective *dicho* 'such' may also occupy this position. On the other hand, encapsulators may also function as cataphors, but this discourse function will not be taken into account in the following pages (López Samaniego, 2014).

4 Antecedent is understood here as any linguistic expression (with or without referring function) to which an anaphor is linked in order to acquire a full semantic interpretation (Pecorari, 2015c).

5 López Samaniego (2013) offers a detailed analysis of the syntactic structures that may function as antecedents for an AE. According to her research, the most common structure is subordinate clause with finite verbs.

6 Speech acts were introduced as fourth-order entities by Dik, but here they are considered as third-order entities (Pecorari, 2015c).

7 There are some exceptions to this claim: when the antecedent is a NP it has referential nature, but NPs referred back by an AE normally express a predication (e.g., *el rechazo del gobierno a reducir la jornada laboral* 'the government's refusal to reduce working hours'—AE: *esta 339ecision* 'this decision'), so their referential nature is more complex than that of NPs codifying first-order entities (e.g., *el libro de matemáticas*, 'the math book'). In these cases the AE does not fulfil a reifying of the predication, but only a recategorization by presenting the refusal as a decision. I then disagree with López Samaniego (2013: 190) who claims that «la reificación conceptual y la creación de nuevos referentes discursivos no se dan en todos los usos de las EEDD [etiquetas discursivas] ni (cabe esperar) en todos los procedimientos de encapsulación en general» ('conceptual reification and the creation of new discursive referents do not occur in

all uses of DLs [discursive labels] nor (expectedly) in all encapsulation procedures in general.'). Other authors prefer to speak of coreference between anaphors of higher order and not of encapsulation in these cases (Pecorari, 2015c). In my view, it is important to distinguish reification and recategorization: the latter comes always into play in the construction of AEs and gives rise to a new discourse referent (see fn. 9 below).

[8] As a consequence, the focus has been on information manipulation through anaphoric encapsulation in different types of texts and discourse genres: scientific (Álvarez de Mon & Rego, 2001; Peña, 2004; Peña & Olivares, 2009); journalistic (Bertucci, 2006; Borreguero Zuloaga, 2006; Borreguero Zuloaga & Octavio de Toledo, 2007; Casado Velarde, 2008; González, 2008, 2010; González & Izquierdo Alegría, 2020; Izquierdo Alegría & González, 2013b; Llamas Saíz, 2010a, 2010b; López Samaniego, 2014; Pecorari, 2015a, 2015b, 2016, 2017); political (Izquierdo & González, 2013a; Korzen, 2016; Ribera, 2016); academic (García et al., 2005; Mattheoudakis & Hatzitheodorou, 2011; Moreno, 2004; Swales, 2001); legal (Magris, 2000; López Samaniego, 2010).

[9] Some authors (López Samaniego, 2014; Gonzáles & Izquierdo Alegría, 2020) consider that a new categorization of the referent takes places only in the case of discursive labels (which excludes not only pronominal forms but also nominalizations and general nouns). Following Conte (1996), Montolío (2013) and Dam (2014), I consider that even nominalizations and general nouns fulfil the operations of reification, recategorization and hypostasis by which a fact or situation is treated as an entity (Dam, 2014). The difference between nominalizations and the so-called 'discursive labels' is that the latter may introduce the evaluative load not only in the modifiers but also in the nominal core guiding the interpretation of the reader with a certain ideological bias. But encapsulating a robbery with the AE *el robo* 'the robbery', *el asombroso robo* 'the amazing robbery', even when the verb *robar* 'to steal' appears in the antecedent, or *el atraco del siglo* 'the heist of the century' are three different ways of recategorizing the referent.

[10] 'Anaphoric encapsulations can thus appear at nodal points of the textual architecture and contribute significantly to the organisation and structuring of the text'.

[11] The use of capital letters indicates specific terms used in the Basel model.

[12] Functional words such as interjections can form a Communicative Unit, but this is a marginal case in written texts. They appear only in written dialogues that mimic face-to-face conversations.

[13] Capital letters are used to refer to sentence Topic, a referent that establishes a relation of aboutness with the rest of the Nucleus and occupies the first position in this Unit. No capital letters are used when 'topic' refers to discourse topic, the global themes that structure a whole discourse. As we will see below, an AE can function both as Topic and topic.

[14] The Basel model takes the notion of Topic from Lambrecht (1994: 131): "A referent is interpreted as the topic of a proposition if in a given situation the proposition is construed as being about this referent, i.e., as expressing information which is relevant to and which increases the addressee's knowledge of this referent".

[15] However, this does not seem to be the case in all discourse genres and writing styles. Pecorari (2016) shows that 62% of the AE found in a corpus of journalistic texts (news releases, informative and argumentative texts) did not occupy a Topic position.

[16] The selection of this NPs vehicles an axiological component to the extent that it expresses a negative evaluation on the part of the writer of the described behaviour, but this aspect would not be considered here as it is not central to our discussion.

[17] Although less frequent it is possible to find examples of linear progression, where the AE in the Topic only encompasses, totally or partially, the information in the Comment of the previous Utterance: (a) Al más puro estilo del expresidente del BCE, Arabia Saudi —el mayor exportador de crudo del mundo— se comprometió el domingo a "hacer todo lo que sea necesario para llevar la estabilidad de nuevo al mercado". // /[**Las palabras**]$_{Topic}$ cristalizaban en hechos:/$_{Nucleus}$ // (I. Ariza, "Arabia Saudí cierra el grifo para sostener el precio", El País, June 6th, 2023) 'In the style of the former ECB president, Saudi Arabia—the world's largest crude oil exporter—pledged on Sunday to "do whatever it takes to bring stability back to the market". [The words]^{Topic} crystallised into action'.

[18] The text has not been modified. Marks of suppressed text (…) were already in the original.

[19] These authors have shown that AEs with evaluative elements are preferably part of the Comment (rhematic position in the terminology of these authors) where the evaluation takes place in a more explicit way. But this can be determined by the type of texts (newspaper editorial) on which the research was conducted.

[20] In a different theoretical framework López Samaniego (2014) signals the presence of AE in paragraph initial positions which are not part of the predicative relations. However, her focus is on the different functions at the utterance information level fulfilled by AES according to their different patterns and positions. These functions are restricted to topic and comment functions (López Samaniego, 2014).

[21] In her definition of AE, Conte (1996, 1999) claims that demonstrative is the typical determiner in this type of NPs. This belief has led some authors to limit their queries to the AEs preceded by demonstratives (Izquierdo Alegría & González, 2013; González & Izquierdo Alegría, 2020). However, corpus-based studies, such as Pecorari (2015c, 2016), have demonstrated that at least in some discourse genres the definite article is much more frequent than the demonstrative. My decision is motivated by the higher probability of finding AEs if they are introduced by determiners indicating proximity to the antecedent, but I would not dare to say that this type of AE is more frequent than the one preceded by a definite article. For an in-depth view of the intricate discussion on the different anaphoric functions of definite articles and demonstratives, Lundquist (2005).

[22] The first letter of the PP was written in caps to limit the query to Frame Units.

[23] *Momento* 'moment' is considered part of an AE when it refers not so much to a precise temporal reference as to an act, situation or process described in the antecedent. This vague reference is blocked when it is accompanied by modifiers such as *mismo* or *preciso* meaning 'that very moment'.

²⁴ When number of types exceeds 20, only those with more than 3 occurrences are reported.

²⁵ Many of these nouns belong to the class of the so-called general nouns (GN; Halliday & Hasan, 1976; Swales, 2001; Flowerdew, 2015), which constitute a semantic class of nouns characterized by the vagueness of their reference. Although Halliday and Hasan (1976) only consider GN such as 'thing', 'affair', 'issue', further studies have widened these list (see López Samaniego, 2014). Some of these GN are prototypical of spoken language (Swales, 2001, on 'thing') and are not found in our corpus. Therefore the category GN is widened to other vague nouns which are usually present in journalistic prose: 'strategy', 'case', 'reason', 'adventure', 'idea', etc. Some authors such as Pecorari (2014a, 2014b, 2015a, 2015b) consider that it is difficult to draw a clear-cut semantic distinction between AEs whose syntactic core is a GN and pronouns such as *esto*, *eso*, *aquello*, *ello* ('this, that, it') usually called 'textual anaphors' (or better 'metatextual anaphors'). Both encapsulate a text segment and unlike other Aes do not categorize the encapsulated segment to the extent that they do not give a label to the new referent. This has led this author to the conclusion that AE may have lexical or pronominal nature (Pecorari, 2014a; also, Moreno, 2004, talks about 'fuzzy labels' in this last case). I do not agree with this point of view: notwithstanding the referential vagueness of GN, the abstract entity denoted by *motivo*, *razón*, *situación*, *objetivo*, *fin*, 'reason, situation, objective', etc., which are usual GN in the journalistic texts, has an incomparably higher degree of precision that the absence of denotation carried out by textual anaphors. More puzzling is the difference between GN and hypernyms because this lexico-semantic differences are blurred in discourse where a noun function as a hypernyms in a particular text and the relation is not systemic but fostered by the world knowledge shared by the participants in the communicative situation. In this regard, I follow López Samaniego's (2014) criterion according to which a nominal core is a hypernym if the antecedent has also nominal nature; otherwise, as in the case of AE, is preferable to speak of GN.

²⁶ For an overview of the different attempts to define discourse topics, Charolles (2020).

²⁷ 'The AE's referent usually coincides or is linked to the general discourse topic'.

²⁸ I am very grateful to Filippo Pecorari for this observation.

²⁹ In cognitive psychology, this process is explained on the basis of a mental model, i.e., a series of mental representations that the reader builds and stores in the short-term memory during the text comprehension processs (López Samaniego, 2013).

³⁰ The complexity of the AEs' interpretation process has motivated the term 'pragmatic anaphors' (Conte 1996, 1999).

³¹ According to Loureda et al. (this volume), in eye-tracking experiments, there are no differences in the processing of AE and coreferential anaphors in total reading times (the differences emerge in the phases of the reading process), but the processing of demonstrative pronouns (what I call textual or metatextual anaphors) is shown to demand higher reading time than AEs, i.e., they are harder to link to their antecedent.

³² AEs violate some of the presuppositions of Ariel's Accessibility Theory. For example, NPs containing a higher quantity of lexical information are used to retrieve highly accessible referents in the immediately preceding Utterance. This excess of information allows the writer to recategorize the referent and introduce evaluative content in the NP (González & Izquierdo Alegría, 2020).

[33] There are, of course, other positions in the Utterance, mainly Topic in propositions in the Frame and the Appendix Units, that can be occupied by AEs so the scale will surely be completed in future studies.