

# Factual Shell Noun Constructions in Discourse

## *Construcciones de encapsuladores factuales en el discurso*

Eladio Duque Gómez

UNIVERSIDAD COMPLUTENSE DE MADRID  
ESPAÑA

eladio.duque.gomez@ucm.es

<https://orcid.org/0000-0001-9964-2895>

**Recibido:** 3-9-2023 / **Aceptado:** 15-12-2023

**DOI:** 10.4067/S0718-09342024000100311

### Abstract

Shell nouns tend to occur in schematic constructions or recurrent patterns, which can operate at the sentence level or across sentences. Although these constructions are not randomly realized with specific shell nouns, a debate exists regarding whether shell nouns with similar semantic features attract similar constructions. Considering that, we intended to study the Spanish factual shell nouns *hecho* (fact) and *caso* (case), in relation to the across-sentences constructions in which they are included and the discursive behavior these constructions exhibit. The analysis is based on the construction of discourse through coherence relations and is conducted by means of a corpus study. The results show that *caso* shows a greater attraction than *hecho* for constructions that operate across sentences. It is primarily realized through the prepositional construction *en este caso* (in this case) and the identificational construction *este es el caso de* (this is the case of), which establish Condition and Background coherence relations. *Hecho*, on the other hand, shows a greater attraction for the characterizational construction *este es un hecho* (this is a fact) and for verbal constructions, such as *este hecho ha de entenderse* (this fact is to be understood), which organize discourse through Evaluation and Interpretation relations. The findings indicate that although they belong to the same semantic class, shell nouns tend to display distinctive constructional profiles and discourse behavior, which depends on the specific constructions in which they are included rather than on the shell noun.

**Keywords:** anaphoric encapsulators, discourse labels, discourse constructions, rhetorical relations, coherence relations.

### Resumen

Los encapsuladores nominales tienden a aparecer en construcciones esquemáticas o patrones recurrentes que pueden tener alcance intra- o interoracional. Aunque estas construcciones no se materializan de forma aleatoria con encapsuladores específicos, existe el debate sobre si los encapsuladores que presentan características semánticas semejantes atraen a las mismas construcciones. Teniendo esto en cuenta, este trabajo

persigue estudiar los encapsuladores factuales en español ‘hecho’ y ‘caso’ en relación con las construcciones con alcance interoracional en las que se incluyen y las funciones discursivas que estas construcciones desempeñan. La propuesta se lleva a cabo mediante un análisis de corpus y se fundamenta en la construcción del discurso mediante relaciones retóricas. Los resultados muestran que ‘caso’ presenta una mayor atracción que ‘hecho’ por las construcciones con alcance discursivo y se materializa, principalmente, en la construcción preposicional ‘en este caso’ y en la construcción copulativa identificativa ‘este es el caso de’ que organizan el discurso, respectivamente, mediante relaciones de Condición y de Fondo. ‘Hecho’, por su parte, presenta una mayor atracción por las construcciones verbales, como ‘este hecho ha de entenderse’ y por la construcción copulativa caracterizadora ‘esto es un hecho’ que organizan el discurso mediante relaciones de Evaluación e Interpretación. Estos resultados demuestran que, a pesar de que los encapsuladores estudiados forman parte de la misma clase semántica, su perfil construccional es diferente y su comportamiento discursivo depende de las construcciones concretas en las que estos se incluyen.

**Palabras clave:** encapsuladores anafóricos, etiquetas de discurso, construcciones de discurso, relaciones de discurso, relaciones de coherencia.

## INTRODUCTION

‘Shell nouns’ (Schmid, 2000), also known as ‘discourse labels’ (Francis, 1994), ‘signaling nouns’ (Flowerdew, 2003), and ‘metadiscursive nouns’ (Jiang & Hyland, 2015), are nominal anaphors whose anaphoric expression is an unspecific abstract noun and whose referent or shell content is a complex notion, usually expressed in a clause or a larger fragment of discourse (i.e., ‘the antecedent trigger’, Cornish, 2022). Because «shell nouns and the clauses and sentences expressing shell contents activate closely related parts of the cognitive models created by texts» (Schmid, 2000: 80), they cannot be identified by their inherent lexical meaning but by their textual behavior and, to a lesser degree, by the specific constructions in which they tend to appear.

Although shell nouns (hereafter SNs) have been automatically identified through certain constructions (Schmid, 2000; Hunston & Francis, 2000) and their main function is primarily cohesive (Flowerdew, 2003; Aktas & Cortés, 2008), most research has not focused on SNs’ constructions as a whole or on their specific roles in discourse organization. Therefore, some authors (e.g., Schmid, 2000: 379) advise that these constructions or «patterns (...) must be looked at less mechanically by looking at grammatical structures rather than just sequences of words» whereas others (e.g., Flowerdew & Forest, 2015: 6-7) advocate «the need for more work» in the role of SNs «in contributing to textual development and coherence.»

Taking both considerations into account, we analyze SNs in the constructions they are part of and focus on those constructions that play a primary role in discourse organization. According to the literature (Flowerdew & Forest, 2015; Duque, 2022), these constructions are: ‘referring item + SN’ and ‘referring item + be + SN’. The latter is mainly realized by a definite phrase with a proximal demonstrative and an

optional adjective, such as ‘this demanding context’ in example 1. The former is mainly realized by a demonstrative pronoun acting as a subject, such as ‘this’ in example 2. Both constructions are opposed to others, such as the ‘SN + postmodifying that-clause’ (example 3), which operate solely within sentence boundaries. It is important to note that the examples throughout this article, extracted from the TenTen corpus family (Jakubiček, Kilgarriff, Kovář, Rychlý & Suchomel, 2013), feature the SN in bold and the antecedent trigger underlined.

- (1) Today, higher education institutions are tasked with multiple and challenging roles. They accommodate evolving student expectations, address changing labour market needs and sustainability pressures, and face unexpected challenges such as the Covid-19 pandemic. In this demanding **context**, the NEWLEAD project aims to build the capacity of university leaders in steering change and in addressing new priorities on the institutional transformation agenda (eua.eu).
- (2) Start by in-lining very small scripts directly in the HTML code. This is the most effective **way** to handle small scripts (www8.org).
- (3) I hated the **idea** that a book about censorship was going to be censored (hnn.us).

The main role of the across-sentences SN constructions (e.g., 1 and 2) is to shape textual organization, and this role is responsible for the high frequency of SNs in discourse (Stubbs, 2012). However, the research on the textual organization of SNs’ functions has primarily taken a broad approach, viewing them as nodal points (Conte, 1996) or signposts (Schmid, 2000) of text structure rather than specific connections in terms of explicit coherence relations (Taboada, 2009; Duque, 2014; Hoek, Zufferey, Evers-Vermeul & Sanders, 2018), as we do in this article.

To sum up, our intention is to study the discursive SN constructions by considering the coherence relations they establish. The object of study is limited to: (a) the SN constructions that operate across sentences, and (b) these constructions including the Spanish factual SNs *hecho* (fact) and *caso* (case). Consequently, we also aim to contribute to the debate on whether SNs with similar semantic features attract similar constructions. The study was conducted on a corpus in Spanish, a language in which research on SNs, mostly called *encapsuladores*, has reached a considerable degree of expansion and complexity in the last few decades (Borreguero, 2006; Borreguero & Octavio de Toledo, 2007; Llamas, 2010; Izquierdo-Alegría & González-Ruiz, 2013; Dam, 2014; López Samaniego, 2014; Ribera & Marín, 2018; Muñoz & Ciapuscio, 2019; González-Ruiz & Izquierdo-Alegría, 2020; López-Mora & García-Aguiar, 2020; Urizar-Ocampo & Rodríguez-Sánchez, 2022).

The article is structured as follows. The theoretical framework is divided into three subsections: 1.1. '*Caso* and *hecho* as factual shell nouns' describes the class of factual SNs and the concerns of previous studies on these two SNs being analyzed; 1.2. 'Shell noun constructions' provides an overview of the main lexico-grammatical SN constructions reported on the literature; 1.3. 'Shell noun constructions in discourse' presents an in-depth examination of the constructions that have a discursive scope in relation to their functions in textual organization. Meanwhile, Section 2, 'Methods,' describes the corpus used and the analysis procedure. Next, Section 3, 'Results and discussion,' presents an examination of the main findings of the corpus study conducted. Finally, the article ends with the conclusions and future research avenues.

## **1. Theoretical framework**

### **1.1. '*Caso*' and '*hecho*' as factual shell nouns**

In the main taxonomies of SNs (Francis, 1994; Schmid, 2000; López Samaniego, 2014, 2015; Borreguero, 2018, for Spanish), *hecho* (fact) and *caso* (case) are part of the same groupings, illustrating their aforementioned similarity. Schmid (2000) outlines a comprehensive semantic categorization, setting apart 5 abstract attributes (factual, mental, linguistic, modal, and eventive) as the «first and most important meaning components shared by larger sets of SN uses» (Schmid, 2000: 88). The attribute [factual], inherent in the SNs 'fact' and 'case,' is used to indicate «which can be said to be the case» (Schmid, 2000: 89). Factual SNs are divided into six types (neutral, causal, evidential, comparative, partitive, and attitudinal), with 'fact' and 'case' being prototypical examples of the neutral subcategory. Flowerdew and Forest (2015) outline a similar semantic categorization, in which the terms 'fact' and 'case' precisely designate the overarching category and the corresponding subcategory containing both SNs: 'fact' and 'case' are included in the 'fact' category, «which represents information about the world, without framing this information in terms of mental or verbal activity», and both also belong to the subcategory 'cases' which correspond to non-modalised propositions», as opposed to «chances, proofs, and needs that exhibit some form of modality» (Flowerdew & Forest, 2015: 31).

*Hecho* and *caso* are also similar nouns according to major Spanish lexicographic references. The Royal Academy dictionary (RAE, 2014) defines *hecho* as an 'action or event' and *caso*, as an 'event, occurrence'. Since both nouns provide minimal semantic content, when used as SNs, they should imply a minor alteration of the antecedent or state of affairs evoked. This semantic emptiness is further associated with their significant usage frequency, placing them in the category that Schmid (2000: 85) names 'prime shell nouns,' those that embody «the central core of the class of shell nouns» in terms of degree of typicality. Likewise, the combination of their limited semantic contribution and high use frequency make *caso* and *hecho* strong candidates for grammaticalization processes.

Few works in Spanish delve into the discursive behaviour of *hecho* as a SN (Rodríguez-Españera, 2015; Urizar-Ocampo & Rodríguez-Sánchez, 2021). Rodríguez-Españera (2015) points out that some uses of *hecho* are far from neutral because they result in the shell contents being assumed, unquestionable, and non-debatable. Similarly, González-Ruiz and Izquierdo-Alegría (2020) note that even though the SN *hecho* lacks evaluation, it can perform argumentative functions because its objectivizing effects make it difficult for the reader to deny the shell content's factuality. Overall, these conclusions fit within the discourse behavior of the English SN 'fact,' at least in the most studied construction 'the fact that' (Schmid, 2000; Gray & Biber, 2014; Jiang & Hyland, 2015).

Regarding the noun *caso*, there is no specialized research on its role as a SN, but research has been conducted focusing on its discursive behavior. Along the same lines, Fuentes-Rodríguez (1996) explores the discourse markers that include the lexeme *caso* and concludes that they are situated in the realm of hypotheses, whether as conditional connectors—*en tal caso* (in such a case)—or as non-operational conditions or concessions—*en cualquier caso* (in any case). The conditional value she finds for *caso* also aligns with the characterization of the English SN 'case'. According to Schmid (2000: 289), 'case' evokes associations with the domain of contingency of events, which is why it is placed not only in the 'factual-neutral' category, but also in the 'situational-conditional' category that is «used to shell the dependence of one event on another.»

To recapitulate, the lexicographic definitions of *caso* and *hecho* are comparable, and both are classified similarly, as prime, factual and neutral SNs, which should affect the constructions in which they are included, since various authors (Hunston & Francis, 2000; Yamasaki, 2008; Gries, 2019) predict that SNs that are semantically close tend to appear in the same constructions (i.e., similar SNs attract similar constructions). However, the two SNs also present certain semantic differences, to some extent derived from their different formation processes. *Caso* often refers to events, circumstances, or situations, but not as a direct result of the verb from which it derives (to fall, from the Latin *cadere*), as it has an independent nominal development from its verbal origin. Meanwhile, *hecho*, as a participial noun from the verb *hacer* ('to do'), is understood, to a greater extent, as the direct result of the action of the verb.

## **1.2. Shell noun constructions**

In this article, the notion of construction is interpreted in a broad sense, in line with usage-based construction grammars (Goldberg, 2006; Hoffman, 2022). That requires two clarifications. First of all, frequency of use is more relevant than emergent meanings (i.e., «word sequences that are often used are constructions even if they do not have idiosyncrasies of meaning or form» [Bybee, 2013: 51]). Secondly, «constructions vary across a continuum of schematicity or abstractness» (Diessel,

2019: 53). At one end of the continuum, there are ‘substantive constructions,’ which are specific instances or tokens that contain actual lexical items; at the other end, there are ‘schematic constructions,’ which are flexible structures realized in a variety of specific instances. Keeping this in mind, the frequent patterns in which SNs tend to occur can be understood as constructions, and they can be classified according to lexico-grammatical (Schmid, 2000) or discursive (Flowerdew, 2003) criteria.

This section focuses on the classifications based on grammatical criteria. In practical terms, we understand these schematic constructions as the queries that have served for the identification of SNs and, as substantive constructions, may or may not develop emergent meanings. The two main schematic constructions for detecting SNs in English are: ‘SN + complement clause’ (e.g., 4) and ‘SN + be + complement clause’ (e.g., 5). Two other constructions frequently mentioned in the literature, although less productive in identifying SNs, are: ‘th-SN’ (‘referring item + sn’) and ‘th-be-SN’ (‘referring item + be + SN’), which can address the anaphoric reference in the sentence or discourse scope (see examples 6-9). The referring items in the th-SN construction are ‘the,’ ‘this,’ ‘that,’ ‘other,’ ‘same,’ and ‘such.’ The referring items in the th-be-SN construction are the pronouns ‘this,’ ‘that,’ and ‘it.’ In this construction, strictly speaking, the pronoun is the anaphoric element, which transfers the characterization of the shell content to the SN through the copula. Therefore, in the examples illustrating this construction, the underlined fragment corresponds to the antecedent trigger of the pronoun and not the SN.

- (4) SN + COMPLEMENT CLAUSE: The turmoil in toy land has also led to **speculation** that Hasbro and Mattel might merge (kq2.com).
- (5) SN + BE + COMPLEMENT CLAUSE: The **problem** is to get doctors to understand the project and what we are trying to do (nlma.nl.ca).
- (6) TH- SN (sentential scope): Norma’s hands and feet were abnormally large and powerful for a girl of her build, but this **fact** enhanced rather than impeded her career (carmelbird.com).
- (7) TH-SN (discursive scope): Sometimes, Excel Spreadsheets contain links to pictures, which cannot be found. This **situation** caused an error in XLSDump in previous versions (gams.com).
- (8) TH-BE-SN (discursive scope): The problem is my parents are very conservative (our family has more arranged marriages). The boy is not of our community, and though we share the same religion, he and his folks eat non-vegetarian food, and we are strict vegetarians. That is the main **cause** of objection from my parents (loveandlearn.com).

- (9) TH-BE-SN (sentential scope): Our city has so much to offer and it is the **reason** that creative, entrepreneurial and visionary people see the promise in Providence (gcpvd.org).

Some of these schematic constructions have developed specialized functions (i.e., associations between form and meanings at the level of the construction). The construction ‘th-SN’ has become specialized in textual linkage (Schmid, 2000; Gray, 2010), and the construction ‘th-be-SN’ has developed a more detailed linkage function: «to characterize a piece of information in a certain way» (Schmid, 2000: 37). The construction ‘SN + be + complement clause’ has focusing effects on the shell content (Tárnyiková, 2018), and the construction ‘SN + complement clause’ has not developed a single dominant function (Schmid, 2000) although some of its associated substantive constructions, such as ‘the fact that,’ have acquired idiosyncratic meanings as ‘stance markers’ (Gray & Biber, 2014; Jiang & Hyland, 2015).

Overall, research on SNs has highlighted four lexico-grammatical constructions in which they tend to occur: (1) ‘SN + complement clause’ (e.g., the fact that), (2) ‘SN + be + complement clause’ (e.g., this fact is), (3) ‘referring item + SN’ (e.g., that fact), and (4) ‘referring + be + SN’ (e.g., it is a fact). These constructions can exhibit broad meaning-form pairings, and some of these pairings involve the potential for organizing discourse, an issue we will examine in more detail in the following section.

### **1.3. Shell noun constructions in discourse**

When SNs «require specific meanings across clause border, they have the potential to perform discourse connective functions at a global level and play dynamic roles in discourse organization» (Yamasaki, 2008: 76). Some authors have explored associations between particular SNs and the textual connections they form. Flowerdew and Forest (2015), for instance, based on conjunctive relations from systemic functional linguistics (Martin, 1992), conclude that associations between SNs and particular relations is probabilistic. In our view (Duque, 2016), although some SNs, such as ‘reason’ (Hoey, 1993), can very often be associated with specific textual connections, the correspondences between SNs and relations are mediated by the constructions in which the SNs occur.

The constructions in which SNs have received the most attention—because they are effective discriminators of SN status for automated analysis (Benítez-Castro, 2015; Flowerdew & Forest, 2015)—are those that display the shell content and the SN in the same clause or nominal group (‘SN + complement clause’ and ‘SN + be + complement clause’), and therefore do not play a prominent role in the organization of discourse. On the contrary, the constructions we focus on in this article (‘th-SN’ and ‘th-be-SN’) can operate at the discourse level, and they serve as the primary means of organizing discourse. In any case, a more detailed classification of these

constructions is required if our objective is to find associations between constructions and specific forms of connections (Duque, 2022).

In this regard, the schematic construction ‘th-SN’ can occur in two main constructions: one prepositional and one verbal (e.g., ‘in this demanding context,’ example 1) and one verbal (e.g., ‘this situation caused,’ example 7). For further illustration, see example 10 and 11 below. These constructions operate similarly to connectives. They function in the explicit cohesion of texts (Stubbs, 2012) and predominantly serve a metadiscursive purpose (Hyland, 2017). However, unlike other grammaticalized connectives—specialized in the procedural or relational function—in these constructions, referential and relational coherence interact (Montolío, 2013; Parodi, Julio, Nadal, Cruz & Burdiles, 2019; Recio-Fernández, Loureda & Sanders, 2021). The interaction entails the division of tasks between the SN, which works within the referential scope, and the rest of the construction that usually introduces the specific connection, either through the verb’s semantic contribution or through the preposition or prepositional phrase.

- (10) PREPOSITIONAL TH-SN CONSTRUCTION: On 9 February Force H carried out a naval and aerial bombardment of Genoa, Leghorn and Spezia (using the Renown, Malaya, Ark Royal and Sheffield), returning to Gibraltar without suffering a single casualty. Despite this offensive **move**, for most of the year the reinforcement of Malta would remain the key task (historyofwar.org).
- (11) VERBAL TH-SN CONSTRUCTION: On Monday, the Spanish FA assembly confirmed that from 2020 the Super Cup will include four teams - the league's champions and second place team, as well as the cup winners and runner-up. This **decision** means that the Spanish Super Cup will transform into a tournament played over a few days (alaraby.co.uk).

When we look at these more substantive constructions, the general connective role of SNs can be described in terms of specific coherence relations, which are triggering renewed interest in discourse linguistics in Spanish (Rodríguez-Ramalle, 2015; Iruskieta, Da Cunha & Taboada, 2015; Garrido & Rodríguez-Ramalle, 2015; Duque, 2016; Santana, Spooren, Nieuwenhuijsen & Sanders, 2018; Mancera-Rueda, 2019; Ibáñez, Moncada, Cárcamo & Marín, 2020; Tordera Yllescas, 2021; García-Pérez, 2022; Fuentes-Rodríguez & Pérez-Béjar, 2022). This line of research—associations between SN constructions and coherence relations—is taken by current developments in discourse relation annotation (Das & Taboada, 2018; Webber, Prasad, Lee & Joshi, 2019), which consider some SN constructions relation-signaling devices, ‘secondary connectives’ (Rysová & Rysová, 2018) or ‘discourse constructions’ (Ruiz de Mendoza & Gómez-González, 2014). In these constructions, the SNs are the ‘variable element’, and the verbs and prepositions are the ‘fixed elements’, which: «capture relational



meaning grounded in high level cognitive models (e.g., addition, exemplification, contrast causa-consequence, etc.)» (Ruiz de Mendoza & Gómez-González, 2014: 303).

Duque (2020, 2022), following the Rhetorical Structure Theory, identifies in corpus studies the verbs and prepositional locutions that in Spanish are usually combined with pronominal encapsulators to set up discourse relations. These include verbs such as *provocar* (to provoke), *generar* (to generate), *contribuir* (to contribute), and *deberse a* (to be due to), which usually establish relations of Cause or Result; *demostrar* (to show), *manifestar* (to manifest), *indicar* (to indicate), and *reflejar* (to reflect), associated with relations of Evidence; *lograr* (to achieve) and *permitir* (to allow), which usually establish relations of Purpose or Means; *suponer* (to suppose), *explicar* (to explain), *implicar* (to imply), and *significar* (to mean), which usually signal relations of Interpretation; *añadir* (to add) and *depender* (to depend), which usually form relations of List and Condition respectively. Regarding prepositions and prepositional locutions, *por* (for), *gracias a* (thanks to), and *debido a* (because of) establish relations of Cause and Result; *frente a* (against), *en vez de* (instead of), *en lugar de* (rather than), and *a pesar de* (in spite of) build relations of Contrast and Antithesis; *ante* (in light of), *en* (in); *por encima de* (above), and *aparte de* (apart from) establish relations of List and Background; *para* (for, to) signals Purpose and Means relations; and *tras* (after) Sequence relations.

In summary, most theoretical approaches to SNs have in common that one of their main functions is the organization of discourse. However, few authors have explored in depth the constructions specialized in this function or have studied the particular forms that discourse organization can take in terms of coherence relations. In this regard, the SN constructions with the greatest potential in discourse organization are ‘th-be-SN’ and ‘th-SN’. To study constructions as signals of relations, it is useful to distinguish within the ‘th-SN’ construction two separate constructions: the verbal construction ‘th-SN + verb’ or ‘verb + th-SN’ and the prepositional construction ‘prep + th-SN’. In these constructions, the lexical contribution of the verb or prepositional locutions—to a greater extent than the SN—guides the coherence relation that is established.

## 2. Methods

This study includes a corpus analysis that is intended (a) to examine the distribution of the constructions in which the factual SNs *becho* and *caso* are included, and (b) to describe the discursive behavior in the textual organization of these constructions. The guiding hypothesis is that, despite both SNs being recognized as factual shell nouns, they shape different coherence relations, due to the distinctive semantic and syntactic features of the constructions in which *becho* and *caso* are included. The analysis follows a semi-automated approach, which is the most widely used for examining SNs (Flowerdew & Forest, 2015; López Samaniego, 2018) and is carried out on the esTEnTen18 corpus from Sketchengine (Kilgarriff & Renau, 2013).

It is a corpus of 16.9 billion words, broken down in 19 Spanish varieties, lemmatized, part-of-speech tagged, and enabled to search complex lexico-grammatical patterns using its own code or query language.

To detect the constructions of our interest automatically, we selected from the patterns discussed in the literature (see section 2) those that usually operate at the discourse level: ‘th-SN’ and ‘th-be-SN.’ These are hereafter referred to as ‘determinative construction’ (e.g., *este hecho* or *ese caso*) and ‘pronominal construction’ (e.g., *ese es el caso* or *este fue un hecho*) to fit the terminology with the actual constructions in Spanish. In the determinative construction, the demonstratives *este/ estos, ese/ esos, aquel/ aquellos, dicho/ dichos*, and *tal/ tales* were considered. We did not consider the definite article. This methodological decision, which other authors have followed (Yamasaki, 2008; Izquierdo Alegria & González Ruiz, 2020, in Spanish), is grounded in the fact that whereas demonstratives tend to function anaphorically, the definite article does not tend to do so. Similarly, for the search of the pronominal construction, the entire set of demonstrative pronouns in Spanish were considered.

Once these patterns were automatically identified, a random sample of 400 concordances was selected: 100 of them correspond to the determinative construction with *caso*, 100 to the determinative construction with *hecho*, 100 to the pronominal construction with *caso*, and 100 to the pronominal construction with *hecho*. The construction of the sample and its size, 200 occurrences per SN, is similar to that of other comparable studies (150 per SN in Yamasaki, 2008, or 200 per SN in Vergaro & Schmid, 2017). In this sample, a manual analysis was conducted following the procedure outlined below. For the determinative construction, firstly, it was verified whether the SN *caso* or *hecho* acted as anaphors in all examples. Secondly, the shell content was identified to annotate the instances in which it appeared within the sentence or across sentences, taking the orthographic sentence into account. Thirdly, the more specific construction in which it appeared was annotated (see section 1.3), distinguishing between the verbal (e.g., *este hecho provocó*) and prepositional constructions (e.g., *en ese caso*). Likewise, adjective modifiers were annotated when they were part of the construction. Lastly, an exploratory analysis of the discursive behavior of each of the identified constructions was conducted, considering the specific coherence relations established, following the RST taxonomy (Mann & Thompson, 1988), and the specific verbs and prepositional phrases shaping substantive constructions.

The annotation procedure for the pronominal construction followed similar steps. Firstly, it was confirmed that the automatically identified instances corresponded to the target construction. Secondly, the shell content was identified to confirm whether it appeared in the same sentence as the construction or in a different sentence, also considering orthography as a criterion. Thirdly, adjective modifiers were annotated when they were part of the construction. Finally, an exploratory analysis of the

established coherence relations as well as the emerging meanings or pragmatic enrichment derived from the studied constructions was conducted.

In summary, the analysis procedure is designed to gather data about various aspects, including the frequencies of occurrence of *hecho* and *caso* in each specific construction; the verbs, prepositions and adjectives that typically participate in these constructions; the frequencies with which these constructions operate at the sentential or discursive level; and the specific coherence relations these constructions tend to build.

### **3. Results and discussion**

Although *caso* and *hecho* belong to the same semantic class as SNs, the results reveal divergent constructional profiles and discursive behaviors for each one. On one hand, in terms of constructional profile (section 3.1), *caso* displays a stronger affinity for SN constructions with a discursive scope, particularly favoring prepositional constructions. On the other hand, *hecho* exhibits a stronger preference for verbal constructions. Additionally, the pronominal construction of both SNs also differs, with *caso* typically appearing in identificative sentences whereas *hecho* tends to appear in characterizing ones. Concerning discursive behavior (section 3.2), the results confirm that it does not solely depend on the studied SNs, but primarily on the specific constructions in which they are integrated.

#### **3.1. Factual shell noun constructions**

*Caso* and *hecho* are among the 100 most frequent nouns in the entire corpus. Although an exhaustive manual analysis is needed to confirm their anaphoric behavior, this data supports their status as candidates for prime SNs in Spanish, as other corpus studies suggest (López Samaniego, 2014, 2018). The overall results reveal that the lemma *caso* is more than twice as frequent as *hecho*. Furthermore, as evidenced by the comprehensive data outlined in Table 1, when we focus on the raw results of the query searches for the constructions with potential discursive scope (determinative and pronominal), the ratio of the occurrences of *caso* to *hecho* is even higher (6,5 million vs. 15,3 million). These data predict that *caso* serves discursive functions to a greater extent than *hecho*. It is also noteworthy that in both cases, the determinative construction is more productive than the pronominal one (with normalized frequencies of 108 vs. 6.9 and 21 vs. 0.78), probably because the former is more flexible, allowing for multiple syntactic configurations.

**Table 1.** Raw search results from EsTenTen18 Corpus (\*NF=normalized frequencies per million words)

	<i>Caso</i>		<i>Hecho</i>	
	N	NF*	N	NF
<b>Lemma</b>	15,366,286	784.27	6,536,316	333.6
<b>Determinative construction</b>	2,117,972	108.1	425,669	21.73
<b>Pronominal construction</b>	135,195	6.9	15,381	0.78

The manual analysis of the 400-item sample indicates that nearly all instances of the lemmas *hecho* and *caso* behave as SNs, whether at the sentence or discourse level, as Table 2 shows. However, the query search used to detect the pronominal construction with *hecho* (demonstrative + *ser* + *hecho*) seems less effective because 47% of the concordances did not match the target construction. Among these discarded concordances, examples such as (12) stood out. In this construction, the demonstrative does not function as a pronoun but as a determiner of another SN, in the example *afirmación* (statement). Example 12 also illustrates how through the SN *hecho*, «speakers (...) accept, or even intend, the change of the conceptual status of the shell content» (Schmid 2000: 99), in this case from the domain of statements to that of facts.

**Table 2.** *Caso* and *hecho* as shell nouns in the sample corpus

	<i>Caso</i>			<i>Hecho</i>		
	N	Sentence (%)	Discourse (%)	N	Sentence (%)	Discourse (%)
<b>Determinative construction</b>	100	34	66	96	34.4	65.6
<b>Pronominal construction</b>	94	11.7	88.3	53	32.1	67.9
<b>Total</b>	194	23.2	76.8	149	33.6	66.4

(12) Los metales también recuerdan. Esta afirmación es un **hecho** que se encargan de demostrar y de desarrollar los científicos del grupo consolidado de Metalurgia Física y Transformaciones de Fase de la Universidad del País Vasco (UPV) (aimme.es).

(12') Metals also have memory. This statement is a **fact** that the research group of Physical Metallurgy and Phase Transformations at the University of the Basque Country (UPV) undertake to demonstrate and elaborate upon.

Regarding the sentential or discursive scope, and according to the information compiled in Table 2, references across sentences stand out (76% of the occurrences for *caso* and 66.4% of occurrences for *hecho*). The discursive scope is particularly prominent in the pronominal construction of *caso*, where it reaches 88%, thereby confirming the greater preference for discursive behavior of this SN. When we examine the determinative construction in detail (see Table 3), the results show that *caso* strongly favors the prepositional construction (97 vs. 1 occurrence) whereas *hecho* leans toward the verbal construction (49 vs. 11 occurrences). The results also illustrate

a greater diversity of substantive constructions for *hecho*, whereas nearly all instances of *caso* are aligned with *en este caso* (in this case) and its less common constructional variants such as *en ese caso* (in that case) and *en tal caso* (in such a case). These constructions seem to be undergoing a process of grammaticalization, evidenced by the fact that the noun *caso* never appears with adjectival modifiers although plural forms such as *en estos casos* (6 occurrences) are reported. Among the less frequent prepositional constructions of *hecho*, the substantive construction *por este hecho* (due to this fact) stands out.

**Table 3.** *Caso* and *hecho* determinative constructions in the sample corpus

	<i>Caso</i>			<i>Hecho</i>		
	N	Sentence (%)	Discourse (%)	N	Sentence (%)	Discourse (%)
<b>Prepositional construction</b>	97	32	65	19	8	11
<b>Verbal construction</b>	1	0	1	72	23	49

The verbal construction is mostly limited to the SN *hecho*. Its substantive constructions exhibit a wide variety, with very few repeated occurrences, except for *ser* (to be) (N=10), *entenderse* (to be understood) (N=3), *producir* (to produce, to cause) (N=3), *contribuir* (to contribute) (N=2), *constituix* (to constitute) (N=2), *poner de manifiesto* (to bring to light) (N=2), and *resultar* (to result) (N=2). Although some of these verbs also appear in analogous constructions with pronominal encapsulators (Duque, 2020, 2022), *hecho* displays a different constructional profile, framed within a clearly argumentative ‘semantic prosody’ (Bednarek, 2008), frequently negative, which is also demonstrated by its combination with verbs such as *criticar* (to criticize), *desmentir* (to disprove), *eludix* (to evade), *lamentar* (to regret), *negar* (to deny), and *renegar* (to disown). Furthermore, in these constructions, the noun *hecho* can take adjectival modifiers, but they are very infrequent in any case. These modifiers highlight the truth value (e.g., *evidente* (obvious)) or the importance (e.g., *fundamental* (fundamental)) of the shell content, which also fits with the aforementioned argumentative prosody.

Regarding the specific demonstrative that takes part in the verbal or prepositional construction, both SNs show a clear preference for the proximal demonstrative *este*, and, to a lesser extent, the medial demonstrative *ese*. Our sample contains no occurrences with *dicho* (such, aforementioned), and the distal demonstrative *aquel/aquellos* only appears once with *hecho*. With *caso*, distal demonstratives form a prepositional construction with different behavior from *en este caso*, because it specifies its reference through a post-modifying prepositional phrase (see the contrasting examples, 13 and 14). On one hand, the specialized demonstrative determiners in the formation of shell nouns are the proximal and medial ones, which are often interchangeable. The distal demonstrative in Spanish, on the other hand, has little influence in constructions with SNs, and when it does, it may not be interchangeable with the other demonstratives, but instead, it usually forms an independent

construction, which operates within the scope of the sentence and with cataphoric resolution, as illustrated by example 13.

(13) En aquellos **casos** en que Docshare pudiera tener acceso a los datos, se compromete a no aplicar, utilizar o revelar los datos tratados (docshare.es).

(13') In those **cases** where Docshare could have access to the data, it commits not to apply, use, or disclose the processed data.

(14) Si estas molestias persisten, se aconseja suspender el uso del producto hasta que desaparezcan los efectos secundarios. En este **caso**, una opción aconsejable es cambiar a una dosis más baja (acne.org.es).

(14') If the discomforts persist, it is advised to discontinue the use of the product until the side effects disappear. In this **case**, a sensible option is to switch to a lower dosage.

Finally, regarding the pronominal construction, *caso* and *hecho* occur in distinct schematic constructions. 'Caso' operates to a greater extent in the discursive realm (see Table 2 data) and typically appears in the form of equative or identificational sentences, with a definite article followed by its specification (e.g., *este es el caso de Madrid* [this is the case of Madrid]). Its negative form (i.e., *este no es el caso* [this is not the case]) is usual, as well [15%]). On the other hand, *hecho* typically appears in characterizational or predicational sentences, preceded by an indefinite article and followed by an adjectival modifier (e.g., *este es un hecho confirmado* [this is a confirmed fact]). The negative constructional variant is rare in the corpus.

Similar to the verbal construction, the frequent modifiers of *hecho* in the pronominal construction highlight its prominence (*importante* (important), *fundamental* (fundamental)) and, above all, its factuality (*indiscutible* (indisputable), *real* (real), *constatable* (verifiable), *objetivo* (objective), *comprobable* (demonstrable), *constatado* (confirmed), and *demostrado* (proven)), which also confirms this construction's argumentative discursive prosody. Conversely, adjectival modifications are not usual in the pronominal *caso* construction.

### **3.2. Factual shell noun constructions in discourse**

Considering that *caso* and *hecho* exhibit a separate constructional profile, this section presents an explorative view of the discursive behavior of the constructions highlighted in the previous section: the verbal constructions with *hecho* (section 3.2.1), the prepositional constructions with *caso* (section 3.2.2), and the pronominal constructions with *hecho* and *caso* (section 3.2.3). The presented results focus exclusively on the exemplars that operate across sentences.

### 3.2.1. Verbal constructions

The verbal construction is the most frequent and therefore central in the discursive behavior of *hecho*, and it is very infrequent and therefore peripheral for *caso*. The performance in the organization of the discourse of the verbal construction with *hecho* varies depending on its specific substantive constructions. This confirms that the verbs that fill the slot of the schematic construction (listed in section 3.1), more than the SN itself, decide the coherence relation established.

*Hecho* verbal constructions tend to signal coherence relations that also confirm the argumentative discursive prosody of which this SN usually forms a part: Evaluation, Interpretation, Cause, and Evidence relations. The most repeated or central exemplar is the substantive construction with the verb *ser* (e.g., *este hecho es* [this fact is]) which usually establishes Evaluation relations between sentences, connecting a situation with ‘an evaluative comment about the situation’ (see the detailed definitions of the relations in Mann & Thompson, 1988), as example 15 illustrates. Interpretation relations established through constructions with verbs such as *constituir* (to constitute) and *entenderse* (to be understood) are also common, such as *este hecho ha de entenderse* (this fact must be understood) in example 16. Cause relations between sentences are established with *hecho* constructions that include verbs such as *generar* (generate) and *producir* (produce) (e.g., 17) as well as evidence relations with constructions that include verbs such as *poner de manifiesto* (to reveal), as in example 18 with *este hecho pone de manifiesto* (this fact reveals). In summary, verbal constructions with *hecho* present a varied discursive behavior in the organization of discourse, ranging from Evaluation to Evidence, passing through Interpretation and Causes.

- (15) [E]n Las Ramblas y otros lugares de Barcelona los comerciantes han decidido vender souvenirs con incuestionable olor a español (...). Este **hecho** es altamente preocupante (vozbcn.com).
- (15') In Las Ramblas and other places in Barcelona, shops have decided to sell souvenirs with an unmistakable Spanish scent (...). This **fact** is highly concerning.
- (16) (...) la guerra y sus consecuencias e interpretaciones, se constituyen en tema central. Este **hecho** ha de entenderse de dos modos: por un lado, como temática específica de un determinado tipo de obras (...); por otro (...) como un prisma evaluador manejado por el censor a partir de la nueva situación creada por la guerra (represura.es).
- (16') War and its consequences and interpretations are a central theme. This **fact** must be understood in two ways: on one hand, as a specific subject matter of a certain type of books (...); on the other (...), as an evaluative prism used by the censor in light of the new situation created by the war.

- (17) La crisis de la industria textil, a finales de los años setenta, provocó el cierre de numerosas empresas, dejando en el paro a gran parte de la población. Este **hecho**, unido a la escasa calidad urbana de los barrios obreros, generaron fuertes movimientos vecinales que exigieron mejoras en la calidad de vida por parte de la administración (upm.es).
- (17') The crisis in the textile industry, toward the end of the 1970s, led to the closure of numerous companies, leaving a large portion of the population unemployed. This **fact**, coupled with the poor urban quality of working-class neighborhoods, led to strong community movements demanding improvements in the quality of life from the administration.
- (18) Si realizamos una búsqueda en la base de datos del AGA sobre Tomás Borrás (...) se obtienen 19 registros que corresponden tanto a obras de producción propia como a adaptaciones de obras extranjeras que también le han sido atribuidas. De esta forma, hemos recuperado registros en los que no aparece el nombre del autor original (...). Este **hecho** pone de manifiesto que los límites entre traducción y adaptación eran muy difusos para los censores, por lo que es necesario afinar las búsquedas (represura.es).
- (18') If we conduct a search in the AGA database on Tomás Borrás (...) we obtain 19 records that correspond to original works and adaptations of foreign works that have also been attributed to him. In this way, we have recovered records in which the name of the original author does not appear (...). This **fact** reveals that the boundaries between translation and adaptation were very blurry for the censors, so it is necessary to refine the searches.

### 3.2.2. Prepositional constructions

The prepositional construction is central in the discursive behavior of *caso*, due to the high frequency of the nominal construction *en este caso*. This construction displays various discursive functions in our corpus, not always clearly differentiated, which depend on the contexts in which it is included. In general, the noun *caso* opens up a range of options in the available mental model, affecting the discursive behavior of the constructions in which it is included. In other words, the frame associated with the noun *caso* shows a base-profile relationship (Langacker, 2000) that affects the organization of the discourse, corresponding the base to a set of units and the profile to an individual unit of that set.

In the central performance of *en este caso*, the options in the set are either unrealized or hypothetical, so the construction is used to organize discourse through Condition relations. The construction is typically placed in an initial thematic position and allows for variants, such as *en ese caso* (in that case) and *en tal caso* (in such a case). In these Condition relations, the reference of *caso* serves as the protasis, as example 19 shows.



The references to *caso* in these exemplars are more inferential and less explicit or textual than those of the constructions with *hecho*. It is also worth noting that *en este caso* appears to have a high propensity to occur in certain textual genres, such as user instructions and terms of service, as example 19 illustrates.

- (19) Dichas comunicaciones comerciales serán relacionadas sobre productos o servicios ofrecidos por el prestador, así como por parte de los colaboradores o partners con los que éste hubiera alcanzado algún acuerdo (...). En este **caso**, los terceros nunca tendrán acceso a los datos personales (ocoval.es).
- (19') Such commercial communications will pertain to products or services offered by the provider as well as by collaborators or partners with whom the provider has reached some commercial promotion agreement (...). In this **case**, third parties will never have access to personal data.

Another less numerous exemplar cluster of *en este caso* consists of instances in which the construction does not function as a conditional marker but as a specificity operator or modifier. It primarily acts over noun phrases with which it is positioned parenthetically, as example 20 shows. It is important to state that these constructions are not interchangeable with *en ese caso* (in that case) or *en tal caso* (in such a case). The references to *caso* in these examples are highly inferential because they often present a vague anaphoric connection of the discourse topic. Although these references have a discursive nature and propose a reactivation of the prior topic, these constructions are not specialized in establishing specific coherence relations.

- (20) Un varón de 41 años de edad (...) asciende a un poste eléctrico de su finca, para hacer una reparación casera. (...) Unos minutos más tarde es encontrado por sus familiares en el suelo, junto al pilote, con una herida en la cabeza (...) La “Marca Eléctrica (...)” puede aparecer en el punto de entrada o salida de la corriente y es de gran ayuda para el diagnóstico de la causa de la muerte. La Marca puede reproducir el objeto conductor, en este **caso** la porción cóncava del mango de unos alicates, y presenta bordes sobreelevados con un lecho deprimido de coloración blanco-grisácea (cuadernosdemedicinaforense.es).
- (20') A 41-year-old man (...) climbs an electrical pole on his property to make a home repair. (...) A few minutes later, he is found by his family on the ground, next to the pole, with a head injury (...) The “Electrical Mark” (...) can appear at the point of entry or exit of the electric current and is very helpful for diagnosing the cause of death. The mark can replicate the conductive object, in this **case** the concave portion of the handle of some pliers, and features raised edges with a depressed, grayish-white bed.

### 3.2.3. Pronominal constructions

As indicated in the previous section (3.1), pronominal constructions take two forms for *hecho* and *caso*. Pronominal constructions with *hecho* are primarily characterizational or predicational copulative sentences whereas constructions with *caso* are mainly equative or identificational copulatives. These two forms are also associated with separated functions in the organization of discourse: the *hecho* construction brings evaluations into the text whilst the *caso* construction often serves as a rhetorical device for presenting a new discourse topic.

Regarding the pronominal constructions with *caso*, the central example is *este es el caso de* which is typically used to organize discourse through Background relations. Such relations are part of the family of contiguity relations (Duque, 2016; Kehler, 2019), which are based on topical continuity. One of the members of the Background relation, usually the one that appears first, provides contextual information that allows for the correct understanding of the second member. In the exemplars from our corpus, these first sentences usually present a ‘general stative,’ typical of the informative discourse mode (Smith, 2003), which is specified in the second sentence that presents the topic of the discourse: the subject that one really wants to talk about. In other words, the construction functions as an introductory discursive formula for a particular discursive topic. Likewise, as it can be seen in example 20, this rhetorical configuration aligns with the aforementioned base and profile relations in that the base is the prior information and the profile is the one introduced by the construction.

- (21) Una enfermedad puede ser rara en una región, pero habitual en otra. Este es el **caso** de la talasemia, una anemia de origen genético, que es rara en el Norte de Europa, pero frecuente en la región del Mediterráneo (orphanet-espana.es).
- (21') A disease can be rare in one region but common in another. Such is the **case** with thalassemia, a genetic anemia that is rare in Northern Europe but frequent in the Mediterranean region.

Another exemplar cluster that stands out in *caso* pronominal constructions is *este no es el caso de* (it is not the case of), which organizes discourse through Antithesis relations. This construction also functions as an introductory formula for a new discursive topic, carrying even greater impact than *este es el caso de*, because it introduces an exceptional or unexpected theme, as example 22 indicates. Generally speaking, in the pronominal construction of *caso*, two substantive constructions stand out: *este es el caso de* and *este no es el caso de* which establish relations of Background or Antithesis. Both constructions are typically used as rhetorical devices to introduce a discursive topic, and both relations are associated with the informative mode of discourse (Smith, 2003), in contrast to the argumentative mode that characterizes SN constructions with *hecho*.

(22) Normalmente, creemos que un monasterio benedictino tiene una vida larga, multisecular, de lento y sosegado fluir. Este no es el **caso** de nuestro Montserratico de Madrid. Su historia sólo alcanza tres siglos y medio, en los que ha pasado por grandes turbulencias y dramáticos avatares (abadiadesilos.es).

(22') Normally, we believe that a Benedictine monastery has a long multi-secular life of slow and tranquil flow. This is not the **case** for our Montserratico in Madrid. Its history only spans three and a half centuries, during which it has gone through significant turbulence and dramatic vicissitudes.

The pronominal constructions with *hecho* form characterizational sentences and mostly establish coherence relations of Evaluation. The evaluation is usually realized with a that-clause (example 23) or with an adjective (example 24). The most frequent adjectives (listed in 3.1) highlight the significance of the reference of *hecho* and above all its facticity (*importante* (important), *constatable* (certifiable), or *indiscutible* (indisputable)). As the examples suggest, these constructions are used in argumentative discourses in which the evaluation introduced by the construction transforms arguments into facts of relevance that are apparently beyond dispute.

(23) Tanto en el caso catalán como en el vasco, parece indudable que el sentimiento nacionalista e independentista es compartido por muchos sectores de la población, incluyendo grupos de izquierda. Y este es un **hecho** que hay que tener en cuenta (andalan.es).

(23') In the Catalan and Basque cases, it seems unquestionable that nationalist and separatist sentiment is shared by many sectors of the population, including left-wing groups. And this is a **fact** that must be taken into account.

(24) Señor Robles, usted sabe muy bien que cuando se reduce la interinidad aumentan las jornadas parciales. Ese es un **hecho** lógico y natural (ccyl.es).

(24') Mr. Robles, you know very well that when temporary employment is reduced, part-time work increases. That is a logical and natural **fact**.

In this line of thought, another prominent exemplar cluster in *hecho* pronominal constructions consists of the standalone construction *esto es un hecho*, without any modifiers, as example 25 shows. In this construction, any characterization or evaluation is confined to the semantic value of *hecho*, which is generally minimal but takes on a strong sense of certainty or marked value of facticity in these contexts, thereby suggesting that the issue at hand is far beyond dispute. Because «the frequent co-occurrence of an inference with a particular construction can lead to the inference being taken as part of the meaning of the construction» (Bybee, 2013: 56), this pragmatic value can be considered an idiosyncratic contribution of the standalone construction, which is similar to that pointed out in the literature for the constructions

'*el hecho de...*' (Rodríguez-Espiñeira, 2015) and 'the fact that' (Gray & Biber, 2014; Jiang & Hyland, 2015).

(25) La acuicultura, como proceso de producción primaria, es una actividad que genera subproductos animales no destinados a consumo humano (SANDACH) y otros residuos. Esto es un **hecho** (ipacuicultura.com).

(25') Aquaculture, as a primary production process, is an activity that generates animal by-products not intended for human consumption (SANDACH) and other waste. This is a **fact**.

## CONCLUSIONS

The semantic similarities between *caso* and *hecho* have led to their classification in the same groups as factual and neutral SNs. However, the results of this study confirm that these semantic similarities do not transfer to their constructional and discourse behavior (Schmid, 2000; Hunston & Francis, 2000; Yamasaki, 2008; Gries, 2019). Furthermore, the results also confirm that this divergent discourse behavior is not dependent on the SNs but instead is mediated by the constructions in which they are embedded (Flowerdew & Forest, 2015).

*Caso* suggests a more vague and inferential referential retrieval than *hecho*, and it exhibits a stronger attraction to SN constructions with a discursive scope (determinative and pronominal). Concerning the determinative constructions, its attraction to the prepositional variant *en este caso* stands out due to the fact that this construction exhibits two distinct discursive behaviors. The more frequent or central one aligns with that of a discourse marker and serves to signal coherence relations of Condition. The less frequent or peripheral behavior aligns with that of a specificity operator and is used for the reactivation of a discursive topic. Regarding the pronominal construction, the equative copulatives *este es el caso de* and *este no es el caso de* are the prevailing ones, and they signal Background and Antithesis relations. These constructions have acquired discursive and pragmatic values, functioning as rhetorical devices for thematic introduction.

*Hecho*, on the other hand, shows a greater inclination toward the verbal construction, which is much less prominent for *caso*. The verbs that fill the slot in this construction significantly determine the coherence relation being established, mainly those relations associated with argumentation: Evaluation, Interpretation, Cause, and Evidence. As for the pronominal construction, its presence in the characterizing construction *este es un hecho* stands out, establishing Evaluation relations and appearing in two distinct variants. The more frequent or central variant is often complemented by adjectives that emphasize the factuality of the shell content. The less frequent variant is the standalone construction *esto es un hecho*, lacking modifiers. This latter

construction has acquired a discursive and pragmatic value, wherein the shell content is presented as an indisputable fact.

All things considered, despite their semantic proximity, *hecho* and *caso* exhibit distinct constructional profiles and divergent discursive behaviors, which depend to a greater extent on the constructions in which they are embedded rather than solely on the SNs. Considering this, future avenues of research on SNs should involve studies focused on specific SNs, taking into account the constructions they appear in and giving greater emphasis to emerging pragmatic and discursive values and to the other elements of these constructions, such as the verbs they include, in line with the research initiated in Duque (2020).

## ACKNOWLEDGMENTS

This work was supported by the Ministerio de Ciencia e Innovación (Spain) [project PID2021-123617NB-C43: Variación gramatical del español: microparámetros en las interfaces de la sintaxis con los niveles morfológico-léxico y semántico-discursivo].

## REFERENCES

- Aktas, R. N. & Cortés, V. (2008). Shell Nouns as Cohesive Devices in Published and ESL Student Writing. *Journal of English for Academic Purposes*, 7(1), 3-14.
- Bednarek, M. (2008). Semantic Preference and Semantic Prosody Re-Examined. *Corpus Linguistics and Linguistic Theory*, 4(2), 119-39.
- Benítez-Castro, M. A. (2015). Coming to Grips with Shell-Nounhood: A Critical Review of Insights into the Meaning, Function and Form of Shell-Noun Phrases. *Australian Journal of Linguistics*, 35(2), 168-194.
- Borreguero, M. (2006). Naturaleza y función de los encapsuladores en los textos informativamente densos. *Cuadernos de Filología Italiana*, 13, 73-95.
- Borreguero, M. (2018). Los encapsuladores anafóricos: Una propuesta de clasificación. *Caplletra*, 64, 179-203.
- Borreguero, M. & Octavio de Toledo, Á. (2007). Presencia y función de los encapsuladores en las crónicas periodísticas del s. XVII. *Philologia hispalensis*, 21, 119-153.
- Bybee, Joan (2013). Usage-Based Theory and Exemplar Representations of Constructions. In Thomas Hoffmann & Graeme Trousdale (Eds.), *The Oxford Handbook of Construction Grammar* (pp. 49–69). Oxford: Oxford University Press.

- Conte, M.-E. (1996). Anaphoric Encapsulation. *Belgian Journal of linguistics*, 10(1), 1-10.
- Cornish, F. (2022). Text, Discourse, Context: A Meta-Trilogy for Discourse Analysis. *Journal of Pragmatics*, 199, 91-104.
- Dam, L. (2014). The Interpretation of Encapsulating Anaphors in Spanish and their Functions. *Folia Linguistica*, 48(1), 37-60.
- Das, D. & Taboada, M. (2018). Signalling of Coherence Relations in Discourse, beyond Discourse Markers. *Discourse Processes*, 55(8), 743-770.
- Diessel, H. (2019). Usage-Based Construction Grammar. In E. Dąbrowska & D. Divjak (Eds.), *Cognitive Linguistics: A Survey of Linguistic Subfields* (pp. 50-80). Berlin: De Gruyter Mouton.
- Duque, E. (2014). Signaling Causal Coherence Relations. *Discourse Studies*, 16(1), 25-46.
- Duque, E. (2016). *Las relaciones de discurso*. Madrid: Arco.
- Duque, E. (2020). Neuter Pronoun ello and Discourse Verbs in Spanish. *Journal of Pragmatics* 155, 273-285.
- Duque, E. (2022). Gramática y estudios del discurso. In C. López-Ferrero, I. Carranza & T. van Dijk (Eds.), *Estudios del Discurso/ The Routledge Handbook of Spanish Language Discourse Studies*. Routledge (pp. 143-156). London: Routledge.
- Flowerdew, J. (2003). Signalling Nouns in Discourse. *English for Specific Purposes*, 22(4), 329-346.
- Flowerdew, J. & Forest, R. (2015). *Signalling Nouns in English: A Corpus-Based Discourse Approach*. Cambridge: Cambridge University Press.
- Francis, G. (1994). Labelling Discourse: An Aspect for Nominal-Group Lexical Cohesion. In Malcolm Coulthard (Ed.), *Advances in Written Text Analysis* (pp. 83-101). London: Routledge.
- Fuentes-Rodríguez, C. (1996). El lexema 'caso' y su rendimiento en el ámbito de la conexión. *Pragmalingüística*, 3(4), 329-349.
- Fuentes-Rodríguez, C. & Pérez-Béjar, V. (2022). Introducción: Macrosintaxis del discurso persuasivo. *Boletín de filología*, 57(2), 11-18.
- García-Pérez, J. (2022). Relaciones macroestructurales y sus fines persuasivos: el caso de las anáforas con valor proposicional. *Boletín de filología*, 57(2), 199-223.
- Garrido, J. & Rodríguez-Ramalle, T. (2015). Constituyentes y relaciones en la oración y en el discurso. *Círculo de Lingüística Aplicada a la Comunicación*, 62, 199-225.

- Gray, B. (2010). On the Use of Demonstrative Pronouns and Determiners as Cohesive Devices. *Journal of English for Academic Purposes*, 9(3), 167-183.
- Gray, B. & Biber, D. (2014). Stance Markers. In K. Aijmer & C. Rühlemann (Eds.), *Corpus Pragmatics: A Handbook* (pp. 219–248). Cambridge: Cambridge University Press.
- Goldberg, A. (2006). *Constructions at Work: The Nature of Generalization in Language*. Oxford: Oxford University Press.
- González-Ruiz, 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(3), 749-788.
- Gries, S. (2019). 15 Years of Collostructions. *International Journal of Corpus Linguistics*, 24(3), 385-412.
- Hoek, J., Zufferey, S., Evers-Vermeul, J. & Sanders, T. (2018). The Linguistic Marking of Coherence Relations. *Pragmatics & Cognition*, 25(2), 276-309.
- Hoey, M. (1993). A Common Signal in Discourse: How the Word Reason is Used in Texts. In J. Sinclair, M. Hoey & G. Fox (Eds.), *Techniques of Description: Spoken and Written Discourse* (pp. 67-82). London: Routledge.
- Hoffman, T. (2022). *Construction Grammar: The Structure of English*. Cambridge: Cambridge University Press.
- Hyland, K. (2017). Metadiscourse: What is it and Where is it Going? *Journal of Pragmatics*, 113, 16-29.
- Hunston, S. & Francis, G. (2000). *Pattern Grammar: A Corpus Driven Approach to Lexical Grammar of English*. New York: John Benjamins.
- Ibáñez, R., Moncada, F., Cárcamo, B. & Marín, V. (2020). Signaling of Causal Relations in Spanish. *Dialogue & Discourse*, 11(1), 40-61.
- Iruskieta, M., Da Cunha, I. & Taboada, M. (2015). A Qualitative Comparison Method for Rhetorical Structures: Identifying Different Discourse Structures in Multilingual Corpora. *Language Resources and Evaluation*, 49, 263-309.
- Izquierdo Alegría, D. & González Ruiz, R. (2013). Encapsulación y etiquetas discursivas en el discurso parlamentario: Función argumentativa a partir de un corpus paralelo. *Oralia*, 16, 185-219.

- Jakubiček, M., Kilgarriff, A., Kovář, V., Rychlý, P. & Suchomel, V. (2013). The TenTen corpus family. In *7th International Corpus Linguistics Conference CL*, 125-127.
- Jiang, F. & Hyland, K. (2015). The Fact that: Stance Nouns in Disciplinary Writing. *Discourse Studies*, 17(5), 529-550.
- Kehler, A. (2019). Cohesion and Coherence. In P. Portner, C. Maienborn & K. von Stechow (Eds.), *Semantics-Sentence and Information Structure* (pp. 450-480). Berlin: De Gruyter.
- Kilgarriff, A. & Renau, I. (2013). EsTenTen, a Vast Web Corpus of Peninsular and American Spanish. *Procedia-Social and Behavioral Sciences*, 95, 12-19.
- Langacker, R. (2000). *Grammar and Conceptualization*. Berlin: Walter de Gruyter.
- Llamas, C. (2010). Interpretación del discurso ajeno: La anáfora conceptual metafórica en la noticia periodística. *Revista de Investigación Lingüística*, 13, 107-126.
- López-Mora, P. & García-Aguiar, L. (2020). La anáfora reiterativa y reasuntiva en documentación notarial malagueña. *Estudios de Lingüística del Español*, 42, 373-396.
- 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, Revista de Filología Hispánica*, 31(2), 435-462.
- López Samaniego, A. (2018). La encapsulación nominal en el discurso académico-científico oral y escrito: Patrones de aparición. *Caplletra*, 64, 129-152.
- Martin, J. R. (1992). *English Text: System and Structure*. Amsterdam: John Benjamins.
- Mann, W. & Thompson, S. (1988). Rhetorical Structure Theory: Toward a Functional Theory of Text Organization. *Text*, 8(3), 243-281.
- Mancera-Rueda, A. (2019). Las relaciones de discurso en la comunicación política en las redes sociales. In N. Pelliser & J. Oleaque (Eds.), *Mutaciones discursivas en el siglo XXI* (pp. 57-74). Valencia: Tirant Humanidades.
- Montolío, E. (2013). Construcciones conectivas que encapsulan: [A pesar de + SN] y la escritura experta. *Cuadernos Aispi*, 2, 115-131.
- Muñoz, V. & Ciapuscio, G. (2019). Los nombres rotuladores: Un estudio de los rótulos cohesivos en artículos de investigación en inglés y español. *Revista Signos. Estudios de Lingüística*, 52(100), 688-714.



- Parodi, G., Julio, C., Nadal, L., Cruz, A. & Burdiles, G. (2019). Stepping Back to Look Ahead: Neuter Encapsulation and Referent Extension in Counter-Argumentative and Causal Relations in Spanish. *Language and Cognition*, 11(3), 431-454.
- Real Academia Española (RAE) (2014). *Diccionario de la lengua española* (23.<sup>a</sup> ed.). Madrid: Espasa.
- Recio-Fernández, I., Loureda, Ó. & Sanders, T. (2021). Constructing Discourse: An Experimental Approach. *Revista Signos. Estudios de Lingüística*, 54(107), 1004-1025.
- Ribera, J. & Marín, M. J. (2018). Lexical Encapsulation and Evaluation in Parliamentary Debate. *ELUA*, 32, 291-315.
- Rodríguez-Espiñeira, M. (2015). El sustantivo ‘hecho’ como ejemplar de nombre encapsulador factual. In I. Álvarez, R. de Andrés, Á. Arias, F. Fernández, T. Fernández, S. Fernández, H. García, A. Martínez, A. Meilán, A. Ojea, J. San Julián & J. Villaverde (Eds.), *Studium grammaticae: homenaje al profesor José A. Martínez* (pp. 207-224). Oviedo: Ediciones de la Universidad de Oviedo.
- Ruiz de Mendoza, F. J. & Gómez-González, M. A. (2014). Constructing Discourse and Discourse Constructions. *Theory and Practice in Functional-Cognitive Space*, 295-314.
- Rysová, M. & Rysová, K. (2018). Primary and Secondary Discourse Connectives. *Journal of Pragmatics*, 130, 16-32.
- Santana, A., Spooren, W., Nieuwenhuijsen, D. & Sanders, T. (2018). Subjectivity in Spanish Discourse: Explicit and Implicit Causal Relations in Different Text Types. *Dialogue & Discourse*, 9(1), 163-191.
- Schmid, H.-J. (2000). *English Abstract Nouns as Conceptual Shells: From Corpus to Cognition*. New York: Walter de Gruyter.
- Smith, C. (2003). *Modes of Discourse: The Local Structure of Texts*. Cambridge: Cambridge University Press.
- Stubbs, M. (2012). Corpora and Texts: Lexis and Text Structure. In J. Mukherjee & H. Magnus (Eds.), *Corpus Linguistics and Variation in English* (pp. 223-230). Amsterdam: Brill.
- Taboada, M. (2009). Implicit and Explicit Coherence Relations. In J. Renkema (Ed.), *Discourse of Course* (pp. 127-140). Amsterdam: Benjamins.

- Tárnyiková, J.(2018). Constructions with Shell Nouns in English. *Caplletra*, 64, 205-225.
- Tordera-Yllescas, J. C. (2021). La evolución de las relaciones retóricas y de la anáfora en niños hispanohablantes. *Lengua y Habla*, 25, 268-294.
- Urizar-Ocampo, C. & Rodríguez-Sánchez, I. (2022). Encapsuladores nominales hecho, idea, tema y problema en un corpus de escritura académica en español. *Estudios de Lingüística Aplicada*, 74,153-191.
- Vergaro, C. & Schmid, H.-J. (2017). Do the Meanings of Abstract Nouns Correlate with the Meanings of their Complementation Patterns? A Case Study on English Commissive Shell Nouns. *Pragmatics & Cognition*, 24(1), 91-118.
- Webber, B., Prasad, R., Lee, A. & Joshi, A. (2019). *The Penn Discourse Treebank 3.0 Annotation Manual*. Philadelphia: University of Pennsylvania.
- Yamasaki, N. (2008). Collocations and Colligations Associated with Discourse Functions of Unspecific Anaphoric Nouns. *International Journal of Corpus Linguistics*, 13(1), 75-98.