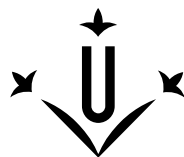
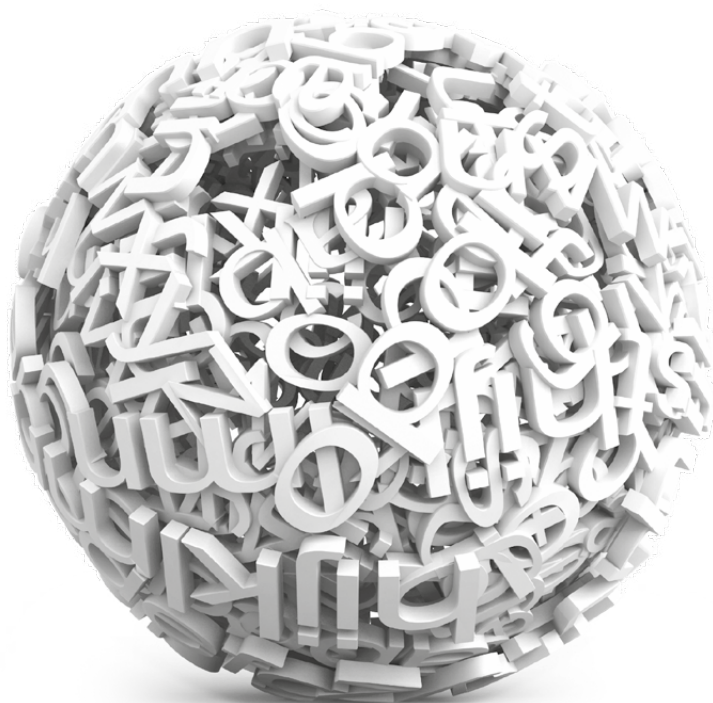


{ sintagma  
{ 37 (2025)



Universitat de Lleida



# Sintagma

**Sintagma**  
Revista de Lingüística / Journal of Linguistics

**COMITÈ EDITORIAL / EDITORIAL BOARD**

**Directora / Editor**

Àngels Llanes

**Sotsdirector / Assistant Editor**

Carlos Ángel Rizos Jiménez

**Editora / Editor**

Patricia Sáez Garcerán

**Secretària administrativa / Administrative assistant**

Olga Rovira

**Consell de Redacció / Executive Committee**

Maria Bargalló - Universitat Rovira i Virgili  
Ana Fernández - Universitat Autònoma de Barcelona  
Enric Llorca - Universitat de Lleida  
Rosa M. Mateu - Universitat de Lleida  
Raquel Serrano - Universitat de Barcelona  
Glòria Vázquez García - Universitat de Lleida  
M<sup>a</sup> Ángeles Calero - Universitat de Lleida  
Patrick Zabalbeascoa - Universitat Pompeu Fabra

**Consell Científic / Scientific Committee**

Beatriz Gallardo - Universitat de València  
Luke Plonsky - Northern Arizona University  
Tracey M. Derwing - University of Alberta  
Christiane Fellbaum - Princeton University  
Peter Garrett - Cardiff University  
Salvador Gutiérrez - Universidad de León  
Juan Carlos Moreno Cabrera - Universidad Autónoma de Madrid  
Catalina Fuentes - Universidad de Sevilla  
Elisa Rosado - Universitat de Barcelona  
Laura Domínguez - University of Southampton  
Victoria Vázquez - Universidade de Santiago de Compostela

# Sintagma

Revista de Lingüística  
Journal of Linguistics

Volum / Volume 37  
2025

Universitat de Lleida  
Lleida, 2025

Departament d'Anglès i Lingüística  
Departament de Filologia Clàssica, Francesa i Hispànica  
Universitat de Lleida  
Plaça Víctor Siurana, 1  
E-25003 Lleida

Maquetació: Edicions i Publicacions de la UdL  
Disseny de portada: Tàngram 3 Disseny Estratègic, SL  
Imatge de portada: ©iStockphoto.com/askhamdesign

ISSN 2013-6455



Creative Commons by 3.0

# Methodological issues in adapting and applying a lexical error taxonomy

**Sanja Marinov Vranješ** - University of Split,  
Faculty of Economics, Business and Tourism  
smarinov@efst.hr

**Višnja Kabalin Borenić** - University of Zagreb,  
Faculty of Economics and Business  
vkabalinb@efzg.hr

Rebut / Received: 10-01-24

Acceptat / Accepted: 21-06-24

## **Resum. Aspectes metodològics per a l'adaptació i aplicació d'una taxonomia d'errors lèxics.**

Tot i que l'anàlisi d'errors per si mateixa és insuficient per avaluar la producció dels aprenents de llengües, és un component essencial de la Complexitat, Correcció i Fluïdesa, un marc generalment acceptat per aquest estudi. La bibliografia revela que les taxonomies d'error utilitzades en estudis similars varien molt, fet que dificulta la comparació dels resultats. Per tal de promoure una aproximació a l'anàlisi d'errors que sigui consistent i uniforme en àrees que no només se centren en l'anàlisi d'errors o vocabulari, proposem una taxonomia amb directrius clares. L'objectiu principal d'aquest estudi és metodològic i consisteix a adaptar la taxonomia d'errors lèxics de Hemchua i Schmitt (2006) tot intentant trobar l'equilibri entre la seva comprensibilitat i la seva claredat i usabilitat. També es discutiran diversos problemes metodològics per quan s'implementi. La taxonomia adaptada s'ha aplicat a un corpus de redaccions escrites per estudiants universitaris croats aprenents d'anglès com a llengua estrangera i estudiants d'Econòmiques i Turisme, i demostra molts avantatges. Els resultats demostren que els errors semàntics són més freqüents que els errors formals, i que els tipus d'errors lèxics més freqüents són l'ús de sinònims pròxims, les combinacions de preposicions i l'elecció de paraules incorrectes.

**Paraules clau:** correcció lèxica, anàlisi d'errors lèxics, taxonomia d'errors lèxics, producció escrita.

## **Abstract. Methodological Issues in Adapting and Applying a Lexical Error Taxonomy.**

While error analysis alone is insufficient to assess language learners' production, it is an

essential component of the Complexity, Accuracy, Fluency triad (CAF), a generally accepted comprehensive framework for its study. The literature review reveals that existent lexical error taxonomies vary widely, making it difficult to compare results. In order to promote a consistent and uniform approach to lexical error analysis in areas of study that are not only or primarily concerned with error analysis or vocabulary learning we suggest a comprehensive taxonomy with clear annotation guidelines. Our main objective is methodological and consists in adapting the Hemchua and Schmitt (2006) lexical error taxonomy by striking a balance between its wide scope and its clarity and usability. We also discuss several methodological issues in its implementation. The adapted taxonomy is applied to a corpus of essays written by advanced learners of English as a foreign language, Croatian university students majoring in business and tourism, and shows numerous advantages. The corpus analysis indicates that semantic errors are more frequent than formal errors and the most frequent lexical errors are Wrong near synonyms, Preposition partners, and Wrong word choice.

**Keywords:** lexical accuracy, lexical error analysis, lexical error taxonomy, written production.

## 1. Introduction

Language production in the foreign language (FL) typically contains various errors that are considered a natural and inevitable process of language acquisition (Augustín Llach, 2011; Hemchua & Schmitt, 2006). Studying the development of FL accuracy helps determine what changes occur in this dimension over time or under a particular influence (e.g. a teaching intervention or a period of exposure to natural language acquisition), but also allows us to compare it to other developmental indicators such as complexity or fluency (Polio & Shea, 2014). Accuracy is defined as language production without deviation from the norm, i.e., without errors committed (Pallotti, 2009). However, this simple definition masks a very complex and inconsistent operationalization, as it is difficult to define the concept of error, and the error coding depends on the subjective interpretation of the rater(s) (Hoffmann, 2022; Juan-Garau, 2014).

Despite the widely acknowledged importance of FL vocabulary mastery (Hemchua & Schmitt, 2006), and notwithstanding the fact that even advanced written production of English second language (L2) users abounds in lexical errors (Lennon, 1991), the amount of research on lexical errors is still insufficient (Picot, 2017). One of the areas of research in which lexical errors are studied as part of the overall profile of an individual's lexical competence is the field of international student mobility research, where they are considered one of the indispensable components of the Complexity, Accuracy, Fluency (CAF) triad (Juan-Garau, 2014; Leonard & Shea, 2017; McManus et al., 2020; Pérez-Vidal & Barquin, 2014). The need to consider a wide range of variables in such complex studies tracking the development of language proficiency during a stay abroad prompted the authors to reduce the number of lexical error categories, which may



have oversimplified the taxonomies applied. In addition, the taxonomies applied in the studies quoted above varied and the interpretation guidelines were either non-existent or difficult to follow due to lack of clarity, making it difficult to compare results. This provided the impetus for our study and fuelled our search for a comprehensive, but not overwhelming, i.e. clear and easy-to-use lexical error taxonomy (LET).

The present study thus contributes to complex studies on language proficiency development by adapting an existing LET, removing the potential hurdles and reflecting the current state of knowledge on important lexical aspects. The main aim of this study is therefore to propose a LET that is both comprehensive, i.e. encompassing a wide range of lexical error types, and clear and easy to use, i.e. without overlapping categories or overwhelming details, and accompanied by transparent guidelines for coding and interpreting the results. Its application to a corpus of advanced written production will help to extend current knowledge about the frequency of different types of lexical errors.

The literature review first outlines the importance of lexical error analysis in the study of written FL production and the study of lexical competence, and then discusses the features of a number of LETs and the challenges of developing and implementing them.

## **2. Literature review**

Lexical errors are important predictors of the quality of written production and authoritative indicators of lexical knowledge (Augustín Llach, 2011). Among a number of measures of lexical competence, they have been shown to have the strongest predictive power for the holistic assessment of written production making them an indispensable component in the overall assessment of lexical competence. For example, Pavičić Takač and Buljan (2017) investigated which of the indices of lexical competence (lexical diversity, lexical density, lexical sophistication, and lexical accuracy) was the best predictor of the holistic assessment of overall written production and the holistic assessment of lexical production. The regression revealed that lexical accuracy had the greatest predictive power, explaining 27.7 % of the variance in the model (out of a total of 70.8 %).

Several problems make lexical errors less susceptible to analysis. Due to the heterogeneity of lexical errors, researchers have developed numerous taxonomies, often tailored to the specific aims of their research, which makes it difficult to compare studies. Furthermore, the literature on the actual processes involved in developing and using an error taxonomy is sparse (Hoffmann, 2022) and the lack of transparency of the annotation process in studies of learner errors is an obstacle to replicability (Lüdeling & Hirshmann, 2015).

Taxonomies of lexical errors vary according to the criteria used by their authors and can range from simple to complex taxonomies (Augustín Llach, 2011). Dissatisfied with the existing simplistic taxonomies, Zimmerman (1986) developed a highly detailed categorization that is reflected in later attempts to find an ideal taxonomy of lexical errors

(e.g., word formation, paraphrasing). His research confirmed the overwhelming presence of errors of semantic nature. Despite the valuable contribution this taxonomy has made, some categories remain questionable or overlap with others. For example, the category of incorrect co-hyponyms is problematic because the example given (*exterminate* instead of *eradicate*) could also be subsumed under the category of wrong near synonyms. The same applies to the example listed under the category of connotations (I *jerked* upstairs instead of I *jumped* upstairs), which makes connotations a subcategory of wrong near synonyms.

Engber (1995) also offered a detailed taxonomy in which she started from two basic categories: lexical choice and lexical form. This taxonomy also contains some irregularities, particularly at the level of implementation. For example, *come back* is treated as a single lexical item even though it is a phrasal verb. This taxonomy certainly influenced later taxonomies, e.g., by introducing the distinction between form and content (word semantics). Both Zimmerman (1986) and Engber (1995) observed errors at the level of individual words as well as at the level of word combinations, which has remained a standard in LETs.

A major problem is the lack of agreement on where the boundary between lexical and grammatical errors lies (Hemchua & Schmitt, 2006; Hoffmann, 2022). To illustrate this we refer to Hoffmann's (2022) overview of the lexical component of three error annotation schemes developed for the analysis of three well-known English learner corpora: International Corpus of Learner English (ICLE; Granger et al., 2009), Cambridge Learner Corpus (CLC; Nicholls, 2003), and Teaching Resource Extraction from an Annotated Corpus of Learner English (TREACLE; MacDonald et al., 2011). The author pays particular attention to how lexico-grammatical categories are treated in the selected schemes. Namely, while certain errors were traditionally considered only grammatical in nature, some LETs acknowledge the fact that knowledge of a word also includes an awareness of the grammatical rules associated with the word, e.g. which prepositions are found alongside it, or which argument structure a verb necessitates. For instance, ICLE has a separate category for lexico-grammatical features (i.e., it counts noun-, verb-, or adjective-dependent prepositions as errors) and CLC also considers certain lexico-grammatical features, but places them under the category 'additional' (i.e., incorrect verb inflections, errors in tense, and violations of a verb's required argument structure, e.g., \*told that instead of told me that). Interestingly, the youngest of the three taxonomies, TREACLE, does not provide for a lexico-grammatical category, but treats morphological and prepositional errors as grammatical.

### ***1.1. Examples of approach to lexical error analysis in complex language development studies***

In the following paragraphs, we refer to a number of studies to illustrate the inconsistencies in the use of LETs, the lack of transparency in annotation and the resulting problems in comparing results and replicability. It should be noted that all the studies

mentioned are high-quality, complex, peer-reviewed studies of language development in oral or written production. The possible reason for the simplified treatment of LETs lies in the array of variables to be considered.

We start with the examples where the lexical taxonomy used is not explained in the published work, which does not necessarily mean that the authors were not rigorous in their coding. McManus et al. (2020) investigated accuracy in the context of the development of oral production of French and Spanish, where lexical errors were defined simply as —selection of an appropriate lexical item— (p. 15). The methodologically well-designed and statistically well-supported study by Köylü & Tracy-Ventura (2022) also contains no information on the taxonomy used. However, the authors recognise the need for a detailed analysis of grammatical, lexical and pragmatic errors, which would paint a more nuanced picture of the development of accuracy.

Juan-Garau (2014) investigated lexical accuracy as one of the components of overall oral accuracy. The author also considered grammatical and pragmatic accuracy, and the taxonomy of lexical errors was simplified to only two categories: lexical choice and collocations. Llanes and Muñoz (2009) applied the classification of errors proposed by Bardovi-Harlig and Bofman (1989), where lexical errors are identified as incorrect idiomatic expressions or words. The proposed categorization also includes the so-called covered errors, which are explained as those that lack precision (*the mother* instead of *their mother*) or the word or expression that is contextually incorrect or a lexical approximation (*bag* instead of *basket*). Judging by the examples, covered errors seem to be a subcategory of lexical errors rather than a category of their own. This category is reminiscent of how the coding of lexical errors of a Spanish corpus was carried out in Leonard and Shea (2017), where an imprecise but not completely incorrect lexical choice was not considered an error (e.g. the use of wrong near-synonyms as in the example above: *bag* instead of *basket*). An influential project in the field of language development research during study abroad - SALA (Beattie, 2014) proposed a simple lexical taxonomy with five categories (Zaytseva, 2016): wrong lexical choice, non-existent words, L1 transfer, phrasal verbs, and fixed expressions. This taxonomy was consistently applied in a number of studies that emerged from this project (Borràs & Llanes, 2021; Pérez-Vidal & Barquin, 2014; Zaytseva et al., 2019) and ensured comparability between these studies. The same/similar approach was used in the analysis of a Spanish corpus by Borràs et al. (2023), who report using a taxonomy with four categories, including: L1 transfer, word choice, non-words and fixed expressions. By adding fixed expressions, the authors clearly indicate the need to account for errors related to multiple expressions and word partnerships. However, in some studies, such as Zaytseva et al. (2019), only three categories are used (wrong lexical choice, non-existent words, L1 transfer), although the authors claim to use the original SALA taxonomy. The possible misuse of word partnerships seems to be often treated as L1 transfer errors (e.g. confusion between verb pairs *make/do* or phrasal verbs). It seems that the way in which L1 transfer errors should be treated in lexical

error research deserves greater attention. This category will be discussed in more detail later in the paper.

Following Picot (2017), in this study we selected Hemchua and Schmitt's (2006) taxonomy (figure 1) since it is influential, widely used and detailed.

<b>A FORMAL ERRORS</b>	
<b>1 Formal misselection</b>	
i	Suffix type: <i>competition/competitiveness</i>
ii	Prefix type: <i>reserve/preserve</i>
iii	Vowel based type: <i>seat/set</i>
iv	Consonant-based type: <i>save/safe</i>
v	False friends: <i>bank/banknote (in Thai)</i>
<b>2 Misformations</b>	
i	Borrowing (L1 words) <i>kopf (instead of head, from German)</i>
ii	Coinage (inventing based on L1): <i>nocive (instead of harmful, from Portuguese)</i>
iii	Calque: <i>to bring us go to the hospital (instead of bring us to the hospital)</i>
<b>3 Distortions</b>	
i	Omission: <i>intresting</i>
ii	Overinclusion: <i>dinning room</i>
iii	Misselection: <i>delitouse</i>
iv	Misordering: <i>littel</i>
v	Blending: <i>travell (travel + travelled)</i>
<b>B SEMANTIC ERRORS</b>	
<b>1 Confusion of sense relations</b>	
i	General term for specific one: <i>equipment &lt;appliance&gt;</i>
ii	Overly specific term: <i>colonels &lt;officers&gt;</i>
iii	Inappropriate co-hyponyms: <i>communication &lt;transportation/public transport&gt;</i>
iv	Near synonyms: <i>regretful &lt;penitent&gt; criminal</i>
<b>2 Collocation errors</b>	
i	Semantic word selection: <i>The city is grown &lt;developed&gt;</i>
ii	Statistically weighted preferences: <i>An army has suffered big losses &lt;heavy losses is preferred&gt;</i>
iii	Arbitrary combinations: <i>hike-hitch &lt;hitch-hike&gt;</i>
iv	Preposition partners: <i>They will travel by foot.</i>
<b>3 Connotation errors</b>	
<b>4 Stylistic errors</b>	
i	Verbosity: <i>I informed my girlfriend of the party through the medium of telephone.</i>
ii	Underspecification: <i>Although cars in the country are lower.</i>

FIGURE 1. HEMCHUA AND SCHMITT'S LEXICAL ERROR TAXONOMY (2006, P. 12)

### ***1.2. Explanation of Hemchua and Shmitt's (2006) lexical error taxonomy***

The taxonomy consists of two basic categories of lexical errors: formal and semantic (Engber, 1995). Formal errors result from the similarity of form between the incorrect word

and the required, correct one. It is subdivided into three subcategories. *Formal misselection* refers to errors resulting from the formal similarity of two words (they are spelled similarly or sound similar), both of which occur in the target language, but in the given context the other word would be appropriate. *Misformations* include words that do not occur at all in the target language and *Distortions* are minor spelling errors. Each of the three subcategories is further divided into their subcategories to capture the nuances of the error type.

The category of semantic errors contains meaning-related errors and it consists of four subcategories. The first concerns *Confusion of sense relations*, which indicates that the user has not fully mastered hyperonymic relations or makes an incorrect choice of synonymous words and the second is intended for *Collocation* errors. These two subcategories are further subdivided into their subcategories, which attempt to refine the type of sense relation or collocation error committed. *Connotation errors* are those based on the incorrect use of a word due to its wrong connotative meaning and have no further subcategories. *Stylistic errors* mainly refer to cases where it is not possible to point out a single incorrect lexeme. Rather, the entire phrase sounds inappropriate due to a series of poor lexical choices. They are further divided into *Verbosity*, which refers to the use of a compensatory paraphrasing strategy in the absence of a more precise lexical item, and *Underspecification*, which is the result of an unclear meaning or a statement that is too brief or vague/incomplete.

This LET was first applied to the analysis of the written language production of Thai third-year English majors (N = 20) with about ten years of experience in learning English. Their essays were about the advantages of rural and urban life. The results showed that (i) the most common types of errors are near-synonyms, prepositional partners, and suffix use, (ii) word semantics is a greater source of error than form, and (iii) the source of error is more often an inherent feature of the target language lexical unit than transfer from the source language (Hemchua & Schmitt, 2006). The study was replicated by Picot (2017) on a sample of Greek students (N = 20) to investigate the applicability, reliability, and validity of the taxonomy. The students were enrolled in an IELTS preparation course at a private language school and had about eight years of experience learning English. Their essay was about the government's right to limit the number of children in a family. Both groups had 1.5 hours to write an argumentative essay of 300-350 words under exam conditions. The two studies found statistically significant agreement in the number, distribution, and frequency of lexical errors, i.e., it was confirmed that learners of English as a foreign language with different source languages have similar problems in achieving lexical accuracy.

### ***1.3. Main issues in applying the Hemchua and Schmitt (2006) lexical error taxonomy***

In addition to confirming the validity of the taxonomy in two L1 contexts, we believe that the main value of the two studies is that they highlight problem areas in

the implementation of the taxonomy that lend themselves to improvement. Below we highlight those confirmed by Picot's (2017) findings.

- *Calque errors* (A2 iii) were almost double in the Picot (2017) study, compared to Hemchua and Schmitt (2006). The author believes this is due to the decision to include all the transfer errors into this subcategory, regardless of the type of error. Accordingly, he perceives it as one of the major flaws in the original taxonomy because by assigning all language transfer errors to one of the *Misformations* (A2) subcategories we lose the information on the type of error, i.e. we confuse the linguistic type (e.g. collocation, preposition partner) and cause of error (L1 interference or complexity of L2).
- The categories of *Inappropriate co-hyponyms*, *Near synonyms*, *Semantic word selection*, and *Connotation* were difficult to code systematically due to overlapping.
- The category of *Semantic word selection* was over inflated in the second study because there were nine times more errors in this category in Picot (2017), compared to Hemchua and Schmitt (2006). This may be the results of its overlapping with some other categories (see above) but also because in Picot (2017) all fixed phrases (e.g. *What's \*a matter?\**) were considered collocational and classified as *Semantic word selection*. Likewise, the author feels it may be the result of "different categorisation procedures and subjective interpretations of the 2006 guidelines" (Picot, 2017, p. 49) regarding the differences between *semantic word selection*, *near synonyms*, and *connotations*.
- A separate category for a wrong word choice (other than wrong near synonym but existing in English) was not envisaged. As a result, these (besides collocations and fixed phrases) were classified under *Semantic word selection*.
- Some lexico-grammatical errors were treated as purely grammatical and excluded from the analysis but might need to be reconsidered (e.g. *It's not difficult \*for getting to a hospital\**). Picot (2017) suggests this should be interpreted as a lexical error since it is the word *difficult* that requires a subsequent full infinitive. The name of the category is not suggested.

The present study contributes to the research field of lexical accuracy by adapting the taxonomy of Hemchua and Schmitt (2006), removing the possible obstacles and reflecting the current state of knowledge on important lexical aspects.

## 2. Methodology

### 2.1. Goal and Research Questions

The objective of this study was twofold. The first was methodological and consisted in identifying a comprehensive lexical taxonomy and adapting it according to the criticisms that arose from the previous research, but also in accordance with our own reflections and assessments. The aim was to maintain the scope of the taxonomy, i.e. to cover a wide

range of lexical error types, but to make the application clear and simple, i.e. without overlapping categories or overwhelming detail, accompanied by transparent guidelines for coding and interpreting the results. The second objective was to identify the main problem areas related to lexical accuracy in a sample of written production of Croatian university students. Thus, our research questions are as follows:

What changes can be made to Hemchua and Schmitt's (2006) taxonomy to improve it and facilitate its application in written corpora?

What are the most common lexical errors in the analysed corpus of written production of Croatian university students?

## ***2.2. Participants***

Twenty-two students from the Faculty of Economics, Business and Tourism at the University of Split, Croatia voluntarily participated in this study. They all had at least 14 years of experience learning English in formal contexts and were exposed to a high proportion of English-language media content, as there is no tradition of dubbing in Croatia. Furthermore, Croatia is a tourist country and English is the primary lingua franca for international communication. Although no official language tests were administered, it is estimated that all students were at least B2 and above, based on their high school leaving exams and the English for Specific Purposes courses they took in the 1st and 2nd years of their undergraduate programme. All respondents applied for the Erasmus+ mobility programme, where English was the expected language of instruction, i.e. all courses they enrolled into were delivered in English. English was also the lingua franca in social encounters. There were seven male and 15 female respondents, eight undergraduates and 14 graduates: one 2<sup>nd</sup> year undergraduate, seven 3<sup>rd</sup> year undergraduates, three 1<sup>st</sup> year graduates, and 11 2<sup>nd</sup> year graduates ( $M=4.091$ ;  $SD=1.019$ ). The participants were in the age bracket between 20 and 23 years ( $M=22.091$ ;  $SD=1.091$ ).

## ***2.3. Materials and procedure***

First, Hemchua and Schmitt's (2006) taxonomy was chosen because it encompasses a wide range of error types and because it was successfully replicated in Picot's (2017) study, which found it appropriate for learners with different L1s. More importantly, the replication study pointed to several problems in the original taxonomy that lent themselves to adaptation (see Literature review). The selected instrument was modified taking into account previous experience with its use (Picot, 2017) and the problems arising from conflicting interpretations of some common error types (e.g. lexical vs. grammatical). A corpus of written production was compiled to which the adapted taxonomy was applied. The essays were written in 60-minute sessions under examination conditions, on the topic of advantages and disadvantages of EU membership, a topic

related to the subject area of the participants' study programme. Mean length of the essays is 494.64 tokens (SD = 97.39, min = 323, max = 773). To further illustrate the lexical complexity of the corpus, we provide the following three measures: frequency (according to COCA reference corpus) (M=1060.336; SD=191.130), lexical density (M=0.529; SD=0.046) and lexical diversity (MTLD) (M=80.410; SD=19.647).

Three raters were involved in the lexical error analysis: an L1 English speaker (an applied linguist by profession) and the authors of the study (L2 English speakers), both ESP (English for Specific Purposes) lecturers in business and tourism. The raters were given detailed guidelines based on a review of the literature and a preliminary lexical error analysis of the essays. The raters identified and coded all lexical errors individually. In cases of disagreement, the raters discussed the individual cases (see for example 4.2.4 Differentiating between Verbosity and Underspecification or 4.2.7 Specific roles of L1 English speakers and the language for specific purposes specialists) and reached consensual decisions (Polio & Shea, 2014). Finally, the relative number of errors was counted, depending on the length of the essays, i.e. the number of tokens (Borràs & Llanes, 2021).

### 3. Results

#### 3.1. *Adapting the Hemchua and Schmitt (2006) taxonomy*

In this section we address our first research question, i.e. we explain the rationale for the changes we have made to the Hemchua and Schmitt taxonomy (2006). The types of adopted lexical errors, the codes used, explanations and examples from the corpus of essays collected for this study are shown in Figure 2.

We start from the category of *Formal lexical errors*, where we retained the three basic subcategories - *Formal misselection*, *Misformations*, and *Distortions* - and eliminated further subclassifications. This decision is supported by the evidence from previous research (Hemchua & Schmitt, 2006; Picot, 2017) which confirmed that most of these (prefix type, vowel-based type, false friend, borrowing, coinage, misselection, misordering, blending) were low frequency errors. Moreover, we assumed that the details given by the subcategories would not have an important impact on the research, but could be overwhelming for the coding.

Regarding the category of *Formal lexical errors* it should be noted here that the original taxonomy took into account both the type and the source of the error by examining cross-linguistic errors. The inadequacy of this approach was acknowledged because it loses information about the nature of the error (see Literature review). Therefore, our adapted taxonomy considers only the type of error. In terms of guidelines for coding this means that L1 transfer errors should not be automatically classified as *Misformations*, as in Picot (2017) but should be classified according to error type.



Three of the four categories of semantic errors have been retained: *Confusion of sense relations*, *Collocations* (in a different hierarchical position) and *Stylistic errors*. The subcategory *Connotations* was merged with that of *Wrong near synonyms* due to the lack of a precise definition and coding guidelines, which led to overlap with the categories *Inappropriate co-hyponyms*, *Near synonyms* and *Semantic word selection* (see Literature review), but also because the interpretation in the earlier studies was insufficient. The table with the original taxonomy does not contain an example (Figure 1), and the rare examples in the observed studies indeed indicate that an overlap with the mentioned categories is to be expected (e.g. \*You will wake up to voice's bird).

The category of *Confusion of sense relations* has been retained but three out of four of its subcategories were merged and one added. The subcategory of *Inappropriate co-hyponyms* was merged with the *Wrong near synonyms* with which it was found to overlap. *The General term for specific term* and *Overly specific term* subcategories were subsumed under the category of *Wrong near synonyms*, since these are essentially its subcategory and add extra complexity to a taxonomy we wish to simplify. The *Wrong word choice* subcategory was added to capture words that exist in English but do not fit the context and do not qualify as either a *Wrong near synonym* or a *Formal misselection*. The need for such a category was recognised in previous research (see Literature review).

The category of *Collocation errors* has been renamed *Word partnerships* to also include other types of lexical combinations, and *Collocations* have been retained as its subcategory. As shown in Literature review, this area suffers from at least two problems: (i) it is not completely clear which word partnerships are considered in particular studies and what the rules for their coding are and (ii) there is a lack of consensus related to the place of lexico-grammatical categories in LETs. Thus, this category has been substantially reshaped. We have excluded three out of the original four subcategories of *Collocations*. *Semantic word selection* has been shown as greatly problematic in previous research, either overlapping with other categories or serving as a repository for the errors for which a category was not envisaged (see Literature review). As for *Statistically weighted preferences*, multiple potential co-raters for this study agreed that it is a confusing category that would cause additional coding difficulties and that attention to such detail is not necessary. The latter also applied to *Arbitrary combinations*. These have been subsumed either under the category of *Collocations* or *Wrong near synonyms* (see also our methodological decision presented in 4.2.2 *Differentiating between Collocations and Wrong near synonyms or Wrong words*). The subcategory of *Preposition partners*, although potentially a subcategory of *Collocations*, has been kept separate because previous research has shown it to be a prominent category of lexical error (Picot, 2017). It is interesting to note that some high profile and relatively recent error taxonomies treat preposition errors only as grammatical errors (TREACLE), while others have recognised their position in the lexico-grammatical category (ICLE and CLC).

	Type of error and code	Explanation	Corpus example
FORMAL ERRORS	<b>FORMAL MISSELECTION</b> (Miss)	The source of the error is a formal similarity with another English word that would fit the given context.	... in which Slovenia wants to <b>enforce</b> Croatia to give away the bay
	<b>MISFORMATIONS</b> (Misf)	The word does not exist in English.	More attractive to foreign <b>investitors</b>
	<b>DISTORTIONS</b> (Dist)	Misspelled word.	For free or for small <b>tution</b>
SEMANTIC ERRORS	<b>CONFUSION OF SENSE RELATIONS</b>		
	Wrong near synonym (WNS)	Word of a similar meaning that does not fit the contexts.	Being part of <b>together</b> market ...
	Wrong word choice (WWC)	The word exists in English but does not fit the context. It is neither a wrong near synonym nor a formal misselection.	Implemented political ideology that is <b>innately</b> eroding our cultural values
	<b>WORD PARTNERSHIPS</b>		
	Collocation (COLL)	Wrong collocation.	They can push us into <b>conducting</b> some law that we don't agree with.
	Preposition partners (PP)	Wrong choice of preposition.	... there are disadvantages <b>in</b> Croatian membership ...
	Fixed phrases (FIX)	Erroneous usage of phrasal verbs, idioms and other types of word partnerships.	<b>On the first sight</b> , that is great.
	Pattern (PATT)	Wrong pattern related to the key word, i.e. word preceded of followed by the wrong structure.	Being a full member leads to progress, <b>to create</b> new opportunities ...
<b>STYLISTIC ERRORS</b>			
Verbosity (VERB)	Using more words than necessary (e.g., paraphrasing a specific word one lacks).	They employ people <b>and give them job and money</b> .	
Underspecification (UNDER)	The thought/idea is incomplete due to the poor choice or lack of vocabulary.	Even now a lot of people <b>have to work to collaborate with EU</b> .	

FIGURE 2. CATEGORIES AND SUBCATEGORIES OF THE ADAPTED HEMCHUA AND SCHMITT (2006) TAXONOMY WITH EXPLANATIONS AND EXAMPLES DERIVED FROM THE CROATIAN CORPUS UNDER STUDY

In addition to renaming *Collocation errors* into *Word partnerships*, and changing the hierarchical position of *Collocations* and *Preposition partners*, two new types of word partnerships have been introduced: *Fixed phrases* and *Pattern*. The *Fixed phrases* subcategory includes lexical units such as phrasal verbs, idioms, but also other fixed word partnerships such as “in conclusion”, “in hope of”, “at first sight”. These differ from collocations because the position of words in these expressions is more strongly determined and also the words together often take on a new meaning which is not simply a sum of meanings of all the included words. *Pattern* errors are ostensibly a matter of syntax, but they are really the result of limited knowledge of the lexical unit,

e.g. “\*told that” instead of “told me that”. Hoffmann (2022) refers to it as “required argument structure”. As shown in the Literature review, CLC classifies these errors in the category “additional” and Picot (2017) refers to such errors as “sentence errors” but recognises this as one of the shortcomings of the taxonomy used. Since knowledge of a word implies knowledge of its syntactic properties (Ellis, 1995), we propose to treat pattern errors as lexical.

Finally, the subcategory *Stylistic errors* seemed to be the most difficult for objective and systematic coding, so its omission was considered. However, the preliminary analysis of the essays showed that the *Verbosity* and *Underspecification* were applicable and that they resulted from a complex set of inappropriate lexical choices that could not be coded by any of the other categories listed. Therefore, we retained them, knowing that their coding would have to be negotiated more often than for other types of errors.

While the literature review and preliminary analysis of the essays allowed us to adjust the taxonomy as shown, several issues arose during the analysis that required further methodological decisions to be made. We present these in the following section.

### ***3.2. Methodological issues resulting from the course of analysis***

When the categories were established, the process of annotation still raised a number of issues on which agreement had to be reached. In this section, we explain the guidelines established for the proposed taxonomy to avoid inconsistencies in the rating process.

#### *3.2.1. Differentiating between Formal misselection and Distortion*

It is possible to encounter such errors, which can be interpreted either as a *Formal misselection* or as a *Distortion* such as the word ‘habitants’ in example 1:

- (1) Some of the *habitants* (Dist) think that membership of the Croatia in EU is a complete disaster.

The word was used instead of the correct word ‘inhabitants’. It was treated as a *Distortion* rather than a *Formal misselection* since it is highly unlikely that the user was aware of the existence of the actual word ‘habitant’ with its specific meaning and its low frequency of occurrence.

#### *3.2.2. Differentiating between Collocations and Wrong near synonyms or Wrong words*

An incorrect collocation is essentially the choice of an incorrect near synonym. Since the taxonomy includes both categories, it was important to find a way to distinguish between the two to allow for consistent coding. For instance, in example 2 the choice of ‘make’ is incorrect because it does not collocate with ‘business’ meaning that the choice of the correct verb depends directly on the noun to which it refers. It was therefore coded as a *Collocation* rather than a *Wrong near synonym*.

- (2) You can easily get some money and *make* (COLL) private *business*.

For the same reason, in example 3, the verb ‘ensure’, although a *Wrong word choice*, has been coded as an incorrect *Collocation* of the noun ‘request’.

- (3) Now there is option to work in all parts of Europe without any special *request* you have to *ensure* (COLL).

### 3.2.3. Differentiating Between Pattern and Preposition Partners

The newly introduced category of *Pattern* often involves unnecessary inclusion or omission of a preposition. This initially caused some coding confusion and required a clearer rule to distinguish between the categories of *Pattern* and *Preposition partners*. When a preposition was needed to correctly complete a pattern, and it was either included where it was not necessary (example 4), or omitted (example 5), these cases were coded as *Pattern*.

- (4) ... it *benefits to* (PATT) the local community.
- (5) ... trying to *explain people* (PATT) by different medias.

Cases in which there was no pattern involved but simply an incorrect preposition was used were coded as *Preposition partners*, as in example 6:

- (6) ... more and more people are going *in* (PP) foreign countries in search *for* (PP) work.

### 3.2.4. Differentiating between Verbosity and Underspecification

As predicted, the cases to be included in the *Verbosity* and *Underspecification* categories required more discussion. Apart from implying a lesser degree of objectivity, these two categories occasionally overlap, although this seems counterintuitive. For example, although the example below was initially coded as *Verbosity* because an inappropriate word choice resulted in an excessively long sentence segment, the opinion prevailed that the idea behind these words is not clear. Therefore, it was coded as *Underspecification* (example 7).

- (7) ... so they decided to go to other countries in European Union *without any problem about life in new country* (UNDER).

### 3.2.5. Number agreement

The analysis revealed numerous errors in number agreement (example 8).

- (8) *One of the more important benefit* (correct: benefits),

A smaller number of such errors could be considered a slip of the pen, but a considerable number were found. It was decided not to consider this a lexical error, although it could be classified under *Pattern*, since the choice of singular or plural depends on a lexical frame (one of ... benefits). The numerical agreement was marked as a lexical error only in case of fixed phrases (e.g., thousand of).

### 3.2.6. *Decisions regarding specific lexemes*

In addition to the need to make a clear distinction between certain categories, decisions had to be made regarding particular lexemes, since their coding could significantly affect the overall result. We illustrate our point with the example of a collocation directly related to the subject of the essay, namely membership in the European Union. Instead of 'accessing the EU' the participants spoke of 'entering the EU' with both the English L1 speaker and the LSP specialists agreeing that this refers to physically entering the EU (crossing the borders) and not joining. However, the referenced corpora confirm the presence of the collocation 'enter EU' in the sense of joining it. Likewise, the frequency of its occurrence in the analysed essays would unjustifiably inflate the number of collocation errors based only on this one collocation. Therefore, it was decided not to code it as an error.

### 3.2.7. *Specific roles of L1 English speakers and the language for specific purposes specialists*

Although the decision to include as raters both an English L1 speaker and the language for specific purposes specialists was made prior to the coding procedure the course of analysis proved this the right choice. The intuition of the English L1 speaker is a very valuable contribution to any such analysis. However, it has been shown that even this intuition can be refuted by evidence from large language corpora. For example, the English L1 speaker considered the word 'majority' (example 9) to be a *Wrong near synonym* while the Sketch Engine English Web 2020 (enTenTen20) corpus (Kilgarriff et al., 2014) confirmed frequent presence of this word partnership.

- (9) Croatia is mostly a tourist country, and *the majority of the GDP* depends on tourism.

In addition, it has been shown that English L2 LSP teachers are more sensitive to specific terminology and its exact meaning than the English L1 speaker (example 10).

- (10) That would have effects on both *work sector* (WWC) and population. (correct: labour market)

### 3.3. Distribution of lexical errors in the written corpus of Croatian university students

In this section, we answer our second research question about the presence of certain types of lexical errors in the analysed written production of a sample of Croatian university students. There was a total of 491 errors. The data are presented as relative to the number of tokens, i.e., the number of lexical errors per 100 tokens (Borràs & Llanes, 2021), to facilitate comparison in future studies that may wish to apply the same taxonomy but have essays of a different length (table 1).

TABLE 1. DESCRIPTIVE ANALYSIS OF TYPES OF LEXICAL ERRORS IN THE CORPUS OF ESSAYS OF CROATIAN UNIVERSITY STUDENTS, RELATIVE TO THE NUMBER OF TOKENS (THE NUMBER OF LEXICAL ERRORS PER 100 TOKENS), IN DESCENDING ORDER OF OCCURRENCE

Type of lexical error	Mean	Std. Deviation	Min	Max
Wrong near synonym (WNS)	1.19	0.60	0	2.45
Preposition partners (PP)	0.54	0.35	0	1.20
Wrong word choice (WWC)	0.52	0.28	0	1.06
Distortions (Dis)	0.51	0.54	0	1.55
Formal misselection (Miss)	0.45	0.33	0	0.99
Pattern (PATT)	0.39	0.32	0	1.09
Verbosity (VERB)	0.29	0.29	0	1.10
Fixed phrase (FIX)	0.25	0.21	0	0.62
Collocations (COLL)	0.24	0.17	0	0.67
Underspecification (UNDER)	0.23	0.25	0	0.73
Misformations (Misf)	0.01	0.05	0	0.21

Table 1 presents the error types in descending order of occurrence. The most frequent errors are three categories of semantic errors, with *Wrong near synonyms* in the first place and more than twice as frequent as the next two, i.e., *Preposition partners* and *Wrong word choice*. Two categories of formal errors, *Distortions* and *Formal misselection*, follow in 4th and 5th place. The third category of formal errors (*Misformations*) comes last in the list and is practically absent.

In 6th place on the list is the second most frequent category of word-partnership errors, *Patterns* (the most frequent being *Preposition partners*). Errors related to the two remaining subcategories of word-partnership errors, *Fixed phrases* and *Collocations*, are slightly less frequent, ranking 8th and 9th, respectively.

Judging from the mean values, the two types of stylistic errors are relatively equally frequent, with *Verbosity* being slightly more frequent than *Underspecification*.

#### 4. Discussion

The aim of this study was to adapt an otherwise comprehensive and empirically validated taxonomy of lexical errors. We reduced the 24-item taxonomy of Hemchua and Schmitt (2006), tested by Picot (2017), to 11 items to ensure clarity and ease of use, as suggested by Hoffmann (2022) and Lüdeling & Hirshmann (2015), without compromising its scope. The comprehensiveness of the new taxonomy is reflected in the sufficiently elaborate framework, which retains the primary division into formal and semantic errors (Engber, 1995) as well as most of the main categories. However, a number of categories/subcategories were excluded, merged or modified to eliminate overlapping, imprecisely defined and less useful ones. All subcategories of the three main categories of formal errors were excluded due to the low frequency of eight of the 13 such categories (Picot, 2017), which would unnecessarily overburden raters and offer no significant practical benefit to complex language proficiency development studies that recognize the need for lexical error analysis. Our corpus analysis revealed a total of 99 formal errors out of a total of 491 lexical errors. This result confirms previous findings on the prevalence of semantic errors (Hemchua & Schmitt, 2006; Picot, 2017; Zimmerman, 1986) and at the same time supports our decision not to further subdivide the three main categories of formal errors. We also chose not to include reference to the source of error since this was shown to hide the nature of error (Hoffmann, 2022) and inflate the category of *Misformations* (Picot, 2017). For example, in the following example from our corpus there is an evident L1 transfer from Croatian in the use of preposition - “using the advantages on\* the best way possible” - but if it were classified as a *Misformation* we would lose the information on its type (*Preposition partners*). This kind of methodological flaw is not uncommon in lexical error coding (Borràs et al., 2023; Pérez-Vidal & Barquin, 2014; Zaytseva, 2016) and should be avoided. This does not ban the study of cross-linguistic errors from lexical error research but suggests that the type and source of error are better not captured by the same taxonomy (Hoffmann, 2022).

A notable example of merging took place within the category of *Confusion of sense relations* where *Inappropriate co-hyponyms* were found to be overlapping with *Wrong near synonyms* (Picot, 2017). Furthermore, *General term for specific one* and *Overly specific term* were merged under *Wrong near synonyms* to further simplify the taxonomy. Even if they were to be retained in the taxonomy, they should be presented at a different hierarchical level, as a subcategory of *Wrong near synonyms* (e.g. “So in most cases \*payments are bigger than in Croatia” (correct: salary); “population’s \*salary stayed very low” (correct: income)). Nonetheless, in a study focusing exclusively on near synonyms these would be worth considering in terms of qualitative analysis that would identify specific lexemes

that occur and therefore require special attention in language instruction. In that case we would also suggest adding an extra subcategory - *Wrong style near synonyms* - to capture the use of a synonym which belongs to a different style or register (e.g. “no matter how crazy that sounds\*” or “get the money\* from European organisations” in formal academic writing).

Another subcategory from the original taxonomy, namely *Semantic word selection*, originally a subcategory of *Collocations*, was also merged with *Wrong near synonyms*, since it proved to overlap with *Inappropriate co-hyponyms*, *Wrong near synonyms* and *Connotations* and was also used for wrong word choice for which there was no designated category (Picot, 2017). The range of overlap between the subcategories of *Confusion of sense relations* and the subcategories of *Collocations* required a specific methodological decision, as described in 4.2.2 *Differentiating between Collocations and Wrong near synonyms or Wrong words*. These clear guidelines enabled our raters to easily distinguish between the three categories and to code the error in “plethora of *opportunities* that were previously either hard to *acquire\** (COLL)” as a *Collocation error*; and the two errors in “Advantages that Croatia has are *ambitious\** (WWC) but there seem to be problems about using them *correctly\** (WNS)” as *Wrong word choice* and *Wrong near synonym*, respectively.

As an example of a substantial modification we highlight the category of *Collocations*, which has been reconceptualised as *Word partnerships*. By doing so we are observing the now common realisation in vocabulary teaching (Lewis, 1993, 1997) that there is a whole range of such partnerships that should be taught as lexis. It follows that they should also be part of LETs. As subcategories of *Word partnerships* our taxonomy recognizes *Collocations*, *Preposition partners*, *Fixed phrases* and *Pattern*. By retaining the category of *Preposition partners* and distinguishing them from other collocations and by adding the category of *Pattern* we show our compliance with the lexical approach (Lewis, 1993, 1997), i.e. our agreement with the idea that while these are traditionally considered to belong to the domain of grammar (MacDonald et al., 2011), they also reflect complete or incomplete knowledge of a word and therefore belong in the taxonomy of lexical errors.

No changes have been made to the third category of semantic errors - *Stylistic errors* - as both subcategories have been retained. It should be noted, however, that despite their obvious connection with vocabulary choice, the two categories also seem to be influenced by competences other than lexical ones. In addition to the lack of a precise term, the categories of *Verbosity* and *Underspecification* were found to be indirectly influenced by the lack of ability to develop an idea, which is then reflected in word choice. It seems that achieving objectivity in the coding of *Verbosity* and *Underspecification* will remain a problem because it cannot guarantee the precision of the analysis and because it is time-consuming and tedious for raters. Therefore, future studies might consider them ambiguous and decide against their inclusion in the lexical taxonomy.

The resulting taxonomy was used to code a corpus of EFL written production (essays) by Croatian university students. In line with the findings of previous research on lexical



accuracy (Engber, 1995; Hemchua & Schmitt, 2006; Picot, 2017), our analysis shows that the most striking lexical errors are semantic in nature, two of which are related to the confusion of sense relations (*Wrong near synonyms* and *Wrong word choice*) and one from the field of word partnerships (*Preposition partners*). As for the confusion of sense relations, this suggests that advanced users have access to a large vocabulary but are still unable to apply it in the correct contexts. This could be because their focus is on fluency and achieving a communicative goal rather than linguistic accuracy, or because our participants tend to use English in less demanding subject areas than that required in the given essay. The pedagogical implication of this finding is that more focus on form is needed to improve awareness of semantic similarity/difference. It is the actual errors made that can be used as a starting point for targeted vocabulary training. Students could be encouraged to look for a better alternative and to support their findings with further examples.

*Word partnerships* are represented at the top of the list of errors by *Preposition partners*. This is not surprising, since prepositions are a very common part of speech in all corpora, making them very frequent, and there is rarely a one-to-one translation. As explained in the methodology section, prepositions are also closely related to the *Pattern* category, which also ranks high. It is interesting to note that *Fixed phrases* and *Collocation errors* are less frequent than *Preposition partner* and *Pattern* errors. This might suggest that it is easier to avoid using a fixed phrase or a good collocation (e.g. by using a paraphrase) when the user is unsure about their use, while they cannot avoid using a preposition or preceding/following a word with a pattern to express an idea. Furthermore, this result confirms the well-established position in language teaching that both prepositions and patterns are part of vocabulary knowledge and should be included more frequently in vocabulary instruction (Lewis, 1993, 1997). This also reinforces our position that prepositions and patterns should be recognised as valid components of LETs, a position not yet fully accepted by all.

Although correct spelling is important, we believe that of the two noticeable formal errors, *Formal misselection* would cause more serious comprehension problems than *Distortion*. It is often due to a lack of knowledge of word families, previously recognised as a challenging aspect of vocabulary knowledge (Brown et al., 2020). The pedagogical implication of this finding underscores the importance of teaching/practising word families together. An example would be to design corpus-based exercises (e.g., concordances) for commonly confused words from the same word family (e.g., economic, economical, economics, economy) and present them to students for comparison. The same could be done with other formally similar words that have been shown to cause confusion (e.g., obtain, retain, maintain) and further look for similar examples in students' language production and create materials for a targeted focus on form.

Finally, the stylistic categories of *Verbosity* and *Underspecification*, although not at the top of the list, are closely related to the highest ranked errors of confusion of sense

relations, as they are due to the lack of a correct synonym. This leads users to resort to circumlocutions that can result in inappropriate language production. As noted above, notwithstanding the obvious link with vocabulary choice, the two categories also seem to be influenced by non-lexical abilities, such as the inability to develop an idea. The pedagogical implication of this finding is that good use of vocabulary can enhance written production, just as well-developed writing skills can lead to better word choice.

## 5. Conclusion

The purpose of this article was to provide a rationale for adapting an otherwise comprehensive and widely used LET. The new, improved and simplified taxonomy strikes a balance between comprehensiveness on the one hand and clarity and usability on the other. It consists of 11 categories, as shown in the report on the analysis of the Croatian corpus of essays. We believe that the offered taxonomy has solved the problems encountered in the application of the original version. We also trust to have clearly stated various methodological issues and offered clear annotation guidelines thus contributing to the much-needed transparency (Hoffmann, 2022; Lüdeling & Hirshmann, 2015) required to facilitate interpretation of results and replication of studies, as well as further adaptations of the taxonomy to specific research needs. Finally, we hope that the rationale for the changes will be acceptable to other scholars in the field and that the transparent guidelines and clearly explained methodological choices will support further studies in this field to test the taxonomy in a range of contexts.

Turning to the research limitations we highlight the limited size and type of the corpus on which the taxonomy was tested. In future studies, its usefulness should also be tested on different age groups, language levels and participants with different L1s. Similarly, it can be used to address the original motivation of the study, i.e. to examine changes in lexical errors after a period of exposure to a foreign language in order to determine which lexical errors, if any, are more likely to be corrected in written production during a period of natural language acquisition. Furthermore, although we judge the study participants to be advanced users of English as a foreign language based on a number of elements of their language user profile, the study lacks an objective measure of their exact language level. This could be helpful in determining more precisely the nature of lexical accuracy problems typical of a given language proficiency. Finally, we must also humbly admit that we do not refer to previous similar empirical studies when discussing the results for the second research question. We sincerely hope that this can be justified at this point by the focus of the study, i.e. the adaptation of the taxonomy, and the challenges posed by the lack of a uniform approach to lexical error annotation in the literature.

## References

- Agustín Llach, María Pilar. (2011). *Lexical errors and accuracy in foreign language writing*. Multilingual Matters.
- Bardovi-Harlig, Kathleen, & Bofman, Theodora. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in Second Language Acquisition*, 11(1), 17-34.
- Beattie, John. (2014). The ‘ins and outs’ of a study abroad programme. The SALA exchange programme. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts*, (pp. 59-84). John Benjamins. <https://doi.org/10.1075/aals.13.04ch3>
- Borràs, Judith, & Llanes, Àngels. (2021). Investigating the impact of a semester-long study abroad program on L2 reading and vocabulary development. *Study Abroad Research in Second Language Acquisition and International Education*, 6(2), 276-297. <https://doi.org/10.1075/sar.21015.bor>.
- Borràs, Judith, Llanes, Àngels, & Prieto Botana, Goretti. (2023). Development of lexical deployment as a result of a short-term study abroad experience in Costa Rica. In I. Checa-García & L. Marqués-Pascual (Eds.), *Current perspectives in Spanish lexical development* (pp. 137-160). De Gruyter Mouton. <https://doi.org/10.1515/9783110730418-006>
- Brown, Dale, Stoeckel, Tim, McLean, Stuart, & Stewart, Jeff. (2020). The most appropriate lexical unit for L2 vocabulary research and pedagogy: A brief review of the evidence. *Applied Linguistics*, 43(3), 596-602. <https://doi.org/10.1093/applin/amaa061>
- Ellis, Rod. (2015). *Understanding second language acquisition*. Oxford University Press.
- Engber, Cheryl. A. (1995). The relationship of lexical proficiency to the quality of ESL compositions. *Journal of Second Language Writing* 4(2), 139-155. [https://doi.org/10.1016/1060-3743\(95\)90004-7](https://doi.org/10.1016/1060-3743(95)90004-7)
- Granger, Sylviane, Dagneaux, Estelle, Meunier, Fanny, & Paquot, Magali. (2009). *International corpus of learner English*. Presses universitaires de Louvain.
- Hemchua, Saengchan, & Schmitt, Norbert. (2006). An analysis of lexical errors in the English compositions of Thai learners. *Prospect*, 21(3), 3-25.
- Hoffmann, Tim. (2022). Measuring lexical accuracy: The categorization of lexical errors in written L2 English. In A. Leńko-Szymańska & S. Götz (Eds.), *Complexity, accuracy and fluency in learner corpus research [Studies in Corpus Linguistics]* (pp.159-180). John Benjamins Publishing Company.
- Juan-Garau, Maria. (2014). Oral accuracy growth after formal instruction and study abroad: onset level, contact factors and long-term effects. In C. Pérez-Vidal (Ed.), *Language acquisition in study abroad and formal instruction contexts* (pp. 87-110). John Benjamins Publishing Company.

- Köylü, Zeynep, & Tracy-Ventura, Nicole. (2022). Learning English in today's global world: a comparative study of at home, anglophone, and lingua franca study abroad. *Studies in Second Language Acquisition* 44(5), 1330-1355. <https://doi.org/10.1017/s0272263121000917>
- Lennon, Paul. (1991). Error: Some problems of definition, identification, and distinction. *Applied Linguistics*, 12(2), 180-196. <https://doi.org/10.1093/applin/12.2.180>
- Leonard, Karen Ruth, & Shea, Christine E. (2017). L2 speaking development during study abroad: fluency, accuracy, complexity, and underlying cognitive factors. *The Modern Language Journal*, 101(1), 179-93. <https://doi.org/10.1111/modl.12382>
- Lewis, Michael. (1993). *The lexical approach: The state of ELT and a way forward*. Commercial Colour Press.
- Lewis, Michael. (1997). *Implementing the lexical approach: Putting theory into practice*. Language Teaching Publications.
- Llanes, Ángels, & Muñoz, Carmen. (2009). A short stay abroad: Does it make a difference? *System*, 37(3), 353-365. <https://doi.org/10.1016/j.system.2009.03.001>
- Lüdeling, Anke, & Hirschmann, Hagen. (2015). Error annotation systems. In S. Granger, G. Gilquin, & F. Meunier (Eds.), *The Cambridge handbook of learner corpus research* [Cambridge Handbooks in Language and Linguistics] (pp. 135-58). Cambridge University Press.
- MacDonald, Penny, Murcia, Susana, Boquera, María, Botella, Ana, Cardona, Laura, García, Rebeca, Mediero, Esther, O'Donnell, Michael, Robles, Ainhoa, & Stuart, Keith. (2011). Error coding in the TREACLE project. In M. L. C. Pastor & M. Á. C. Mora (Eds.), *Actas del III congreso internacional de lingüística de corpus: Las tecnologías de la información y las comunicaciones: presente y futuro en el análisis de corpus* (pp. 725-740). Universitat Politècnica.
- McManus, Kevin, Mitchell, Rosamond, & Tracy-Ventura, Nicole. (2020). A longitudinal study of advanced learners' linguistic development before, during, and after study abroad. *Applied Linguistics* 42(1), 136-163. <https://doi.org/10.1093/applin/amaa003>.
- Nicholls, Diane. (2003). The Cambridge Learner Corpus – Error coding and analysis for lexicography and ELT. In D. Archer, P. Rayson, A. Wilson, & T. McEnery (Eds.), *Proceedings of the Corpus Linguistics 2003 Conference* (pp. 572-581). University of Lancaster.
- Pallotti, Gabriele. (2009). CAF: Defining, refining and differentiating constructs. *Applied Linguistics*, 30(4), 590-601. <https://doi.org/10.1093/applin/amp045>
- Pavičić Takač, Višnja, & Buljan, Gabrijela. (2017). Exploring EFL learners' lexical competence: What numbers tell us about words. In K. Cergol Kovačević & S. L. Udier (Eds.), *Applied linguistics research and methodology: Proceedings from the 2015 CALS Conference* (pp. 55-70). Peter Lang.
- Pérez-Vidal, Carmen, & Barquin, Elisa. (2014). Comparing progress in academic writing after formal instruction and study abroad. In C. Pérez-Vidal (Ed.), *Language*

- acquisition in study abroad and formal instruction contexts* (pp. 217-235). John Benjamins Publishing Company.
- Picot, Anthony. (2017). Lessons in lexical error analysis. Revisiting Hemchua and Schmitt (2006); an analysis of the lexical errors in the compositions of Greek learners. *International Journal of English Language Teaching*, 5(4), 40-59.
- Polio, Charlene, & Shea, Marc C. (2014). An investigation into current measures of linguistic accuracy in second language writing research. *Journal of Second Language Writing*, 26, 10-27. <https://doi.org/10.1016/j.jslw.2014.09.003>
- Zaytseva, Victoria. (2016). *Vocabulary acquisition in study abroad and formal instruction: An investigation on oral and written lexical development*. (Doctoral dissertation). Universitat Pompeu Fabra.
- Zaytseva, Victoria, Miralpeix, Imma, & Pérez-Vidal, Carmen. (2019). The effects of contexts on the acquisition of oral lexical ability in English as a foreign language. *The Language Learning Journal*, 49(5), 1-17. <https://doi.org/10.1080/09571736.2019.1656766>



# When number is not neutral: Translating Chinese bare nouns into English

Ludovica Lena - Xiamen University

ludolena@xmu.edu.cn

Yue Liu\* - Xiamen University

liuyue@xmu.edu.cn

Rebut /Received: 11-01-24

Acceptat / Accepted: 09-04-24

**Resum. Quan el número no és neutre: Traduir noms escuets xinesos a l'anglès.** Aquest estudi explora els noms escuets xinesos, centrant-se específicament en la seva característica de neutralitat numèrica. Com a cas d'estudi, investiguem el nom escuet més utilitzat per expressar la referència humana general, *rén* 'persona, persones'. Basant-nos en una anàlisi d'un corpus paral·lel xinès-anglès, la recerca examina les propietats a nivell de frase i a nivell de discurs que influeixen en la interpretació de significats singulars vs. plurals inclosos en el nom escuet font quan es tradueix a un idioma com l'anglès, que marca gramaticalment el número. L'anàlisi revela la relació multifactorial entre la marca de número en anglès i l'estructura semàntica (en termes de partició i genericitat), així com les propietats pragmàtiques (relacionades amb l'estat del discurs) del nom escuet font. Es demostra que certs contextos exclouen inequívocament una interpretació singular o plural, especialment en afirmacions genèric-partitives i en oracions que introdueixen línies de discurs directe.

**Paraules clau:** noms escuets, xinès, número, traducció.

**Abstract. When number is not neutral: Translating Chinese bare nouns into English.** This study explores Chinese bare nouns, with a specific focus on their feature of number neutrality. As a case study, we investigate the bare noun most frequently used to convey general human reference, i.e., *rén* 'person, people'. Building upon a parallel corpus composed of Mo Yan's novels and their aligned English translations, the research investigates the sentence-level and discourse-level properties influencing the interpretation of singular vs. plural meanings included within the number-neutral source bare noun when translated into a target language such as English, which

grammatically marks number. The analysis uncovers the multifactorial relationship between number marking in English and the semantic structure (in terms of partitivity and genericity) as well as the pragmatic properties (related to the discourse status) of the source bare noun. It is shown that certain contexts unambiguously preclude a singular or plural interpretation, particularly in generic-partitive statements and sentences introducing direct speech lines.

**Keywords:** bare noun, Chinese, number, translation.

## 1. Introduction

This corpus-based investigation is concerned with Standard Chinese (hereafter Chinese) bare nouns, that is, nouns that appear without any prenominal determiner. Bare nouns in Chinese, as in other languages, are highly multifunctional, since they are undetermined with respect not only to number, but also to definiteness, specificity and genericity (Cheng & Sybesma, 1999; Hole, 2012; LaPolla, 1995; Lena, 2023b; Li, 1997; Rullmann & You, 2006). Specifically, as pointed out by Rullmann & You (2006), Chinese bare nouns convey a general number, referring to both singular and plural referents.

This article builds upon a parallel corpus of Chinese-to-English aligned translations to observe how Chinese bare nouns are translated into a target language that grammatically marks number such as English, and specifically in which contexts the singular vs. plural meaning included in the source bare noun is activated. The bare noun *ren* ‘person, people’ is taken as our case study.

Our working hypothesis is that number marking in the target language is influenced by the semantic-pragmatic properties of the source bare noun, of which it is an indirect representation. Since prior research has pointed out that sentence position is relevant for the interpretation of bare nouns (see Sect. 2), the bare noun *ren* is examined in different syntactic positions, in the light of its semantic and pragmatic properties. It will be shown that, although one-to-one correspondences between a given property and the aligned form in English cannot be expected in most cases, tendencies emerge when the semantic properties of genericity (Kind vs. Entity), partitivity (Whole vs. Part), and the discourse status (Discourse-new vs. Discourse-old) of the source bare noun are considered.

The rest of the article is structured as follows. Sect. 2 illustrates how (in)definiteness oppositions are realized in an article-less language such as Chinese, with a focus on the properties of bare nouns. After presenting the parallel corpus and the methodology adopted in this study (Sect. 3), the article turns to the presentation of the data (Sect. 4), discussing the distribution and types of bare *ren* in the source subcorpus (4.1), the overall number marking in the aligned subcorpus (4.2), and finally linking number marking to the semantic-pragmatic properties of the source bare noun (4.3). A discussion is



offered in Section 5. Finally, Section 6 concludes the article and sums up its main contributions.

## 2. Background: Bare nouns in Chinese

As an article-less language, Chinese lacks the grammatical category of (in)definiteness, and relies on different strategies to mark (in)definiteness oppositions (Chen, 2004). The following examples are extracted from the same novels used to build our corpus (described in Section 3) to briefly illustrate the distinction between bare and non-bare nouns in Chinese. As this contribution is concerned with the bare noun *ren* ‘person, people’, the sentences below will focus on noun phrases (NPs) with *ren* as the nominal head.

The [numeral + classifier] sequence has both referential and non-referential uses (Chen, 2004; LaPolla, 1995). Accordingly, the NP *yī-ge ren* [one-CL person] ‘a person’ denotes a specific individual in (1), while in (2) each occurrence points to a generic individual (through this contribution, the relevant referring expression – if available – in the aligned English sentence is marked in italics):

- (1) Ěr-biān túrán xiǎng-qǐ-le **yī-ge rén** huānkǔ  
 ear=side suddenly sound-INC-PFV<sup>t</sup> one-CL person cheerful  
 de hǎnjiào shēng: (...).  
 SUB shout sound  
 Suddenly my ears filled with the happy shouts of *a man* somewhere: (...).  
 (Life and Death are Wearing Me Out)
- (2) **Yī-ge rén** hèn lìng **yī-ge rén** jìng néng  
 one-CL person hate another one-CL person actually can  
 dá dào rú cǐ qiángliè de chéngdù, zhè wúyí  
 reach such intense SUB degree this undoubtedly  
 shì yī zhǒng měi, yī zhǒng duìyú quán  
 be one kind beauty one kind to all  
 rénlèi de wěidà gòngxiàn.  
 humanity SUB great contribution  
 If *one person’s* hatred for *another* could reach such proportions, it was an  
 unquestioned form of beauty, a magnificent contribution to humanity.  
 (The Republic of Wine)

Nouns modified by demonstrative determiners, i.e. the proximal demonstrative *zhe* ‘this’ and the distal demonstrative *na* ‘that’, are used to encode identifiable referents,

discourse-old or inferable, (LaPolla, 1995)<sup>1</sup>. Example (3) features two ‘definite’ NPs with *ren* as their nominal head, marked with the proximal (*na-ge ren* ‘that person’) and distal (*zhe-ge ren* ‘this person’) demonstrative determiner, respectively.

- (3) Fùqīn qiánbian            **nà-ge rén**      kēngkēng-de      késou-qilai,  
 father in.front            that-CL person ONOM-ADV cough-INC  
**zhè-ge rén**      de      kèsòu shēng fēicháng      shúxī.  
 this-CL person SUB cough sound very            be.familiar.with  
*The man* in front of Father coughed loudly. It was a familiar cough (Lit.  
 ‘The/this person’s cough was familiar’), (...). (Red Sorghum Clan)

Besides nouns that are marked either by a quantifying expression or by a demonstrative determiner, Chinese speakers can make use of bare nouns as referring expressions, as mentioned above. Bare nouns are an interesting site of study because of their formally unmarked status and their highly multifunctional nature. Bare nouns in Chinese are not only number-neutral, they can also denote discourse-old, discourse-new and generic referents. They resemble English bare plurals in that they can have a generic or specific interpretation, but in addition, they can also convey definite reference (Bremmers et al., 2021; Cheng & Sybesma, 1999; Dayal & Jiang, 2022; Jenks 2018; Li, 1997).

From the point of view of propositional semantics, it has been acknowledged in the literature that Chinese bare nouns have both Kind-level and Entity-level interpretations (Li, 2013; Li & Thompson, 1981; Rullman & You, 2006; Yang, 1998). Previous studies also pointed out a semantic correlation between the interpretation of bare nouns and the type of predicates (Li, 1997; Kuo, 2008). That is, as subjects of individual-level predicates (i.e. predicates denoting characterizing properties), bare nouns have a generic interpretation, while as subjects of stage-level predicates (i.e. predicates denoting transient properties), they are interpreted as Entity-referring (Carlson, 1977; Kratzer, 1988; Krifka et al., 1995).

In turn, the Entity-level meaning can be both definite and indefinite, and pragmatics play a role in the interpretation (Kuo, 2008). From the point of view of the discourse structure, bare nouns as referring expressions present a high degree of ambiguity, given that they can designate both unidentifiable and identifiable referents (Sun, 1988; LaPolla, 1995)<sup>2</sup>. Generally speaking, bare nouns in preverbal position tend to receive

1. The glosses used through the article are as follows. ADV: adverbial particle, CL: classifier, COLL: collective marker; CRS: Current Relevant State aspect marker, DUR: durative aspect marker, INC: inchoative aspect marker; NEG: negation, ONOM: onomatopoeia; PASS: passive marker, PFV: perfective aspect marker, PROG: progressive aspect marker, SFP: sentence final particle, SG: singular, SUB: subordinative particle, VEN: venitive marker.

2. On accessibility and identifiability (Prince 1981; Ariel 1990; Gundel et al., 1993; Chafe 1994; Lambrecht 1994).

a definite (i.e., generic or discourse-old) interpretation, while they are interpreted as indefinite in postverbal position (Cheng & Sybesma, 1999; Hole 2012; Lena, 2020a).

On these theoretical grounds, the current work argues that number marking can be understood in the light of the following interacting properties of the source bare noun: **GENERICITY** (Kind vs. Entity), **PARTITIVITY** (Whole vs. Part), and **DISCOURSE STATUS** (Discourse-old vs. Discourse-new). Let us begin by presenting the corpus and methodology adopted, before turning to the discussion of the data.

### 3. Corpus and Method

The parallel corpus compiled for this study consists of six Chinese novels written by the Nobel literature prize laureate Mo Yan (i.e. *The Garlic Ballads*, *Sandalwood Death*, *The Republic of Wine*, *Big Breasts & Wide Hips*, *Life and Death are Wearing Me Out*, *The Red Sorghum Clan*) and their aligned translations into English (all translated by Howard Goldblatt). The corpus was constituted via the *English-Chinese Parallel Concordancer* made available by the Department of Linguistics and Modern Language Studies at the Education University of Hong Kong (<https://corpus.eduhk.hk/paraconc>).

TABLE I. THE CORPUS

Novels	English title	Tokens (words)	Sentences
<i>Tiāntáng suàntái zhī gē</i>	The Garlic Ballads	126.075 (101.939)	6.571
<i>Tānxiāng xíng</i>	Sandalwood Death	195.849 (163.225)	8.954
<i>Jiǔ guó</i>	The Republic of Wine	137.288 (113.921)	6.491
<i>Fēng rǔ fēi tún</i>	Big Breasts and Wide Hips	330.308 (271.730)	17.542
<i>Shēngsǐ píláo</i>	Life and Death are Wearing Me Out	276.532 (229.790)	11.430
<i>Hóng gāoliang jiāzú</i>	The Red Sorghum Clan	161.140 (134.257)	7.730
Total		1.227.192 (1.014.862)	74.243

The data extraction was done in two steps. Firstly, a search of the character 人 [rén] was performed within the corpus, which gave 7895 raw results. A randomized selection of 200 hits was then coded by the author of this work with the objective of identifying instances of bare *ren*, as distinguished by both *ren* preceded by a prenominal modifier and instances of lexicalized items including *ren* as a bounded morpheme (table 2).

This procedure allowed us to identify the contexts of occurrence of *rén*, i.e. as a subject (*ren\_S*), an object (*ren\_O*), as a pivot (*ren\_P*), as the agent in passive constructions

(*ren\_A*). “Pivot” is a term used in the literature on existential constructions (Bentley et al., 2015; McNally, 2011; Sarda & Lena, 2023). In this study it denotes the bare noun *ren* appearing in existential-presentational constructions, namely *you ren* [exist person] ‘there’s someone, there are people’ (Lena, 2023b). The label “agent” (*ren\_A*) is used for commodity and denotes a specific syntactic position whereby *ren* is introduced by the passive marker *bei*.

Next, a random sample of each *ren*-type (*ren\_S*, *ren\_O*, *ren\_P*, *ren\_A*, *ren\_CAUS*) was extracted from the corpus (n = 50; N = 250; cf. table 3).

The English aligned sentences were then coded according to the type of referring expression aligned with *ren*. Assuming that “a high quality translation creates a grammatically correct mirror image of the original meaning in the target language” (Tellings et al., 2022), the parallel corpus methodology aims at identifying some of the interacting factors that contribute to activating, in the English translation, the singular or plural meaning included in the general number of the source bare noun *ren* ‘person, people’.

As a study based on a limited corpus featuring novels from a single author, this work serves as exploratory research that lays the groundwork for future studies utilizing larger and more diverse datasets.

## 4. Results

### 4.1. The source subcorpus

Table 2 below shows the general distribution of the morpheme [ren] in the randomized sample of 200 occurrences extracted from the corpus (the extraction method was presented in Sect. 3). The occurrence of bare *ren* is rather infrequent (14 %) with respect to its appearance with prenominal modifiers (41.5 %) and as a lexicalized item (36 %).

TABLE 2. EXTRACTION OF THE MORPHEME [REN] FROM THE CORPUS

Category	Forms	Examples	Count	%
Pre-nominal modifiers + N	RC + <i>ren</i>	<i>huo-zhe de ren</i> [live-DUR SUB person]	27	13.5 %
	ADJ + <i>ren</i>	<i>hao ren</i> [good person]	24	12 %
	QUANT + <i>ren</i>	<i>yi ren</i> [one person]	21	10.5 %
	DEM + <i>ren</i>	<i>na ren</i> [that person]	11	5.5 %
	Subtotal		83	41.5 %

Bare noun	<i>ren_S</i>	<i>ren</i> V [person V]	2	1 %
	<i>ren_O</i>	V <i>ren</i> [V person]	7	3.5 %
	<i>ren_P</i>	<i>you ren</i> [exist person]	10	5 %
	<i>ren_A</i>	<i>bei ren</i> V [PASS person V]	3	1.5 %
	<i>ren_CAUS</i>	<i>rang ren</i> V [make person V]	5	2.5 %
	<i>ren_DITR</i>	<i>gei ren</i> O [give person O]	1	0.5 %
	Subtotal		<b>28</b>	14 %
<i>Ren-men</i>	<i>ren-men_S</i>	<i>ren-men</i> V [person-COLL V]	<b>6</b>	3 %
Lexicalized	<i>x-ren</i>	<i>fanren</i> ‘criminal’	44	22 %
	<i>ren-x</i>	<i>renjia</i> ‘the others’	28	14 %
	Subtotal		<b>72</b>	36 %
Other		<i>ge ren</i> ‘personally’	11	5.5 %
Total			200	100%

Figure 1 below illustrates the distribution of *ren* as a bare noun with respect to the other macro-categories.

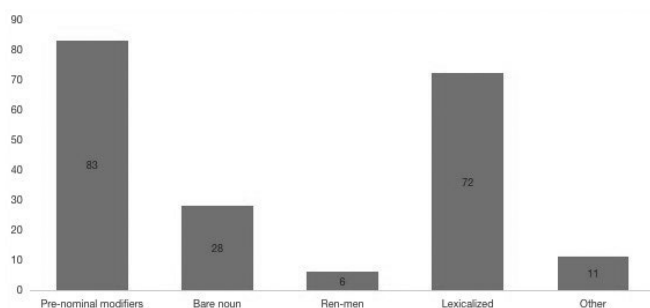


FIGURE 1. RELATIVE FREQUENCY OF *REN* AS A BARE NOUN

This exploratory search further allowed us to identify the contexts of occurrence of *ren* as a bare noun. As previously discussed, *ren* can appear as a subject (*ren\_S*) – although infrequently, as an object (*ren\_O*), including the separately-calculated object position after a causative verb (*ren\_CAUS*) and the pivot position in existentials (*ren\_P*). It can also

be an agent introduced by the passive marker (*ren\_A*) or, marginally, it can be a recipient in ditransitive constructions (*ren\_DITR*).

On these grounds, we then extracted 50 random examples for each type of bare *ren* based on its position and function in the sentence, with the exception of *ren* appearing in ditransitive constructions (*ren\_DITR*), due to an insufficient number of relevant occurrences, as seen in table 3. The total 250 Chinese sentences were then analysed by considering the referring expressions aligned with *ren* in the corresponding English sentences, as discussed in the next subsection.

TABLE 3. THE BARE-REN CORPUS

Types	Count	%
<i>Ren_S</i>	50	20 %
<i>Ren_O</i>	50	20 %
<i>Ren_A</i>	50	20 %
<i>Ren_P</i>	50	20 %
<i>Ren_CAUS</i>	50	20 %
Subtotal	250	100 %

#### 4.2. Number marking in the aligned subcorpus

Let us now turn to the discussion of the aligned data. To begin with, table 4 below illustrates the distribution of the number marking in the English aligned referring expressions according to the types of *ren*. Importantly, the categories “singular” and “plural” co-exist with the category “irrelevant”. This encompasses all the cases in which number is not expressed because a relevant referring expression is not provided in the English sentence.

TABLE 4. NUMBER MARKING IN THE ALIGNED SENTENCES

	<i>Ren_S</i>	<i>Ren_O</i>	<i>Ren_A</i>	<i>Ren_P</i>	<i>Ren_CAUS</i>	Total
Plural	41	13	11	15	13	93 (37.2 %)
Singular	9	15	6	26	16	72 (28.8 %)
Irrelevant	0	22	33	9	21	85 (34 %)
Subtotal	50	50	50	50	50	250 (100 %)

When all sentences are considered together, the mean values do not show a preference toward one particular category (the plural category being only slightly more represented

at 37.2 %). However, when each type of *ren* is considered separately, interesting tendencies emerge. For the sake of clarity, the data are represented in figure 2 below. First, the relevant human referent is always expressed in the aligned translation when the source bare noun *ren* has the function of a subject (*ren\_S*). Secondly, the aligned referring expression strongly tend to have plural number. By contrast, when *ren* is an existential pivot (*ren\_P*), the aligned referring expression has more often singular number. The categories *ren\_O* and *ren\_CAUS* show a similar distributive behaviour, with similar proportions of singular and plural marking, and a relatively more frequent portion of sentences in which number is unmarked. Finally, when *ren* is the agent in passive constructions (*ren\_A*), it strongly tends to remain unexpressed in the aligned sentence.

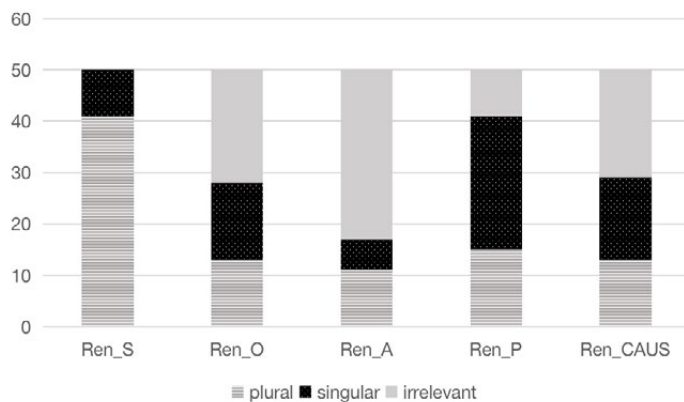


FIGURE 2. NUMBER MARKING IN THE ALIGNED SENTENCES

### 4.3. Linking number marking to the semantic-pragmatic properties of the source bare noun

#### 4.3.1. General observations

Let us begin the discussion by considering the types and frequency of the aligned referring expressions in the English sentences, as seen in Table 4. A great deal of variation can be observed, in which potentially Kind-referring expressions such as *people* coexist with potentially Entity-referring expressions, in turn likely to be Discourse-new (e.g. *someone*,  $a + N_{sc}$ ) or Discourse-old (e.g. personal pronouns). Interestingly, lexical categories such as *humans* or *every human* (and also the bare singular *man*) unambiguously point to the class of human beings in their entirety.

TABLE 4. FORM OF THE ALIGNED REFERRING EXPRESSIONS

Aligned referring expressions (REs)	Total
people	36
personal pronoun	27
<i>someone (somebody)</i>	24
humans / every human	12
$a + N_{SG}$	12
anyone (anybody, whoever)	8
everyone (everybody)	7
<i>some (+ N<sub>PL</sub>)</i>	7
the people	6
<i>no one / nobody / no N<sub>SG</sub></i>	4
others	4
<i>one of + N<sub>PL</sub></i>	3
<i>the + N<sub>PL</sub></i>	3
<i>another (+N<sub>SG</sub>)</i>	2
man	2
<i>the + N<sub>SG</sub></i>	2
Other REs (n=1)	8
Unexpressed	83
Total	250

In what follows, each type of *ren* is analysed individually, since interpretative differences – likely to influence number marking in the aligned sentences – are expected between the different positions. In particular, recall from Sect. 2 that previous studies pointed out a correlation with (in)definiteness: bare nouns tend to receive a definite reading in preverbal position and are interpreted as indefinite in postverbal position.

#### 4.3.2. The subject position

Preverbal *ren* (*ren<sub>S</sub>*) is often found in generic statements where it refers to the kind of “human beings”. In (4), the English sentence features a noun modified by the indefinite article (i.e. *a man*) – although a bare plural (e.g. *men* or *people*) would be a perfectly acceptable option (and often is, cf. 4.2). By contrast, in (5) *ren* is Entity-referring, as it points to some particular individual. Additionally, this entity is evoked in discourse, hence the alignment with the English pronoun *him* is made possible, signaling an



element in high activation state (Gundel et al., 1993). In other words, its referent is Entity-referring and Discourse-old, identifiable by the interlocutor.

- (4) Kind-referring (Whole).

Rén guò liú míng, yàn guò liú shēng.  
 person pass leave name wild.goose pass leave sound  
*A man* leaves behind his good name, a wild goose leaves behind its call.  
 (Big Breasts and Wide Hips)

- (5) Entity-referring (Discourse-old)

Rén fǎnzhèng sǐ-le, yīqiè dōu yào kǎolù  
 person anyway die-PFV everything all must consider  
 huó-zhe de rén.  
 live-DUR SUB person  
 Shuōchuān le, jiù shì qián!  
 disclose CRS then be money

It's too late to do anything for *him* now (Lit. 'The person has died anyway'),  
 so it's time to think about yourselves – in other words, money.  
 (The Garlic Ballads)

In its Entity interpretation, number marking in the aligned sentence is unpredictable due to the context-dependent retrieval. However, *ren\_S* is prevalently Kind-referring in the corpus, and that motivates its frequent alignment with plural referring expressions in English (cf. Sect. 4.2). In both cases, *ren\_S* is treated as a topical element and cannot remain unexpressed in the English translation.

#### 4.3.3. The object position

The occurrence of *ren\_O* was examined after the verbs *da* 'hit', *ma* 'insult', *jian* 'see', *sha* 'kill', i.e., the most frequent verbs observed in the exploratory corpus search (Sect. 4.1). In this context, *ren\_O* is often Entity-referring and nonspecific. In such nonspecific uses, it is either unexpressed (e.g. (6)), either aligned with *someone* or *anyone* (e.g. (7)).

- (6) Entity-referring, nonspecific.

“Bù xǔ mà rén!”  
 NEG allow insult person  
 “No cursing allowed,” (...). (The Garlic Ballads)

- (7) Entity-referring, nonspecific.

Wǒ shuō Zhào Jiǎ, shā rén shí nǐ shì

1SG say Zhao Jia kill person time 2SG be  
 pà háishì bú pà?  
 be.afraid or NEG be.afraid  
 I say, Zhao Jia, when you kill *someone*, are you afraid? (Sandalwood Death).

However, if *ren\_O* appears in contrastive constructions where human beings are contrasted with other species, the Kind-interpretation can be activated:

(8) Kind-referring (Whole).

Jiàn rén dǎ rén, jiàn gǒu dǎ gǒu.  
 see person hit person see dog hit dog  
*Every human* who got in their way tasted the whips, as did the dogs. (Sandalwood Death)

(9) Kind-referring (Whole).

Wǒ zhīdào tāmen huì kāi qiāng shā rén, dàn  
 1SG know 3PL can open fire kill person but  
 tāmen bú huì kāi qiāng shā lú.  
 3PL NEG can open fire kill donkey  
 I knew they wouldn't hesitate to shoot *a person*, but would never shoot a donkey. (Life and Death Are Wearing Me Out).

Despite the object position, *ren\_O* can then receive a Kind (definite) interpretation, although less frequently than *ren\_S* in our corpus.

As the object of the causative verb *ràng* 'let, make', *ren\_CAUS* does not behave much differently than *ren\_O*. The only difference being that when *ren\_CAUS* is Entity-referring and specific, it can denote discourse-old entities:

(10) Entity-referring (Discourse-old)

...dāihuì ràng rén sòng-lai.  
 in.a.while make person deliver-VEN  
 I'll have *one of my men* bring it over. (Big Breasts & Wide Hips).

#### 4.3.4. The agent position

As previously mentioned, *ren\_A* stands out insofar as it is most often unexpressed in the aligned sentence (Sect 4.2), as seen in (11). In a few cases, however, the Kind interpretation can be activated by the context, as in (12), and the aligned English sentence must express the relevant human referent.

(11) Entity-referring, non-specific.

“Shì-bù-shì bèi rén yān le?”

be-NEG-be      PASS      person castrate PFV  
 Not castrated  $\emptyset$ , are you? (The Republic of Wine).

- (12) Entity-referring, non-specific.

Xiànzài tā shì yī-piàn fèixū le, rén chuàngzào  
 now 3SG be one-CL ruins CRS person build  
 de, yòu bèi rén cuīhuǐ.  
 SUB again PASS person destroy  
 Now the village lay in ruins; man had created it, and *man* had destroyed it.  
 (Red Sorghum Clan).

#### 4.3.5. The pivot position

Finally, when appearing after the existential verb *you* ‘there be, exist’, *ren\_P* can be Entity-referring (as in (13)) or Kind-referring (as in (14))<sup>3</sup>:

- (13) Entity-referring (Discourse-new).

“Nà shì shéi?” Hēi“àn=li yǒu rén hǎn.  
 that be who darkness=in exist person scream  
 “Who’s there?” came a voice out of the darkness. (Lit. ‘...someone shouted  
 in the darkness’). (The Garlic Ballads).

- (14) Kind-referring (Part).

...yǒu rén shuō nà xiē huāwén shì tiānshàng  
 exist person say that few pattern be heaven  
 de kēdǒu wén.  
 SUB tadpole script

A design had been scorched into her back, which *some people* said was the script of heavenly tadpoles. (Red Sorghum Clan)

Hence, *ren\_P* can denote both Entity-referring and Kind-referring referents, with however some peculiarities. Kind-referring *ren\_P* belong to the generic-partitive subtype, that is, in the overwhelming majority of cases, they refer to a part of the generic set of “people”. This contrasts with sentences previously seen in which Kind-referring *ren* (often *ren\_S*, but also possible in the other positions except for *ren\_P*) refer to the overall generic set of “people”. The data show that when the source bare noun is positive with respect to the two parameters of partitivity and genericity, the plural meaning of the aligned referring expression is systematically activated in the English sentence (e.g. in (14)).

3. See Lena (2024a, 2024b) for a detailed analysis of the bare noun *ren* in the pivotal position.

One aspect complicates the picture emerged so far. The partitive interpretation is not limited to generic sentences. Entity-referring *ren\_P* can also denote discourse entities that stand in Set-Member relation with a Discourse-old referential group, as in (15):

- (15) Entity-referring (Discourse-anchored).  
 Zhàlán=li chòuqìxūntiān, yǒu rén zài dǎhūlu,  
 fence=in stink.to.high.heaven exist person PROG snore  
 yǒu rén wǎng zhàlán=biān jiǎo=shang  
 exist person towards fence=side corner=on  
 nà-ge tiěpí shuītǒng=li sāniào, (...).  
 that-CL iron pail=in urinate  
*Some of the men* snored loudly; *others* got up to piss in a tin pail, (...).  
 (Red Sorghum Clan).

Both (13) and (15) include entity-referring participants, which are Discourse-new in the former case, and Discourse-anchored in the latter (Prince, 1981). This also means that, although *ren\_P* can pick up a member of a discourse-old group, it cannot directly denote a discourse-old referent, hence the “definite” (discourse-old) interpretation is not available. As in the case of discourse-old *ren\_S*, nonetheless, the interpretation of discourse-anchored entities is also context-dependent, and the choice of the aligned referring expression in English, including the singular vs. plural marking, must take the previous context into account.

In the case of Discourse-new referents, number marking is dependent on the propositional semantics at the sentence level. An illustrative example is provided in (16), where the sentence-initial locative expression *daolu liang bian* ‘at the two sides of the road’ implies that at least one person is located on each side of the road. The English translation consequently uses the plural referring expression *crowds of gawkers*:

- (16) Entity referring (Discourse-new)  
 Dàolù liǎng biān shǐzhōng yǒu rén guānkàn (...).  
 road two side all.along exist person watch  
 ...but always there were *crowds of gawkers*, (...). (The Garlic Ballads)  
 Lit. ‘There were people watching on both sides of the road.’

In this respect, one context systematically activates the singular meaning in the aligned English sentence, that is, when the sentence introduces a direct speech line, as in (17–18):

- (17) Entity referring (Discourse-new)  
 Xià yǔ le, yǒu rén rǎngjiào.  
 fall rain CRS exist person scream  
 ‘It’s raining,’ *someone* shouted, (...). (Big Breasts & Wide Hips).

## (18) Entity referring (Discourse-new)

Ān tīngdào yǒu rén shuō: “Dūtǒng dàrén,  
 ISG hear exist person say Commander Great  
 zhè shì zhīxiàn dàrén de sīzhái!”  
 this be Magistrate Great SUB private.residence  
 “Commander,” *someone* said, “these are the Magistrate’s private quarters.”  
 (Sandalwood Death).

Direct speech not only unambiguously activates Entity-referring reading of *ren*<sub>P</sub>, it also systematically identifies the relevant entity as a single participant.

## 5. Discussion

In different syntactic positions, the bare noun *ren* can receive Kind or Entity readings. The Kind reading can be further distinguished in UNIVERSAL (when it denotes the Whole) and PARTITIVE (when it denotes the Part). Universal kind readings can be found in all syntactic positions except for *ren*<sub>P</sub>, but are particularly represented in the subject position (*ren*<sub>S</sub>). The relevant referents are typically encoded by plural referring expressions in the English sentence (e.g., *people*, *humans*), although singular forms with generic interpretation are not excluded (e.g., *a man*, *a person*).

In its Entity-referring interpretation, *ren* can be Discourse-new, Discourse-old (or Discourse-anchored, as far as *ren*<sub>P</sub> is concerned). Number marking for Discourse-old *ren* is dependent on the previous context, and can result in forms typically high in activation scales (e.g., personal pronouns). In the case of Discourse-new *ren*, by contrast, the number interpretation is dependent on the propositional meaning of the sentence.

One context, in particular, systematically activates the singular meaning of the aligned referring expression, namely, when *ren*<sub>P</sub> introduces a direct speech line. Despite the number-neutrality property of the source bare noun *ren*, the choice of an aligned referring expression with plural number in these contexts would often result in an infelicitous outcome. The opposite situation is observed with Kind-referring pivot *ren*. When *ren* is a pivot in existential constructions (*ren*<sub>P</sub>), it can only receive a generic-partitive reading, as far as the Kind interpretation is concerned. By definition, the generic-partitive interpretation refers to more than one individual, and the corresponding aligned referring expression must be plural (*some people*, *certain individuals*). In this scenario, selecting a singular referring expression in English would similarly result in an inappropriate choice.

In most cases, however, English has to choose one option (e.g., a referring expression with singular reference) while the other option (e.g., a referring expression with plural reference) could also be accepted, as the relevant opposition are semantic-pragmatic in nature, of which the singular vs. plural morphological marking is only an opaque representation. As a language morphologically marking number, the number-neutral

property of bare nouns in Chinese is generally lost in translation. Different strategies, however, allow to keep the number indeterminacy. Recall in this respect the high portion of sentences in which *ren* was *not* expressed in the English sentence (Sect. 4.2). One such strategy is the use of an agent-less passive construction, in which the relevant entity is inferred from the meaning of the predicate, as previously seen. Another strategy is found in sentences such as (19), involving the thematization of the location, while the human referent is only inferred from the meaning conveyed by the predicate:

- (19) Wéiqiáng=shang hěn píngjìng, sìhū méi yǒu rén.  
 enclosing.wall=on very quiet seemingly NEG exist person  
 They pressed forward, drawing closer and closer to the wall, which appeared **unmanned**. (Sandalwood Death)<sup>4</sup>

In sentences such as (19), the essential meaning is the human presence (or absence), rather than the specific number of individuals involved. This concept is further exemplified in English sentences like (20), where the term *manned* synthetically captures this idea without overloading the sentence with momentarily irrelevant information such as the number of astronauts involved.

- (20) Elon Musk's SpaceX COO says **manned** missions will reach Mars by 2030.<sup>5</sup>

## 6. Conclusion

Building upon parallel corpus methodology, this research delved into the number-neutral nature of Chinese bare nouns, with specific attention on the translation of the bare noun *ren* 'person, people' into English.

It revealed that the distinction between singular and plural markings in the English translations indirectly reflects the semantic and pragmatic properties of the source bare noun, which in turn are dependent from its position in the sentence.

Certain contexts unequivocally rule out the singular or plural interpretation respectively. When the sentence introduces a direct speech line, the English translation invariably opts for singular referring expressions. Conversely, the Kind-referring reading of the generic-partitive subtype activates the use of plural referring expressions in English (e.g., *some people*, *certain individuals*, etc.). This stands in contrast to the generic-universal interpretation of *ren*, which can be conveyed in translation through the generic

4. The backgrounding of the human entity is often found in negative you ren constructions, as in (19) (see Lena, 2024c, for a full discussion).

5. Retrieved from: <https://www.euronews.com/next/2022/05/09/elon-musk-s-spacex-coo-says-manned-missions-will-reach-mars-by-2030-nasa-says-otherwise>.

usage of the indefinite article (e.g., *a man*), although the use of the bare plural *people* is dominant.

Finally, it has been pointed out that the referent of *ren* is often left unexpressed in the aligned subcorpus, resulting in several strategies allowing the number-neutral properties of the source bare noun to be kept.

In conclusion, this work has served as a preliminary investigation on the translation strategies of Chinese bare nouns. However, it is important to highlight that this case study – based on works by one author and translated by a single translator – requires corroboration from larger and more diversified datasets to substantiate the hypotheses it presents.

## References

- Ariel, Mira (1990). *Assessing Noun-Phrase Antecedents*. London & New York: Routledge.
- Bremmers, David, Liu, Jianan, van der Klis, Martijn & Le Bruyn, Bert (2021). Translation mining: Definiteness across languages (A reply to Jenks 2018). *Linguistic Inquiry*, 1-18.
- Bentley, Delia, Ciconte, Francesco Maria & Cruschina, Silvio (2015). *Existentials and Locatives in Romance Dialects of Italy*. Oxford University Press.
- Carlson, Gregory N. (1977). *Reference to kinds in English*. PhD dissertation, University of Massachusetts at Amherst.
- Chafe, Wallace L. (1994). *Discourse, consciousness, and time: The flow and displacement of conscious experience in speaking and writing*. The University of Chicago Press.
- Chen, Ping (2004). Identifiability and definiteness in Chinese. *Linguistics* 42(6), 1129-1184.
- Cheng, Lisa Lai-Shen & Sybesma, Rint (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic Inquiry*, 30(4), 509-542.
- Dayal, Veneta & Jiang, Li Julie. (2022). The puzzle of anaphoric bare nouns in Mandarin: A counterpoint to *Index!*. *Linguistic inquiry*, 54(1), 147-167.
- Gundel, Jeanette K., Hedberg, Nancy & Zacharski, Ron (1993). Cognitive status and the form of referring expressions in discourse. *Language* 69 (2), 274-307.
- Hole, Daniel (2012). The information structure of Chinese. In Manfred Krifka & Renate Musan (Eds.), *The expression of information structure*, 45-70. De Gruyter.
- Jenks, Peter (2018). Articulated definiteness without articles. *Linguistic Inquiry* 49(3), 501-536.
- Kratzer, Angelika (1988). Stage-level and individual-level predicates. In Manfred Krifka (ed.), *Genericity in natural language*, 247-84. University of Tubingen.
- Krifka, Manfred, Pelletier, Francis J., Carlson Gregory N., ter Meulen, Alice, Link, Godehard & Chierchia, Gennaro. (1995). Genericity: An introduction. In Gregory N. Carlson & Francis J. Pelletier (eds.), *The generic book*, 1-124. The University of Chicago Press.

- Kuo, Jenny Yi-chun (2008). A pragmatic approach to the interpretations of Mandarin bare nouns. *Journal of Pragmatics* 40(6). 1082-1102.
- Lambrecht, Knud (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge University Press.
- LaPolla, Randy J. (1995). Pragmatic relations and word order in Chinese. In Pamela A. Dowling & Michael Noonan (eds.), *Word order in discourse* (pp. 297-329). John Benjamins.
- Lena, Ludovica (2017). Les énoncés « présentatifs » chez les apprenants sinophones de français L2. *SHS Web of Conferences*, 38. EDP Sciences.
- Lena, Ludovica (2020a). Presentational sentences: A study on the information structure of Path verbs in spoken discourse. In Dongyan Chen & Daniel Bell (eds.), *Explorations of Chinese theoretical and applied linguistics*, 45-81. Cambridge Scholars.
- Lena, Ludovica (2020b). Referent introducing strategies in advanced L2 usage: A bi-directional study on French learners of Chinese and Chinese learners of French. In Jonathon Ryan & Peter Crosthwaite (eds.), *Referring in a second language: Studies on reference to person in a multilingual world*, 164-183. Routledge.
- Lena, Ludovica (2022). An unbridgeable gap? The treatment of definiteness restrictions in French and Chinese presentational constructions. *Discours*, 31.
- Lena, Ludovica (2023a). The (non-)acquisition of the Chinese Definiteness Effect: A usage-based account. In Howard Hao-Jan Chen, Keiko Mochizuki & Hongyin Tao (eds), *Learner corpora: Construction and explorations in Chinese and related languages*. Springer.
- Lena, Ludovica (2023b). Partition and existence: The case of *you ren* 'there's someone, there are people' in Chinese. In L. Sarda and L. Lena (eds.), *Existential constructions across languages: Forms, meanings and functions*, 245-282. John Benjamins.
- Lena, Ludovica (2024a). Encoding indefinite human reference without indefinite pronouns: the case of Chinese presentationals. *Folia Linguistica*, 58(3), 659-698.
- Lena, Ludovica (2024b). Looking for someone: The encoding of indefinite human reference in Chinese/English aligned translation. *Lingua*, 308, 1-24.
- Lena, Ludovica (2024c). Negative existential and presentationals in Chinese: A parallel corpus investigation. *Languages in contrast* 24(2), 226-247.
- Li, Charles. N. & Thompson, Sandra. A. (1981). *Mandarin Chinese: A functional reference grammar*. University of California Press.
- Li, Jie (1997). Predicate type and the semantics of bare nouns in Chinese. In Liejiong Xu (ed.), *The referential properties of Chinese noun phrases*, 99-115. Ecole des hautes études en Sciences sociales, Centre de recherches linguistiques sur l'Asie orientale.
- Li, Xuping (2013). *Numeral classifiers in Chinese: The syntax-semantics interface*. De Gruyter.
- McNally, Louise (2011). Existential Sentences. In Klaus von Stechow, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 1829-48. De Gruyter.



- Prince, Ellen F. (1981). Toward a Taxonomy of Given-New Information. In Peter Cole (ed.), *Radical Pragmatics*, 223-256. Academic Press.
- Rullmann, Hotze & You, Aili (2006). General number and the semantics and pragmatics of indefinite bare nouns in Mandarin Chinese. In Klaus von Heusinger & Ken Turner (eds.), *Where Semantics Meets Pragmatics*, 175-196. Brill.
- Sarda, Laure & Lena, Ludovica (2023). Existential constructions: In search of a definition. In Laure Sarda & Ludovica Lena (Eds), *Existential constructions across languages: Forms, meanings and functions*, 1-32. John Benjamins.
- Sun, Chaofen (1988). The discourse function of numeral classifiers in Mandarin Chinese. *Journal of Chinese Linguistics* 16(2), 298-323.
- Tellings, Jos, Fuchs, Martín, van der Klis, Martijn, Le Bruyn, Bert & de Swart, Henriëtte. (2022). Perfect variations in dialogue: A parallel corpus approach. *Proceedings of SALT* 32.
- Yang, Ronglan (1998). Chinese bare nouns as kind-denoting terms. *RuLing Papers*, 1, 247-288.



# The importance of corrective feedback in foreign language learning

Mirjana Matea Kovač - University of Split  
mirjana@ffst.hr

Rebut / Received: 25-01-24

Acceptat / Accepted: 21-06-24

**Resum. La importància del feedback correctiu en l'aprenentatge de llengües estrangeres.** En contextos d'aprenentatge de llengües, la retroalimentació correctiva oral i escrita és un element essencial de l'aprenentatge de la llengua. Pel que fa a les opinions dels estudiants sobre el tipus d'errors i correccions que es produeixen, els estudiants estan d'acord que corregir errors simples, així com errors més complexos, comporta canvis en l'interllenguatge dels aprenents. També, segons ells, els treballs escrits s'han de corregir a fons i proporcionant una explicació metalingüística. En general, es pot concloure que els estudiants van mostrar actituds positives pel que fa a la retroalimentació correctiva oral i escrita. A més, la retroalimentació correctiva retardada és una estratègia significativament més preferida en comparació amb la retroalimentació correctiva immediata. Els resultats obtinguts indiquen una comprensió més exhaustiva del fenomen de la retroalimentació correctiva, tenint en compte que aquest tema de recerca no està prou investigat.

**Paraules clau:** *feedback* correctiu, *feedback* immediat i retardat, errors de pronunciació, errors gramaticals.

**Abstract. The importance of corrective feedback in foreign language learning.** In language learning contexts, oral and written corrective feedback are essential elements of language learning. In this research study the affective perspective of learner engagement with corrective feedback was analysed. Regarding students' views on the type of occurring errors and corrections, students agree that correcting simple, as well as more complex errors, leads to changes in the learners' interlanguage. Also, in their opinion, written assignments should be thoroughly corrected, followed by metalinguistic explanation. Overall, it can be concluded that students displayed positive attitudes regarding corrective feedback in spoken and written performance. Moreover, delayed corrective feedback is a significantly more preferred strategy compared to immediate

corrective feedback. The obtained results indicate a more comprehensive understanding of the phenomenon of corrective feedback, considering that this research topic is insufficiently researched.

**Keywords:** corrective feedback, immediate and delayed feedback, pronunciation errors, grammatical errors.

## 1. Introduction

Error correction has always been a controversial topic. It refers to interventions that alert the student to a grammatical, lexical, stylistic and pronunciation error that needs to be corrected. In both behavioural and cognitive theories of language learning, feedback is an essential element of language learning (Şakiroğlu, 2020). There is an idea that it can be harmful and ineffective on the one hand, as well as essential and beneficial for some grammatical errors, on the other hand (Veliz, 2007). In order to fully grasp the underlying mechanism of error correction, it is important to briefly explain the nature of speech production in L2. The production of output in spoken and written performance differs significantly, and these differences are also evident in the ways of providing corrective feedback. Spoken communication takes place in real time, where the cognitive processing loads are greater. Planning and the production of utterances take place simultaneously, posing significant demands, especially in a foreign language where the use of grammatical rules is non-automated. This is compounded by limited morphosyntactic and lexical knowledge, as well as restricted attention and working memory capacity. On the other hand, in a written performance, the language learner usually has time to plan, and therefore the corrective feedback is delayed in most cases, unlike in a spoken performance, where the learner, in most cases, is corrected immediately after the error occurs. In addition, speech errors may or may not result in a communication breakdown, and corrective feedback usually refers to language errors that can seriously impede mutual understanding. In the case of written communication, corrections are not only related to linguistic form, but also to content, discourse and organizational aspects of language use (Ellis et al., 2020).

The main purpose of this paper is to provide some crucial information on the nature of corrective feedback, taking into account the fact that research has mostly studied this phenomenon using English as a foreign language. Therefore, the present research study aims at finding out the attitudes of future German language teachers about the following issues: i) the effectiveness of immediate and/or delayed teacher and peer corrective feedback, ii) types and number of errors to be corrected, iii) teacher's feedback related to written assignments, iv) the effectiveness of implicit and explicit error correction and v) influence on motivation. Thus, the goal is to achieve a better understanding of the different aspects of corrective feedback, especially in the context of learning German as L2.

## 2. Corrective feedback

The decision of correcting or not correcting the occurring errors in output production is an interesting phenomenon in applied linguistics research. In the late 1990s, two publications —Lyster and Ranta (1997) and Truscott (1996)— marked the beginning of the development of the phenomenon of corrective feedback, thus, the first detailed taxonomy was designed. Lyster and Ranta (1997) introduced a term, ‘uptake,’ referring to the learner’s reaction after receiving corrective feedback. In other words, uptake is a product of perception and an intermediate stage in the acquisition process. The learner shows that he/she noticed the corrective feedback and stored it in short-term memory. This reaction after an error is particularly important because it can point to deep cognitive processing, which, consequently, accelerates the development of the learner’s interlanguage. Thus, the reaction after an error can result in the production of a new linguistic form, which, according to Swain (2005), indicates the following: i) the learner recognizes certain deficiencies in his/her knowledge and pays conscious attention and ii) the learner will reach for previously acquired knowledge (metalinguistic knowledge).

The discussion about the effectiveness of corrective feedback began when John Truscott (1996) questioned the purpose of error correction, pointing out that grammar correction had no place in teaching writing. According to Truscott, correcting errors is not purposeful, because mastering a language involves a specific acquisition order of individual structures, and therefore, correcting errors will not significantly affect the development of those language structures for which the learner is not developmentally ready. Truscott (1996) emphasized that corrective feedback is ineffective, even harmful, in the language acquisition process. Namely, according to Truscott, corrective information would only affect the learner’s explicit knowledge, which is not functional in authentic communication situations. Therefore, corrective feedback should focus exclusively on meaning, not on linguistic accuracy.

On the other hand, Long (2015) advocated a completely different approach —he believed that error correction is crucial in the development of a learner’s interlanguage. Emphasis on the accuracy of the linguistic form is highly recommended, especially with those forms that do not contribute to the semantic information of the lexical unit (for example, the suffix *-s* for the third person singular in English).

Lyster and Ranta (1997) established two basic taxonomies: explicit and implicit. The former are those in which the error is overtly pointed out, as opposed to implicit strategies, in which the teacher prompts for decoding errors. In other words, they can be implicit (the corrective power is not obvious) and explicit (learners are aware of the correction) (Ellis et al., 2020). Implicit corrections are indirect evidence of in-depth language processing and self-monitoring (Kovač, 2020; Levelt, 1989).

Other authors proposed similar taxonomies. For example, Li and Vuono (2019) divided corrective feedback into implicit (reformulation, clarifying question and repetition) and explicit strategies (metalinguistic comment and explicit correction),

whereby explicit feedback displayed a greater immediate effect, but shorter long-term effects, in contrast to implicit feedback, which resulted in longer lasting effects.

Ellis and Shintani (2014), in a comprehensive study, summarised in detail the research studies, finding out the similarities and differences between the correcting strategies. They concluded that explicit correction is one feedback strategy, unlike prompting an error. Thus, implicit feedback includes as many as four subcategories - request for clarification, repeating the error, elicitation and metalinguistic comment. Therefore, using multiple strategies is more effective than relying on an isolated one. In other words, signalling an error or prompting for correction involves a combination of explicit and implicit strategies, leading to an expectation of more frequent representation of this corrective feedback than explicit feedback alone. When it comes to teachers' practice of using implicit or explicit feedback, the body of research has yielded different conclusions. Choi and Li (2012), as well as Li and Vuono (2019) concluded that teachers generally prefer recasts or explicit feedback. Lyster and Saito (2010) and Brown (2016) pointed out that teachers are more likely to use explicit corrective feedback with elementary school students and adults, while implicit feedback is a more preferred strategy with high school students. The author explained the findings in the following way – explicit correction is a simpler and safer strategy in teaching children and is often used by teachers to avoid incorrect self-corrections. Implicit feedback, on the other hand, is deemed effective for high school students, focusing attention on errors and promoting self-correction and self-monitoring. Also, Brown (2016) concluded that more experienced teachers tend to prefer implicit, such as recasts, over explicit feedback. Other studies (Li & Vuono, 2019) showed that the learner will notice the correction if it is related to shorter speech segments and if a smaller number of such corrections are accompanied by a rising intonation. The authors also pointed out that corrective feedback positively affects implicit learning. However, referring to previous studies, Li and Vuono (2019) found out that most self-corrections (after the teacher's explicit correction of the linguistic form) are just routine repetitions, and that only the inclusion of the corrected form in a new context will result in qualitative changes of the learner's interlanguage.

Veliz (2007) found that even though the recasts are the most widely used type of error feedback in the classroom, they tend to be ineffective. Students are often not aware that they are being corrected, believing that the teacher is just interacting with them. However, more advanced students are more capable of perceiving the correction in the form of a recast. Finally, the author pointed out that teachers need to be more consistent and systematic when correcting faulty utterances. Lyster et al. (2013) in their extensive review study concluded that oral correction is significantly more effective than no correction and also revealed a tendency for learners receiving prompts or implicit correction to demonstrate more gains on some measures than learners receiving recasts. It is interesting to note that recasts were more effective for improving pronunciation of familiar, but not unfamiliar items (Saito & Lyster, 2012). Recasts were as effective as

prompts for young ESL learners with high pretest score but less effective than prompts for young learners with low pretest scores (Ammar & Spada, 2006).

Alsolami (2019) pointed out that adequate feedback is crucial to enhance the overall learning process and to improve the pronunciation.

Shen and Chong (2023) compiled a qualitative research synthesis using a perception-based framework. The study was based on 14 articles, dealing with factors affecting how learners engage with written corrective feedback. There are three kinds of learner engagement with written corrective feedback: affective, behavioural and cognitive. The affective perspective focuses on learners' emotions, values and attitudes towards written corrective feedback. The behavioural perspective is examined based on learners' uptake or revisions prompted by written corrective feedback. The cognitive perspective can be examined by looking at students' awareness of error patterns and learners' attitudes towards written corrective feedback. Learners' perception of corrective feedback will also affect its effectiveness and overall motivation (Li, 2017). In this regard, Loewen et al. (2009) noted very positive attitudes among students towards corrective feedback, while among teachers, the attitudes depended on teaching experience. Teachers with less teaching experience expressed more doubts about the influence of corrective feedback on language learning. Nevertheless, research on attitudes and beliefs related to corrective feedback is generally positive (Li & Vuono, 2019). Li and Vuono (2019) summarized their conclusions as follows: a) teachers and students show a very positive attitude towards written feedback; b) students prefer direct feedback – the teacher identifies the error, explains the nature of the error and offers the correct form; c) the absolute majority of students want to receive feedback or explicit correction; d) students do not prefer prompting an error without explanation and correction; e) students at a higher level of language proficiency want the teacher to correct content errors, then linguistic (grammatical) and discourse errors; f) among learners at a lower level of language proficiency, opinions are divided; g) teachers and learners point out that written assignments should be corrected; h) the absolute majority of learners state that they will read the teacher's comments and correct the errors.

In another review study by Ellis and Shintani (2014, p. 257) the following conclusions were outlined: a) it is necessary to correct the learners, however, the teacher has to carefully monitor the process in order to avoid negative impact on motivation; b) if teaching emphasizes the communicative aspect and the development of fluent speech, immediate corrective feedback could have negative consequences on fluent speech and/or motivation; c) in the case of written corrective feedback, students should redo their written works in accordance with the teacher's corrections, and the teacher should review the tasks again; d) excessive correction is not desirable and teachers should selectively correct errors, depending on their importance; e) regardless of the existence of a whole range of possible strategies for correcting spoken and/or written errors, teachers generally show a tendency to motivate students to self-correct; f) whenever possible, students should correct each other.

### 3. Methodological procedures

Twenty-three graduate students of German studies took part in this research study, who encountered the concept of corrective feedback within the course *Development of speech fluency* in the academic year 2022/23 and 2023/24. During the lectures the students learned about the different strategies used for error correction. For twelve consecutive weeks, over one academic semester, the students performed speaking tasks that were organized according to the principles of task-based teaching. Each week students performed different tasks – role plays, argumentation tasks, retelling stories, and problem-solving tasks. Each task was performed twice. In the first performance of the task the students were given comprehensible input and instructions from the teacher. They were also allowed to use the dictionary. Also, a list of useful formulaic phrases was at their disposal which helped them to produce more coherent utterances. While performing the task for the first time, the students corrected each other using a strategy of their preference. During the repeated task performance, the teacher provided explicit and implicit corrective feedback. Afterwards, the students discussed their performance, comparing the first and second performance in terms of the frequency of used fillers, the occurrence of different errors, speech rate, as well as the usefulness of teacher and peer corrective feedback. Each week two different tasks were performed. While correcting, the teacher used the following strategies (Lyster & Ranta, 1997):

1. *Recast* – the teacher replaced the erroneous form with the correct one, and the rest of the sentence remained unchanged.
2. *Explicit correction* – consisted of a comment by which the teacher informed the student about the error and then corrected it. The correction was preceded by the word *Nein*.
3. *Metalinguistic comment* – the teacher explained the nature of the error using metalanguage, but did not correct the error.
4. *Elicitation* or an attempt to encourage the student to self-correct. This attempt was usually preceded by a rising intonation with which the teacher indicated an error (pronunciation, lexical, syntactic or pragmatic).
5. *Error repetition with rising intonation* – the teacher repeated the erroneous linguistic form, expecting the student to self-correct.
6. *Clarification question*. The teacher asked the student a question to clarify the used language structure.

At the end of the semester, students filled out a questionnaire with the aim of obtaining information on the expediency of providing corrective feedback and on the possible effects on language learning/acquisition, as well as on motivation. The questionnaire consisted of 28 statements that were divided according to several subcategories – immediate/delayed corrective feedback, type of errors, implicit or explicit correction, type of errors that ought to be corrected, and teacher's feedback on written assignments.



The participants expressed their level of agreement with a certain statement by choosing answers on a 5-degree Likert scale (5 = strongly agree, 4 = agree, 3 = neither agree nor disagree, 2 = disagree, 1 = strongly disagree). A Likert scale assumes that the strength/intensity of an attitude is linear, i.e., on a continuum from strongly agree to strongly disagree, and makes the assumption that attitudes can be measured. Attitude was determined based on the argument that a mean score of 3 in Likert scale represents neutral attitude, mean score of less than 3 represents negative attitude and greater than 3 represents a positive attitude. The range of interpreting the Likert scale mean score was given as follows: 1.0-2.4 (Negative attitude), 2.5-3.4 (Neutral attitude), and 3.5-5.0 (Positive attitude) (McLeod, 2019). The items can be seen in Table 1. As can be seen in table 1, the bold titles of individual statements are the basis for mutual comparison.

The selected statements were based on the analysis of research interests in the field of corrective feedback studies. Thus, in this research the affective perspective of learner engagement with corrective feedback was analysed. The compared statements include student attitudes towards the essential determinants of providing corrective feedback – timing of correction (immediate/delayed); corrector factor – teacher/student; tendency to reflect on grammatical and pronunciation errors; correction of simple/complex errors in two modalities – speaking and writing; implicit/explicit corrective feedback; the amount of errors to be corrected (selected/all) in two modalities – speaking and writing (Brown, 2016; Li & Vuono, 2019; Lyster & Saito, 2010).

The rationale for the comparison is the examination of:

- a. the existence of significant differences between the statements related to the correction timing (immediate/delayed) depending on the corrector factor – teacher/peer in the case of pronunciation errors (1 – 2, 3 – 4);
- b. the existence of significant differences between the statements related to the correction timing (immediate/delayed) depending on the corrector factor – teacher/peer in the case of grammatical errors (6 – 7, 8 – 9);
- c. the existence of a significant difference with regard to the tendency to reflect on pronunciation/grammatical errors (5 – 10). In other words, do students reflect more on pronunciation or grammatical errors?
- d. the existence of significant differences with regard to the type of error – complex/simple – considering the two modalities – speaking/writing (12 – 13, 14 – 15);
- e. the existence of a significant difference between the two strategies of implicit correction – elicitation/error repetition with rising intonation (17 – 18);
- f. the existence of a significant difference between implicit correction (elicitation) and explicit corrective feedback (recast) (17 – 22);
- g. the existence of a significant difference between two categories of explicit correction (recast/explicit corrective feedback with metalinguistic comment (22 – 23);

- h. the existence of a significant difference between the amount of errors the teacher should correct in speaking (only those being essential for understanding the message/all errors) (25 – 26);
- i. the existence of a significant difference between the amount of errors the teacher should correct in writing (only those being essential for understanding the message/all errors) (27 – 28).

TABLE 1. STATEMENTS TO BE JUDGED ALONG A FIVE-POINT LIKERT SCALE.

Statement No.	Statement
1	Immediate corrective feedback (teacher) – I prefer when the teacher corrects my pronunciation errors – interrupts and corrects immediately.
2	Delayed corrective feedback (teacher) – I prefer when the teacher does not correct my pronunciation errors right away, but after the end of the lesson.
3	Immediate corrective feedback (peer) – I prefer when my peer colleague corrects my pronunciation errors immediately when they occur.
4.	Delayed corrective feedback (peer) – I prefer when a peer colleague does not correct my pronunciation errors right away, but after the end of the lesson, in private.
5.	Self-reflection – I reflect on the corrected pronunciation error and sometimes repeat it silently.
6.	Immediate corrective feedback (teacher) – I prefer when the teacher corrects my grammatical errors immediately.
7.	Delayed corrective feedback (teacher) – I prefer when the teacher does not correct my grammatical errors right away, but later, in private.
8.	Immediate corrective feedback (peer) – I prefer when my peer colleague corrects my grammatical errors right away.
9.	Delayed corrective feedback (peer) – I prefer when my peer colleague does not correct the grammatical errors right away, but later, in private.
10.	Self-reflection – I reflect on the corrected grammatical error and sometimes repeat it silently.
11.	Error correction has a negative effect on motivation (causes anxiety and fear of speaking).
12.	Correction of more complex errors in speaking – The teacher's corrective feedback of more complex structures in spoken performance (e. g. word order in German dependent clauses) is effective for successful acquisition/learning.
13.	Correction of simple errors in speaking – Correcting simple errors in spoken performance (e. g. verb ending) is effective for successful acquisition/learning.
14.	Correction of more complex errors in writing –Corrective feedback on more complex errors in written performance (e. g. word order in German dependant clauses) is effective for successful acquisition/learning.

15.	Correction of simple errors in writing – Correcting simple errors in written performance (e. g. verb ending) is effective for successful acquisition/learning.
16.	Implicit feedback (elicitation) – I prefer when the teacher encourages self-correction by signalling non-verbally the error.
17.	Implicit corrective feedback (elicitation): I prefer when the teacher points out an error with rising intonation and expects me to self-correct.
18.	Error repetition with rising intonation: I prefer when the teacher repeats the erroneous form expecting me to self-correct.
19.	I have noticed that self-corrections lead to the acquisition of target language structures.
20.	Effectiveness of implicit corrective feedback – It is only effective for less complex structures.
21.	Effectiveness of implicit corrective feedback – It is only effective for pronunciation errors.
22.	Explicit corrective feedback (recast) – I prefer when the teacher corrects the error and I repeat it afterwards.
23.	Explicit corrective feedback with metalinguistic comment – I prefer when the teacher corrects the error and explains the rule using metalanguage.
24.	Routine repetition – When I repeat the correct form, it is usually routine repetition, not deep processing (i. e. most often I just repeat it, without thinking about the nature of the error).
25.	Number of errors to be corrected in speaking – I think that not all errors should be corrected in spoken performance, but only those errors that are essential for the understanding of the message.
26.	Number of errors to be corrected in speaking – I think that all errors in spoken performance should be corrected.
27.	Number of errors to be corrected in writing – I prefer when the teacher thoroughly corrects all errors in written assignments and provides metalinguistic explanation.
28.	Number of errors to be corrected in writing – I prefer when the teacher corrects only those errors in written assignments that are necessary for the understanding of the content.

#### 4. Results and discussion

Research results are based on the group of 23 graduate students of German studies who expressed their attitudes towards corrective feedback. Table 2 illustrates arithmetic means and standard deviations for each of the statements from the questionnaire.

TABLE 2. ARITHMETIC MEANS AND STANDARD DEVIATIONS OF THE STATEMENTS

Statement number	Arithmetic mean	Standard deviation
1.	2,913	1,041
2.	3,826	1,302
3.	2,304	1,105
4.	3,435	1,409
5.	4,000	1,087
6.	2,870	1,217
7.	3,609	1,234
8.	2,522	1,123
9.	3,261	1,453
10.	4,174	0,834
11.	2,522	1,123
12.	4,174	0,778
13.	4,174	0,717
14.	4,478	0,730
15.	4,391	0,499
16.	3,696	1,146
17.	3,870	1,058
18.	3,130	1,140
19.	3,870	0,869
20.	3,174	1,072
21.	3,739	0,964
22.	4,044	0,638
23.	3,783	0,998
24.	2,609	1,158
25.	2,435	0,788
26.	3,609	0,891
27.	4,174	0,984
28.	2,087	0,900

Considering that the Likert scale is of ordinal type, and two dependent samples are compared, we used the non-parametric Wilcoxon Signed-rank test to investigate significant differences between the compared statements. The results are presented in table 3 (level of significance 0.05).

TABLE 3. WILCOXON SIGNED-RANK TEST RESULTS

Wilcoxon Signed-rank test	Sum of positive ranks	Sum of negative ranks	p	Statistically significant difference
Statement comparisons				
1 - 2	120	- 33	0,035	Yes
3 - 4	139	- 32	0,019	Yes
6 - 7	130,5	- 40,5	0,051	No
8 - 9	91,5	- 28,5	0,070	No
5 - 10	4,5	- 1,5	0,750	No
12 - 13	15	- 13	1,000	No
14 - 15	5	- 10	0,750	No
17 - 18	12,5	- 78,5	0,019	Yes
17 - 22	64,5	- 40,5	0,465	No
22 - 23	13,5	- 41,5	0,186	No
25 - 26	165,5	- 24,5	0,002	Yes
27 - 28	206	- 4	< 0,001	Yes

There are in general two strategies for corrective feedback depending on the timing of the correction. The first strategy implies an immediate intervention and ought to be practiced when a learner commits an error. The second is the delayed corrective feedback, where the correction is usually practiced after the end of the lesson. As indicated in table 2, the students prefer delayed corrective feedback of pronunciation and grammatical errors (statements 2 and 7). Similar conclusions were reported by Ellis and Shintani (2014) in a review paper on various aspects of providing corrective feedback, stating that if teaching emphasizes the communicative aspect and the development of fluent speech, immediate corrective feedback could have negative consequences on fluent speech and/or motivation. As it can be seen in table 3, it is evident that the delayed correction of pronunciation errors is significantly more preferred than immediate correction (statements 1 – 2). The findings of the current study are consistent with the results reported in Şakiroğlu (2020). Her results revealed that 90 % of the learners would like to be corrected during the process of speaking English. However, the majority of the students indicated the preference to be corrected after finishing turn with nice and friendly manners. On the other hand, in several studies, students were found to favour immediate over delayed correction in pronunciation (Alghazo, 2015). Immediate correction will positively affect the development of procedural knowledge,

while delayed correction will encourage metalinguistic understanding, that is, students will be encouraged to think about the error and the correction (Ellis et al., 2020).

Regarding immediate and delayed correction by peer colleagues, the results also point to the conclusion that delayed corrective feedback of pronunciation errors is significantly more preferred than immediate (table 3, statements 3 – 4). As far as research on peer evaluation is concerned, Saito and Lyster (2012) wanted to examine corrective feedback as a way of enhancing L2 development. It was the first study to examine its effectiveness on utterance fluency. The authors conducted this research study in university-level required English classes over one academic semester. Corrective feedback provided by peers has been shown to be a feasible option for helping learners to attend to form in effective ways.

As for the implicit corrective feedback, students prefer when the teacher non-verbally signals an error. They also believe that implicit corrective feedback is more efficient for pronunciation errors. A significant difference has been obtained for the two strategies of implicit feedback. Namely, elicitation or an attempt to encourage the student to self-correct (preceded by a rising intonation) is significantly more preferred than error repetition (the teacher repeated the erroneous linguistic form, expecting the student to self-correct) (table 3, statements 17 – 18).

Regarding explicit corrective feedback, students prefer when the teacher provides the students with the correct form, which is followed by the student's repetition of the correct form. They also support the use of metalinguistic explanation for error correction (table 2, statements 22 and 23). Several studies confirmed the effectiveness of explicit corrective feedback for learners at a higher level of language proficiency, which was not the case for less advanced students (Lyster et al., 2013; Ammar & Spada, 2006). According to Ellis et al. (2020), explicit correction will result in the activation of cognitive mechanisms, where repetition and restructuring will have a positive effect on interlanguage development. The authors' interpretation is in accordance with Krashen's theory of language acquisition (1985), according to which explicit teaching is effective for the acquisition of simpler language structures, while more complex structures can be acquired only implicitly. However, the concept of structural complexity and the way to master the target structures is relatively unknown (Ellis et al., 2020). For example, the ending -s for the third person is a structurally simple form, yet, it is adopted at later stages of the learner's interlanguage development. Nevertheless, it should be pointed out that the students in this research are at a higher level of command of the German language, and therefore it would be interesting to examine whether the conclusions are valid for less proficient students.

It is interesting to note that students agree with the statement that while repeating the correct form, they usually just repeat it, without thinking about the nature of the correction (table 2, statement 24). This is in line with the results suggested by Li and Vuono (2019), who found out that most self-corrections (after the teacher's explicit correction of the linguistic form) are just routine repetitions, and that only the inclusion

of the corrected form in a new context will result in qualitative changes of the learner's interlanguage. However, the students in this research study tend to later reflect on the corrected pronunciation and grammatical errors (table 2, statements 5 and 10). This indicates a higher level of awareness of the need for self-monitoring, which is involved at almost every level of language processing (Kovač, 2020).

The students agree that the teacher's corrective feedback of simple and more complex structures, both in spoken and written performance, is effective for successful language learning (table 2, statements 12, 13, 14, 15, 16). Furthermore, the students want the teacher to correct all pronunciation errors (table 2, statement 26). Also, they do not agree that only those errors which are essential for understanding the message should be corrected (table 2, statements 25 and 28). These attitudes are in contrast to those expressed in Ellis and Shintani (2014), namely, the learners in the compared studies believe that excessive correction is not desirable and teachers should selectively correct errors, depending on their importance. In order to further confirm this claim, a corresponding statistical test was conducted. There is a significant difference in students' views related to error corrections, that is, a significantly higher level of agreement has been obtained for the statement that teachers should correct all errors in spoken performance, compared to the statement that only those errors that are essential for understanding the message should be corrected (table 3, statements 25 – 26). The obtained significant difference can be explained by the level of command of the German language. Li and Vuono (2019) found that students at a higher level of language proficiency want the teacher to correct content errors, then linguistic (grammatical) and discourse errors.

The same conclusions apply to the written modality. In the context of correcting errors in written performance, a significantly higher level of agreement has been recorded for the statement *I prefer when the teacher thoroughly corrects all errors in written assignments and provides metalinguistic explanation* compared to *I prefer when the teacher corrects only those errors in written assignments that are necessary for the understanding of the content*, table 3, statements 27 – 28). This was partially confirmed in the research conducted by Van Beunigen et al. (2008). The authors pointed out the superiority of focused corrective feedback, that is, the selective correction of target language structures (not all errors). The effectiveness of corrective feedback will also depend on the learner's willingness to respond positively to such feedback. Their results showed that corrective feedback can be effective in improving students' accuracy.

An interesting result of this study is that no significant differences have been recorded between the correction strategies – implicit and explicit (table 3, statements 17 and 22). In other words, the statistical analysis has not proven the superiority of the students' preference of the implicit corrective feedback, which was emphasized as more efficient by Ellis and Shintani (2014), as well as Lyster (2004). Those authors explained that this correction strategy will encourage students to make self-corrections, which are, in turn, an indirect evidence of the acquisition of the target form, as a consequence of deep language processing and enhanced monitoring.

The students of this study emphasize that the correction of both simple and more complex structures in spoken and written performance leads to successful acquisition of the target forms. Based on the analysis of previous research, Ellis et al. (2020) concluded that explicit teaching (and corrective feedback) is effective primarily for simple language structures (e. g. the definite article in English, irregular forms of the past tense, and the ending *-s* for the plural of nouns), but that the same conclusions do not apply to more complex language forms (e. g. indefinite article, ending *-ed* for past tense, relative clauses, passive and ending *-s* for third person). However, as it has already been mentioned previously, the division into simpler and more complex structures is relatively arbitrary, and the developmental process of acquiring grammatical forms is still unknown, especially for other foreign languages, such as German.

## 5. Conclusion

One way to improve accuracy in oral and written communication is the constant corrective feedback. In this research, the affective perspective of learner engagement with corrective feedback was analysed. The analysis of the obtained results displayed that the students agree on the effectiveness of teacher's corrective feedback of simple, as well as more complex structures in spoken and written performance (e. g. word order in German dependent clauses). A significant difference in students' attitudes was confirmed by comparing two strategies of implicit feedback. Namely, elicitation or an attempt to encourage the student to self-correct (preceded by a rising intonation) was significantly more preferred than error repetition (the teacher repeated the erroneous linguistic form, expecting the student to self-correct). Furthermore, there was a significantly higher degree of agreement with the statement that teachers should correct all errors, compared to the statement that only errors which are essential for the understanding of the message should be corrected, both in written as well as in spoken performance. With respect to students' views on correcting errors in written performance, students agree that correcting simple errors in written performance is effective for successful language learning, and that written assignments should be thoroughly corrected by means of metalinguistic explanation. Thus, the research pointed to overall positive attitudes regarding corrective feedback in spoken and written performance.

Further research, targeting the effectiveness of various corrective feedback strategies and corrector factor (teacher versus peer colleague), would be an interesting research field which might help the teachers develop more contemporary teaching strategies.

## References

Alghazo, Sharif M. (2015). Advanced EFL Learners' Beliefs about Pronunciation Teaching, *International Education Studies*, 8(11), 63-76.



- Alsolami, Radiah (2019). Effect of Oral Corrective Feedback on Language Skills. *Theory and Practice in Language Studies*, 9(6), 672-677.
- Ammar, Ahlem & Spada, Nina (2006). One size fits all? Recasts, Prompts, and L2 Learning. *Studies in Second Language Acquisition*, 28, 543-574.
- Brown, Dan (2016). The type and linguistic foci of oral corrective feedback in the L2 classroom: a meta-analysis. *Language Teaching Research*, 20(4), 436-458.
- Choi, Seung-Yi & Li, Shaofeng (2012). Corrective feedback and learner uptake in a child ESOL classroom. *RELC Journal*, 43 (3), 331-351.
- Ellis, Rod & Shintani, Natsuko (2014). *Exploring Language Pedagogy through Second Language Acquisition Research*. Routledge.
- Ellis, Rod; Skehan, Peter; Li, Shaofeng; Shintani, Natsuko & Lambert, Craig (2020). *Task-based language teaching: Theory and practice*. Cambridge University Press.
- Kovač, Mirjana M. (2020). *Speech fluency in a foreign language*. Split: Faculty of Humanities and Social Science in Split.
- Krashen, Stephen (1985). *The input hypothesis: Issues and implications*. New York: Longman.
- Levelt, Willem J. M. (1989). *Speaking: from intention to articulation*. MA: MIT.
- Li, Shaofeng (2017). Teacher and learner beliefs about corrective feedback. In Nassaji, Hossein & Kartchava, Eva (Eds.), *Corrective Feedback in Second Language Teaching and Learning*, (pp. 143-157). Routledge.
- Li, Shaofeng & Vuono Alyssa (2019). Twenty-Five Years of Research on Oral and Written Corrective Feedback in System. *System*, 84, 93-109.
- Loewen, Shawn, Erlam, Rosemary & Ellis, Rod (2009). The incidental acquisition of 3<sup>rd</sup> person- s as implicit and explicit knowledge. In Ellis, Rod; Loewen, Shawn; Elder, Chaterine; Erlam, Rosemary; Philp, Jenefer & Reinders, Hayo (Eds.), *Implicit and Explicit Knowledge in Second Language Learning, Testing and Teaching* (pp. 262-280). Multilingual Matters.
- Long, Mike (2015). *Second Language Acquisition and Task-Based Language Teaching*. Wiley-Blackwell.
- Lyster, Roy (2004). Differential effects of prompts and recasts in form-focused instruction. *Studies in Second Language Acquisition*, 26, 399-432.
- Lyster, Roy & Ranta, Leila (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19, 37-66.
- Lyster, Roy & Saito, Kazuya (2010). Oral feedback in classroom SLA: A meta-analysis. *Studies in Second Language Acquisition*, 32, 265-302.
- McLeod, Saul (2019). Sampling Methods. *Simply Psychology*. Retrieved from <https://www.simplypsychology.org/sampling.html>
- Saito, Kazuya & Lyster, Roy (2012). Effects of Form-Focused Instruction and Corrective Feedback on L2 Pronunciation Development of Japanese Learners of English. *Language Learning*, 62(2), 595-633.

- Şakiroğlu, Hülya Ü. (2020). Oral Corrective Feedback Preferences of University Students in English Communication Classes. *International Journal of Research in Education and Science*, 6(1), 172-178.
- Shen, Rongyi & Chong, Sin Wang (2023). Learner engagement with written corrective feedback in ESL and EFL contexts: a qualitative research synthesis using a perception-based framework. *Assessment & Evaluation in Higher Education*, 48(3), 279-290.
- Swain, Michael (2005). Legislation by hypothesis: The case of task-based instruction. *Applied Linguistics*, 26, 376-401.
- Truscott, John (1996). The case against grammar correction in L2 writing classes. *Language Learning*, 46, 327-369.
- Van Beuningen, Catherine, de Jong, Nivja & Kuiken, Folkert (2008). The Effect of Direct and Indirect Corrective Feedback on L2 Learners' Written Accuracy. *International Journal of Applied Linguistics*, 156 (3), 279-296.
- Veliz, Leonardo C. (2007). Corrective Feedback in Second Language Classrooms. *Literatura y lingüística*, 19, 283-292.

# Towards the characterization of WORK problem schemas in the DIAPASON Ontology

Ángel Felices Lago – Universidad de Granada  
afelices@ugr.es

Rebut / Received: 12-02-24

Acceptat / Accepted: 08-04-24

**Resum. Cap a la caracterització dels esquemes de problemes de TREBALL a l'ontologia DIAPASON.** Els sistemes que analitzen i interpreten els continguts generats pels usuaris poden extreure informació útil a Internet per millorar la qualitat de la vida dels ciutadans. Aquest article pretén aplicar un model ontològic (DIAPASON) basat en el paradigma dels sensors socials per facilitar l'avenç de les societats intel·ligents. L'objectiu principal d'aquesta ontologia és ajudar al mòdul de PLN en les tasques de categorització de temes i reconeixement de paraules clau quan s'analitzen textos curts en anglès procedents de plataformes de mitjans socials per dur a terme la detecció de problemes socials. Com a mostra, s'ha descrit i formalitzat un nombre representatiu d'esquemes de problemes pertinents al domini WORK. Aquest domini està inclòs en la dimensió ECONOMIA de la categoria SOCIAL (a diferència del descrit en l'ontologia com a FÍSICO). La plantilla que desenvolupa la descripció de cada problema inclou aspectes essencials per a la seva detecció, com la tipologia (categoria, dimensió, domini, tipus de problema i esquema de problema) i definició inspirat pel llenguatge d'interfície COREL.

**Paraules clau:** Construcció d'ontologia, sensors socials, modelització conceptual, problemes socials, processament de llenguatge natural

**Abstract. Towards the characterization of WORK problem schemas in the DIAPASON Ontology.** Systems that analyze and interpret user-generated content can extract actionable information from the Internet to improve citizens' quality of life. This article aims to apply an ontological model (DIAPASON) based on the social-sensing paradigm to support the advance of smart societies. The main goal of this ontology is to assist the NLP module in the tasks of topic categorization and keyword recognition when analyzing English short texts from social media platforms to perform social problems detection. As a sample, a representative number

of problem schemas belonging to the WORK domain have been described and formalized. This domain is included in the ECONOMY dimension of the SOCIAL category (as opposed to the one described in the ontology as PHYSICAL). The template that develops the description of each problem includes essential aspects for its computer detection, such as typology (category, dimension, domain, problem type and problem scheme) and definition inspired by the COREL interface language.

**Keywords:** Ontology building, social sensors, conceptual modeling, social problems, natural language processing

## 1. Introduction

Citizens convey their thoughts, opinions and emotions through text data, photos, videos and audios on various social and private topics in social networks (Facebook, Twitter(X), Instagram, LinkedIn, blogs, wikis or YouTube). This is what is known as user-generated digital content or user-generated content (UGC). Consequently, understanding the natural language with which we express ourselves in social media in terms of immediacy is one of the areas of interest not only for artificial intelligence (AI) (McStay, 2020), but also for natural language processing (NLP). In that sense, emotion detection (Jurafsky & Martin, 2021) or the classification of speech acts in connection with dialogues have been some of the most studied areas. However, as explained in Perrián-Pascual (2023), it is worth highlighting the current line of research that inspires the contribution we present here: the one related to the so-called social sensing or, more specifically, social-media crowdsensing (Li et al., 2017). This new approach takes advantage of the data provided by users from “social sensors” or what are known as agents that provide information about their environment through social networks. This is the context of the ALLEGRO project (*Adaptive multi-domain social-media sensing framework*)<sup>1</sup>, which consists of two modules (Data Analysis and Data Fusion) that make use of a multimodal data repository and a knowledge base (ontological model). In both modules, high-performance computing (HPC) is applied. Data analysis, for example, integrates components at various levels depending on the multimodal nature. In this study we confine ourselves exclusively to the text analysis module defined as DIAPASON (*Unified hybrid approach to microtext analysis in social media crowdsensing*), which is especially designed for the analysis of microtexts<sup>2</sup>.

---

1. To know more about the ALLEGRO project, go to the following website: <http://allegro.ucam.edu/>

2. The construction of the DIAPASON ontology is based on METHONTOLOGY (Fernández López et al., 1997) guidelines for conducting the entire development process. Moreover, the ontology-engineering paradigm of ODPs from Hitzler et al. (2016) was adopted.

Therefore, the main purpose of this contribution is, on the one hand, to apply the ontological model designed by Perrián-Pascual (2023) for the development of DIAPASON, by means of the practical demonstration of the categorization of a series of social problems perceived by citizens with considerable level of emotionality and, on the other, to show how a representative sample of problem schemas belonging to the WORK domain has been described and elaborated. That domain is integrated into the ECONOMY dimension of the SOCIAL category (as opposed to the one described in the ontology as PHYSICAL). The template that develops the description of each problem includes essential aspects for its computer detection such as the typology (i.e. problem realm, problem dimension, problem domain, and problem type) and a definition of the problem in the form of a problem schema, whose syntax is inspired by the interface language known as COREL (Perrián-Pascual & Mairal-Usón, 2010).<sup>3</sup>

Consequently, the structure of the article is as follows: section two briefly describes the foundations on which social problems are categorized, particularly the problems related to the ECONOMY dimension (where the WORK domain is eventually integrated). Section three describes the DIAPASON ontology and the description of its formalization schemas. Section four describes the modeling of the WORK domain and, at the same time, presents the specific WORK problems that have been detected revising social networks like *Twitter* among other references. Then a detailed analysis of two specific problems schemas is provided: JOBDISCRIMINATION and NOBENEFITS. Finally, a series of conclusions are listed.

## 2. Background

It is a well-known fact that social and economic issues are frequently interrelated, and even more so when the microtexts from tweets in *Twitter* (now *X*) provide countless examples. In fact, in the ALLEGRO project previously mentioned, the analysis of a considerable number of tweets referring to social issues have undoubtedly an economic origin, or at least this is the opinion of many of the sources consulted. A limitation for this approach is the fact that examples from other social networks such as *Facebook* or *Instagram* could also be informative and have not been included in this study. In fact, this is a pending task to be accomplished in further research once the searcher tools are refined.

### 2.1. Social problems

There are two main ways of characterizing social problems. We can identify what a social problem is from an objectivist approach, i.e., social problems are identified

---

3. It should be noted that the COREL notational system is greatly influenced by the computational potential of the Role and Reference Grammar (RRG) model created by Van Valin and LaPolla (1997), which facilitates the development of NLP from the perspective of deep semantics.

with conditions that are detrimental to society (Davis, 1937), or from a subjectivist approach, i.e., social problems are identified with conditions that are perceived as problems (Best, 1989; Holstein & Miller, 2003; Spector & Kitsuse, 1977). Although the objectivist approach continues to dominate sociology textbooks (DeFronzo & Gill, 2020), the alternative paradigm is increasingly influential in academic literature. This is due to various reasons. As explained by Best (2017), objectivists often do not specify what constitutes a harmful perception; instead, they present that kind of perception as something so vague that it often becomes meaningless. However, the subjective approach focuses not only on the reality of the problem's existence, but also on the level of the public concern aroused by the same issue. Consequently, the focus of interest is the social process by means of which people identify social problems, that is, what in sociology is called "social construction" (Heiner, 2002). In other words, Best (1995, p. 45) concluded that "social problems are what people view as social problems [...] No condition is a social problem until someone considers it a social problem", which represents the subjectivist approach to social problems.

As a result of this constructionist context, the DIAPASON approach (Perián-Pascual, 2023) shares some theoretical postulates of the model for studying social problems proposed by Best (2017) and defined in the first phase as claims-making (the process of directing the audience attention to the problem through a complaint). In this sense, three factors play a key role in this process:

- (a) complaints, i.e., statements about social problems that need to be solved,
- (b) the producers of complaints, i.e., the people who formulate the complaints,
- (c) the problem conditions, i.e., the conditions that become the subject of the complaints.

The foundations of social problems are rooted in various factors. Charon and Vigilant (2005), Mooney et al. (2021), among others, coincide in listing the categories of social problems as originating not only in the areas of the economy, education, political matters, health care, family circumstances, violence and crime, injustice or discrimination, but also in the areas of technology, environment, population, globalization, individual or group behaviors.

The objective and subjective perception of our quality of life has been another crucial factor for the construction of the DIAPASON ontology related to social problems in the LIVING dimension and these problems were primarily determined by considering Quality of Life reports (Bundesregierung, 2018; Eurostat, 2017; Helliwell et al., 2020; OECD (Organization for Economic Co-operation and Development), 2020; Office for National Statistics, 2019). Sociological encyclopedias and textbooks (Eitzen et al., 2014; Parrillo, 2008; Seccombe & Kornblum, 2020) were also examined to explore this dimension. Eventually, eleven domains have been identified: CRIMEJUSTICE, CULTURE, EDUCATIONTRAINING, HEALTH, INEQUALITY, LEISURE, POPULATION, RESEARCHINNOVATION, SOCIALRELATIONSHIP, SUBJECTIVEWELLBEING

and **TERRORISMWAR**. As a result, the **DIAPASON** ontology has been modeled<sup>4</sup>. However, even if there is a considerable amount of consensus about the inventory of social problem domains in this dimension, it is necessary to explain the decisions taken by Perrián-Pascual (2023) when this part of the **DIAPASON** ontology was modelled: He considers that economy, politics, and the natural environment are multidimensional concepts, so they are treated as separate problem dimensions rather than as basic problem domains in the **LIVING** dimension. **ECONOMY** and **GOVERNANCE** are regarded as two different social dimensions, whereas **ENVIRONMENT** is considered as a problem dimension in the **PHYSICAL** realm of the Smart City framework. It should be noted that the economic dimension, for example, cannot be completely separated from domains of other dimensions, such as housing, health, environment or inequality.

## **2.2. The *ECONOMY* dimension**

In Eurostat 2017, the directors of Social Statistics Panel promoted the *Quality of Life* Expert Group, whose main task was to draw up a list of quality of life indicators. Among the nine domains covered by the report, there are some that affect the **ECONOMY** dimension integrated in **DIAPASON**, especially those called “material living conditions”, “production” and “economic security”. These domains cover a series of themes and indicators to refer to them. Some are objective (e.g. long-term unemployment rate) and others are subjective (e.g. level of job satisfaction) since, while the former are objectifiable, the latter also contribute to assessing quality of life. In this context, it is worth mentioning the contribution made by the OECD in its biannual report *How's Life?* (2020), where the conditions for improvement in the quality of life are projected. The **ECONOMY** dimension is linked to satisfaction with a variety of economic concerns: personal and family financial situation, perceived income adequacy, material possessions and the extent to which physical needs are met (i.e., standard

---

4. For a better understanding of the taxonomic hierarchy and the four levels of thematic granularity of the **DIAPASON** ontology, the following concepts defined by Perrián-Pascual (2024b) should be taken into consideration: “The upper level is structured into two problem realms: the social realm, which subsumes the **LIVING**, **ECONOMY** and **GOVERNANCE** problem dimensions, and the physical realm, which subsumes the **MOBILITY**, **INFRASTRUCTURE** and **ENVIRONMENT** problem dimensions. The modelling of the dimension level primarily results from Smart City frameworks (Giffinger et al., 2007; Govada et al., 2017; Appio et al., 2019). At the next level, each problem dimension is organised in problem domains (e.g. **COMMUNITYFACILITY**, **CRIMEJUSTICE**, **ECOLOGICAL HAZARD**, **HEALTH**, **MOTORIZEDVEHICULARTRAFFIC**, and **POPULATION**, among others) on which citizens can show an attitude of disapproval towards some specific aspect of the community. Finally, the lower level describes particular types of community problems that can affect some, most or all citizens for each problem domain (e.g. **BEACHCROWDING**, **BEACHPARKING**, **BEACHSECURITY**, **MARINESPILL**, and **SEAWATERTRANSPARENCY**, among others).”

of living), feelings of financial security or access to necessary goods and services or aspirations and attainment of material goods (Sirgy, 2018).

According to Eurostat (2017), there are two types of domains in this dimension: (i) those related to the material conditions of life, such as CONSUMPTION (e.g. personal expenditure), INCOME (e.g. purchasing power) and WEALTH (e.g. financial assets), and (ii) those related to productive activities like WORK (e.g., stability, over-qualification, unemployment, and work-life balance), all of which focus on what economic science defines as microeconomics. However, MACROECONOMICS (e.g., inflation, investment rate, national debt, and public expenditure) does not appear in this classification, even if other previously mentioned studies (Bundesregierung, 2018; Office for National Statistics, 2019) point out how the basic macroeconomic indicators are factors that, depending on their level or amount, affect the economy of countries and, eventually, affect their citizens' quality of life. In consequence, macroeconomic indices are a necessary dimension to be added to the other four in DIAPASON.

Despite the importance of the macroeconomic domain, in this study we have decided to choose the WORK domain for being a central microeconomic factor that is closely connected with various economic challenges, including income inequality, poverty, health, social security, taxation, informal economy, social security or retirement policies. The way in which this domain has been modeled in the DIAPASON ontology will be further explained in section 4. Before that, it is necessary to describe the DIAPASON ontology and the way in which problem schemas are formalized.

### 3. DIAPASON Ontology

As can be observed in Perinián-Pascual (2023, 2024a, 2024b), Fernández-Martínez (2024) and Jiménez-Briones and Felices-Lago (2024), each type of problem in the DIAPASON ontology is represented with a problem schema. As an example, the schemas corresponding to the problem types DISAPPOINTMENTMONARCHY (SOCIAL problem realm / GOVERNANCE dimension / POLITICALSYSTEM domain) and BEACHSECURITY (PHYSICAL problem realm / INFRASTRUCTURE dimension / COMMUNITYFACILITY domain) are shown in (1) and (2), respectively.

- (1) (DISAPPOINTED | DISTRUST | PESSIMISTIC): (monarchy-108363812 | \$monarch)
- (2) ((P-police-108209687 ^ N-police-108209687) & beach-109217230)

Formally, problem schemas have a macrostructure and a microstructure. On the one hand, the macrostructure of problem schemas is based on the theory of speech acts (Searle, 1969), whereby it consists of two components: the locutionary part expressing factual information and the illocutionary part, which expresses non-propositional



meaning. It should be noted, however, that the illocutionary component of the problem schema can be irrelevant for the recognition of some problem types as shown in (2). In other words, the locutionary part of the schema is sometimes sufficient to recognize some problems. Therefore, problem schemas can represent both the content and the intention of messages posted by social network users. On the other hand, the microstructure of problem schemas consists of simple elements (i.e., concepts, functions, named entities and operators) and complex elements (i.e., expressions), which are described below:

- a. The concepts have the form of *WordNet synsets* (Fellbaum, 1998), which are linguistically constructed using lexical units from different languages. *WordNet* is a lexical database where nouns, verbs, adjectives, and adverbs are organized into synonym sets (i.e., synsets). Each *synset*, representing a different concept, is connected to other *synsets* through lexical and semantic relationships.<sup>5</sup> For example, 108363812 in schema (1) and 108209687, 109217230 in schema (2) are *WordNet 3.0 synsets*. English words accompanying the *synsets* in the above schemas are not strictly necessary, since lexical units from different languages can be automatically retrieved from a *synset*. However, the most representative word of the *synset* is included to assist in the reading of these knowledge representations.
- b. Functions are categories of pragmatic meaning that can be linguistically realized through formulaic constructions, i.e., lexical, syntactic and semantic patterns that can be implemented in various linguistic realizations. For social problem recognition, Fernández-Martínez (2024) collects the communicative functions of three types that were used by Blundell et al. (1982): informative (e.g. saying that something is wrong), attitudinal (e.g. expressing disagreement), and actional (e.g. warning of danger). For example, DISAPPOINTED and PESSIMISTIC in (1) are linguistically projected to constructions such as “it’s a real pity” and “That’s too bad”, respectively, among many others. However, it was soon realized that there was a need to include not only new functions (e.g. saying you do not have confidence in someone or something) but also new patterns for existing functions different from the accepted standard for written text.
- c. Named entities are proper names (e.g., individuals, places, organizations, etc.). From an ontological perspective, named entities in problem schemas can take the form of instances, e.g., \$European\_Union, or categories of instances, e.g. \$monarch in schema (1), which can be instantiated by the named entity of any member of the royal family of a country, e.g., King Charles III in the United Kingdom.

---

5. It should be noted that *WordNet* is a lexical database that groups words into *synsets* and stores the connections that these *synsets* establish among themselves by means of the semantic relations of hyponymy, hyperonymy, troponymy, holonymy and meronymy.

- d. There are two types of operators. On the one hand, conceptual operators, such as quantity modifiers (M, P) or negation (N), provide more semantic specificity to the description of the type of problem. Modifiers intend to represent a particular quality, entity, event or situation at a higher (M) or lower (P) quantity, degree or intensity than expected for the state of affairs being described, where the referents modified can be nouns (e.g., many/few dogs), verbs (e.g. drink a lot/little), or adjectives (e.g., highly/slightly influential). In contrast, negation (N) refers to any construct that introduces the lack of a given quality, entity, event or situation. On the other hand, logical operators, i.e., conjunction (&), inclusive disjunction (|) and exclusive disjunction (^), connect two or more elements of the same type.
- e. Expressions hold a set of concepts, functions, named entities or other expressions, only if they are linked with logical operators. For that purpose, expressions employ round brackets.

It should be noted that the construction and formalization of these social problems in DIAPASON is carried out using the functionalities that this ontology module makes available to the researchers in charge of this task, namely, an editor, a browser, a search engine, and a validator (Periñán-Pascual, 2023). In addition, for this knowledge modeling task, the methodology illustrated in figure 1 below is followed. As an example, the problem of low pensions is provided (Jiménez-Briones & Felices-Lago, 2024):

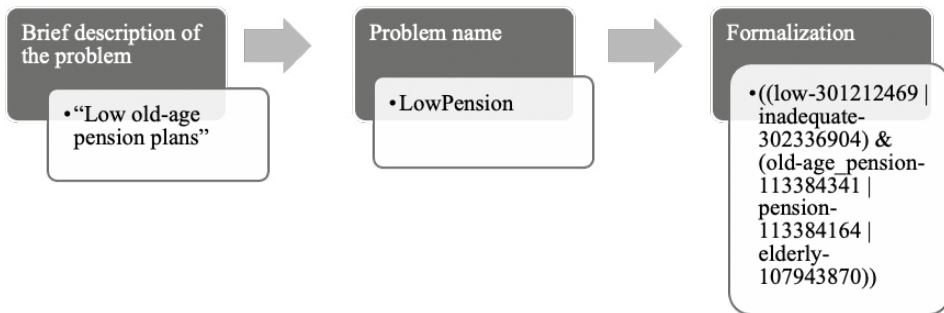


FIGURE 1. PROBLEM SCHEMA DEVELOPMENT IN DIAPASON

#### 4. Listing and modeling of WORK domain problems in DIAPASON

Well-known social problems encyclopedias (Parrillo, 2008, Eitzen et al., 2014, Seccombe & Kornblum, 2020) have compiled various lists of social problems based on economic factors. Even if they do not coincide with each other on their breakdown, at least work-related problems are identified by all of them as one of the crucial axes for understanding the scope of this domain.

#### 4.1. Methodology

The selected list of specific problems presented in table 1 or table 2 (see section 4.3) and included in the DIAPASON knowledge organization system corresponds to the following methodological criteria:

- a. The social problems linked to the WORK domain and presented in section 4.3 have been preliminary noted in terms of a bibliographical consultation of the chosen topic in the previously mentioned encyclopedias and some classical textbooks on labor economics, such as Kaufmann and Hotchkiss (2000), Borjas (2006); Ehrenberg and Smith (2009), among others.
- b. Specialized academic journals such as *The Journal of Labor Economics (JOLE)*, *Labour Economics*, *The Journal of Economic Inequality* and *Industrial and Labor Relations Review (ILR Review)* have also been consulted.
- c. The social media network *Twitter* (now *X*) has been preliminary tracked to know how experts, on the one hand, and ordinary citizens, on the other, have been expressing their complaints linked to the target problems in their tweets in English.

It is important to bear in mind that the list given below (4.2, 4.3) is by no means exhaustive and in many cases the way of referring to a problem subsumes other similar ways of doing so. In consequence, not all of them have been included here to avoid excessive redundancy.

#### 4.2. Description of the problems

Probably the most cited and widespread labor problem at macroeconomic level and whose effects are most noticeable in microeconomics (as perceived and suffered by citizens) is that of unemployment, especially when the rate is high (HIGHUNEMPLOYMENT), without forgetting that of precarious employment (UNDEREMPLOYMENT), but the problem of informal employment (UNDERTHE TABLEWORK) must also be highlighted in those countries where tax control is limited and the underground economy is very powerful. In poor countries or those under the impact of serious economic crises, other labor problems associated to survival or severe poverty are common (NODOLE, NOINCOME). Abusive working conditions as well as lack of protection at work (UNSAFEWORKINGCONDITIONS) are a very common and permanent set of complaints in social networks and collected in the literature on social problems. The most serious abuse is the one affecting children (CHILDLABOR), but no less cruel is work in semi-slavery (SWEATSHOP) or non-payment for work already done (WAGETHEFT) or deserved (NOBENEFITS). No less serious is harassment at work (WORKPLACEBULLYING), preventing workers from unionizing (NOUNIONS) or overwork leading to collapse (BURNOUT).

Income differences are also a recurring issue in the media, whether they are of a general nature (INCOMEDISPARITY) or, frequently, with a gender bias (WAGEGAP). Wages that

are too low (SALARYISLOW) or salary reductions (WAGEREDUCTION) are other problems frequently associated with the aforementioned. Discrimination in the workplace for various reasons (sex, race, nationality, etc.) is included in the following umbrella concept: JOBDISCRIMINATION, but age-related discrimination has a specific concept: AGEISM. The fear of losing one’s job (JOBINSECURITY) is a big problem in times of crisis, but the rapid advance of new technology and automation (WORKFORCEAUTOMATION) also causes a lot of fear. Now, the lack of good communications in rural areas, especially the Internet, (DIGITALDIVIDE) makes employment difficult, but also the lack of job opportunities leads to the flight of the most qualified young people (BRAINDRAIN) and when they manage to work, too much dedication often prevents them from reconciling work and family life (WORKLIFEIMBALANCE). No less inconvenient is the scarce reward that many older workers have once they have retired (LOWPENSION). Other no less relevant problems could be mentioned, but the ones that are included in this section are the most recurrent in the sources consulted.

**4.3. Listing of problems**

When exploring all the accessible resources mentioned in the three methodological criteria explained in 4.1, claims and complaints about WORK problems can be subdivided into two categories (see table 1). The more technical ones are usually voiced by experts and literate citizens, while the informal ones are expressed by the average citizen in more colloquial terms, particularly in social networks. The extent to which the two categories should be merged will have to be determined at some point in the final phases of the ALLEGRO project.

TABLE 1. DISTRIBUTION OF PROBLEMS  
RELATED TO THE WORK DOMAIN

<b>Complaints of a more technical nature</b>	<b>More informal citizen's complaints</b>
CHILDLABOR	AGEISM
DIGITALDIVIDE	BRAINDRAIN
HIGHUNEMPLOYMENT	BURNOUT
INCOMEDISPARITY	NOBENEFITS
JOBDISCRIMINATION	NODOLE
JOBINSECURITY	NOINCOME
LOWPENSION	NOUNIONS
UNDEREMPLOYMENT	SALARYISLOW
UNSAFEWORKINGCONDITIONS	SWEATSHOP
WAGEREDUCTION	WAGEGAP
WORKFORCEAUTOMATION	WAGETHEFT
WORKLIFEIMBALANCE	UNDERTHETABLEWORK
WORKPLACEBULLYING	

Following the application of the methodology explained in sections 3 and 4.1, the problems detected and finally collected in table 2 (also represented in the DIAPASON knowledge organization system<sup>6</sup>) have been formalized in alphabetical order as follows:

TABLE 2. LIST OF PROBLEMS OR COMPLAINTS RELATED TO THE WORK DOMAIN AND THEIR CORRESPONDING SCHEMAS

<b>Problem</b>	<b>Problem schema</b>
AGEISM <sup>7</sup>	ageism-101154373
BRAIN DRAIN	brain_drain-101062468
BURNOUT	(stress-114376188   M-work-100584367   overworking-100623370   exhaustion-114017206   syndrome-114304060)
CHILD LABOR	((abuse-100419908   exploitation-100418903   N-human_right-105176846) & work-201162754 & child-109917593)
DIGITAL DIVIDE	((N-access-202248808 & (internet-103580615   e-mail-106279326))
HIGH UNEMPLOYMENT	((high-301210854   excessive-301533974) & unemployment-113968308)
INCOME DISPARITY	((disparity-104752530   inequality-104752221) & income-113255145)
JOB DISCRIMINATION	((discrimination-101153548 & work-104602044)   (employee-110053808   applicant-109607280)   (gender-105006898   homosexuality-100856847   age-104924103   religion-105946687   disability-114548343   nationality-107949463))
JOB INSECURITY	WORRIED: (job-100582388   loss-113327676   instability-113975659)
LOW PENSION	((low-301212469   inadequate-302336904) & (old-age_pension-113384341   pension-113384164   elderly-107943870))
NO BENEFITS	((N-benefit-105142641   No-perk-113298011) & (job-100582388   government-108050678))
NO DOLE	N-dole-101088304
NO INCOME	((N-income-113255145   N-earnings-113279262)   time_period-115113229   unemployment-113968308)

6. In the DIAPASON Knowledge Organization System, it is possible to visualize not only the name of the problem and the problem schema, but also the problem description and the top-down hierarchy of the domain in the ontology.

7. AGEISM has not been included as part of the JOB DISCRIMINATION problem because it is generally perceived as a very specific set of problems focused on elderly people, even if it could also be integrated in the latter problem under certain circumstances.

NOUNIONS	N-union-108233056
SALARYISLOW	(low-302336759 & salary-113279262)
SWEATSHOP	sweatshop-104370600
UNDEREMPLOYMENT	((P-employment-113968092 & P-wage-113279262)   P-skilled-302225510   P-time-115270431)
UNDERTHE TABLEWORK	((undeclared-300688581   illegal-301401854   informal-301634027) & work-100575741)
UNSAFEWORKING CONDITIONS	((dangerous-302058794   unhealthy-302559534) & work-104602044 & environment-113934596)
WAGEGAP	((disparity-104752530 & wage-113279262)   gender-105006898)
WAGEREDUCTION	(reduction-100351638 & wage-113279262)
WAGETHEFT	((N-wage-113279262   N-benefit-105142641   No-perk-113298011) & (illegal-301401854   (labor_contract-106526961 & violation-100770543)))
WORKFORCE AUTOMATION	((workforce-108212347 & N-human-100007846)   (software_system-106566077   hardware-103493333   robot-102761392   automation-100102457   (repetitive-301964367 & task-100719705)))
WORKLIFEIMBALANCE	((imbalance-114004149   conflict-113980845) & (work-100575741 & (personal-301767329 & life-113961399)))
WORKPLACEBULLYING	(workplace-104602044 & bullying-101222477)

Regarding the problem schemas that encode each citizen complaint in table 2, a detailed explanation of their composition and structure is provided in Perrián-Pascual (2023). However, in the following lines we describe the steps followed in the formalization of the semantic content of two very common problems in citizens' complaints regarding the WORK domain, one of a more technical nature (JOBDISCRIMINATION) with applications to diverse situations; the other, more informal, as it is the case of NOBENEFITS, when employees suffer the effects of company or institutional decisions to cut down on their salaries in a more indirect way. As for the first example, table 3 shows its description, name and outline.

In this schema only the locutionary information provided by the relevant *WordNet synsets* is available and no illocutionary information is added, since it is better to opt for the minimal expression. This circumstance is repeated in practically all the similarly collected problem schemas in this domain, except in JOBINSECURITY, in which the function WORRIED is necessary.

In JOBDISCRIMINATION the *synset* discrimination-101153548 is linked by the conjunction logical operator “&” to a series of nominal *synsets* such as work-104602044, employee-110053808 or applicant-109607280, all of them linked by the inclusive disjunction logical operator “|”, as they refer to the entities that might suffer from this kind

of discrimination. At the same time, the inclusive disjunction logical operator “|” also adds the group of entities referring to the typologies of discrimination such as gender-105006898, homosexuality-100856847, age-104924103, religion-105946687, disability-114548343 and/or nationality-107949463. Evidently, the diversity of nominal *synsets* opens the range for the detection of the maximum number of representative tweets. Thus, combining examples of *synsets* from the two parts united by parentheses would facilitate the detection of a much larger number of tweets.

TABLE 3. COMPONENTS OF THE JOBDISCRIMINATION PROBLEM

Description	Name	Schema
Making an adverse decision against the employee or the job applicant on the grounds of gender, sexual orientation, age, religion, physical disability or nationality	JOB DISCRIMINATION	((discrimination-101153548 & work-104602044   employee-110053808   applicant-109607280)   (gender-105006898   homosexuality-100856847   age-104924103   religion-105946687   disability-114548343   nationality-107949463))

The other example, NOBENEFITS in table 4 offers different schema characteristics if it is compared to the previous case in table 3.

TABLE 4. COMPONENTS OF THE NOBENEFITS PROBLEM

Description	Name	Schema
Situation with no well-being financial assistance in time of need or job without additional advantages	NOBENEFITS	((N-benefit-105142641   N-perk-113298011) & (job-100582388   government-108050678))

This problem schema also relies solely on propositional or locutionary information supplied by the relevant *WordNet synsets* for the same reasons explained above. However, this time the *synsets* benefit-105142641 and perk-113298011 are modified by a conceptual operator of negation (N), which provides more semantic specificity to this type of problem. These two *synsets* are delimited by parentheses and linked to two *synsets*, job-100582388 and government-108050678, by the logical conjunction operator “&”. This means that jobs in general or governmental bodies are usually affected by this lack of incentives. The combinatorial flexibility of this problem schema can rescue large numbers of examples referring to the same problem, although this last statement will have to be verified at a later stage of this research.

## 5. Conclusions

The selected domain, WORK, as well as the representative examples of problem schemas analyzed in more detail, JOBDISCRIMINATION and NOBENEFITS, show how the DIAPASON module, integrated in the ALLEGRO project, can be a useful tool for analyzing microtexts coming from social networks and referring to dimensions or domains of all kinds.

It is worth noting that a significant number of problem schemas for the WORK domain at the ontological level can be found in more than one domain of the same ECONOMY dimension, as is the case, among others, of HIGHUNEMPLOYMENT and SALARYISLOW, which are also part of the MACROECONOMICS domain, or LOWPENSION, AGEISM and DIGITALDIVIDE, which are part of the POPULATION domain in the LIVING dimension (Jiménez-Briones & Felices-Lago, 2024). This fact facilitates the understanding of the multidimensional and complex nature of social problems, which, in this case, go even beyond the purely economic aspects and merge with other domains of other dimensions.

Another relevant aspect is the methodology followed for the selection of problems as well as the criteria used to provide the lexical and syntactic features for the modeling of the problem schemas explained in section 3, whose characteristics are inspired by COREL interface language and eventually allowing computer processing. However, the validity of the collected problem schemas offered must be decisively checked in the following phase of this project, which consists in the collection of a corpus of User Generated Content (UGC) units from the *Twitter (X)* social platform and determine to what extent these problem schemas are able to detect a wide range of examples. For that purpose, a specific application will be used, namely, *Twitter Searcher*, specifically developed for this project (Perián-Pascual, 2024a).

## Acknowledgements

This publication is part of the R&D&I project PID2020-112827GB-I00, funded by MICIU/AEI/10.13039/501100011033.

## References

- Appio, Francesco Paolo, Lima, Marcos & Paroutis, Sotirios (2019). Understanding smart cities: Innovation ecosystems, technological advancements, and societal challenges. *Technological Forecasting & Social Change*, 142, 1-14.
- Best, Joel (1989). Dark Figures and Child Victims: Statistical Claims about Missing Children. In: J. Best (Ed.). *Images of Issues* (pp. 21-25). Aldine de Gruyter.
- Best, Joel (1995). Typification and social problems construction. In: J. Best (Ed.), *Images of issues: Typifying contemporary social problems* (pp. 1-10). Routledge.



- Best, Joel (2017 [2008]) *Social Problems*, (3rd edition). W.W. Norton & Company.
- Blundell, John; Higgens, Jonathan & Middlemiss, Nigel (1982). *Function in English*. OUP.
- Borjas, George (2006). *Labor Economics*. McGraw Hill.
- Bundesregierung (2018). *Government Report on Wellbeing in Germany*. Berlin: Federal Press Office. Retrieved from <http://www.gut-leben-in-deutschland.de/downloads/Government-Reporton-Wellbeing-in-Germany.pdf>.
- Charon, Joel & Vigilant, Lee (2005). *Social Problems: Readings with Four Questions*. Wadsworth.
- DeFronzo, James & Gill, Jungyun (2020). *Social Problems and Social Movements*. Rowman & Littlefield.
- Davis, Kingsley (1937). The Sociology of Prostitution. *American Sociological Review*, 2(5), 746-755.
- Ehrenberg, Ronald & Smith, Robert (2009). *Modern Labor Economics*. Prentice Hall.
- Eitzen, Stanley, Zinn, Maxine & Smith, Kelly (2014). *Social Problems* (13th edition). Allyn and Bacon.
- Eurostat (2017). *Final report of the expert group on quality-of-life indicators*. Publications Office of the European Union.
- Fellbaum, Christiane (Ed.) (1998). *WordNet: An Electronic Lexical Database*. MIT Press.
- Fernández López, Mariano, Gómez-Pérez, Asunción & Juristo, Natalia (1997). METHONTOLOGY: From ontological art towards ontological engineering. In *Proceedings of the AAAI-97 Spring Symposium on Ontological Engineering (Palo Alto [California], March 24-25 1997)* (pp. 33-40). AAAI ([www.aaai.org](http://www.aaai.org)). Retrieved from <https://cdn.aaai.org/Symposia/Spring/1997/SS-97-06/SS97-06-005.pdf>
- Fernández-Martínez, Nicolás José (2024). Taxonomía de funciones comunicativas negativas para su identificación automática en el contexto de las ciudades inteligentes. In Françoise Olmo-Cazeville, (Ed.). *Investigación lingüística en entornos digitales*, 217-246. Tirant Lo Blanch.
- Giffinger, Rudolf, Fertner, Christian, Kramar, Hans, Kalasek, Robert, Pichlermilanovic, Natasha & Meijers, Evert (2007). *Smart Cities Ranking of European Medium Sized Cities*. Retrieved from [http://www.smart-cities.eu/download/smart\\_cities\\_final\\_report.pdf](http://www.smart-cities.eu/download/smart_cities_final_report.pdf).
- Govada, Sujata, Spruijt, Widemar & Rodgers, Timothy (2017). Smart city concept and framework. In Vinod, Kumar (Ed.) *Smart Economy in Smart Cities. Advances in 21st Century Human Settlements*, 187-198. Springer.
- Heiner, Robert (2002). *Social problems: An introduction to Critical Constructionism*. OUP.
- Helliwell, John, Layard, Richard, Sachs, Jeffrey & De Neve, Jan Emmanuel (Eds.) (2020). *World Happiness Report 2020*. Sustainable Development Solutions Network.
- Hitzler, Pascal, Gangemi, Aldo, Janowicz, Krzysztof, Krisnadhi, Adila & Presutti, Valentina (Eds.). (2016). Ontology Engineering with Ontology Design Patterns – Foundations and Applications. *Studies on the Semantic Web* (vol. 25). IOS Press, 3-21.

- Holstein, James & Miller, Gale (2003). *Challenges & Choices*. Gruyter.
- Jiménez-Briones, Rocío & Felices-Lago, Ángel (2024). La formalización del conocimiento en DIAPASON a través de una muestra de problemas poblacionales y macroeconómicos. In Françoise Olmo-Cazevieuille, (Ed.). *Investigación lingüística en entornos digitales*, 187-216. Tirant Lo Blanch.
- Kaufmann, Bruce & Hotchkiss, Julie (2000). *The Economics of Labor Markets*. Dryden Press.
- Wei, Li, Wen-jun, Wu, Huai-min, Wang, Xue-qi Cheng, Hua-jun, Chen, Zhi-hua, Zhou & Rong, Ding (2017). Crowd intelligence in AI 2.0 era. *Frontiers of Information Technology & Electronic Engineering*, 18, 15-43.
- McStay, Andrew (2020). Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data and Society*, 7(1).
- Mooney, Linda; Clever, Molly & Van Willigen, Marieke (2021). *Understanding Social Problems*. Cengage learning.
- OECD (2020). *How's Life? 2020: Measuring Well-Being*. OECD Publishing. Retrieved from [https://www.oecd-ilibrary.org/economics/how-s-life/volume-/issue-\\_9870c393-en](https://www.oecd-ilibrary.org/economics/how-s-life/volume-/issue-_9870c393-en).
- Office for National Statistics (2019). *Measuring National Well-Being in the UK: International Comparisons 2019*. Retrieved from <https://www.ons.gov.uk/people-populationandcommunity/wellbeing/articles/measuringnationalwellbeing/international-comparisons2019>.
- Parrillo, Vincent (Ed.). (2008). *Encyclopedia of Social Problems*. Sage.
- Perián-Pascual, Carlos (2023). From Smart City to Smart Society: A quality-of-life ontological model for problem detection from user-generated content. *Applied Ontology*, 18 (3), 263-306.
- Perián-Pascual, Carlos (2024a). Modelización de las quejas de los ciudadanos como artefactos digitales culturales: DIAPASON. In Françoise Olmo-Cazevieuille, (Ed.). *Investigación lingüística en entornos digitales*, 187-216. Tirant Lo Blanch.
- Perián-Pascual, Carlos (2024b) Exploring the problems through social media: The case of Beach Quality. In Jiménez-Briones, Rocío & Corral Esteban, Avelino (Eds.). *Approaches to Knowledge Representation and Language*. (pp. 11-26). Colección Interlingua. Comares.
- Perián-Pascual, Carlos & Mairal-Usón, Ricardo (2010). La gramática de COREL: Un lenguaje de representación conceptual. *Onomázein* 21, 11-45.
- Searle, John (1969). *Speech Acts*. C.U.P.
- Secombe, Karen & Kornblum, William (2020). *Social Problems*. Pearson.
- Sirgy, M. Joseph (2018). The psychology of material well-being. *Applied Research Quality Life*, 13, 273-301.
- Spector, Malcolm & Kitsuse, John (1977). *Constructing Social Problems*. Transaction Publishers.
- Van Valin, Robert D. & LaPolla, Randy J. (1997). *Syntax: Structure, Meaning, and Function*. CUP.

# Los actos de amenaza a la imagen en la entrevista sociolingüística: un estudio comparativo entre entrevistas libres y entrevistas estructuradas

Alba Fernández Sanmartín - Universidade de Santiago de Compostela  
alba.fernandez@usc.es

Rebut / Received: 22-02-24

Acceptat / Accepted: 05-07-24

**Resum. Els actes d'amenaça a la imatge a l'entrevista sociolingüística: un estudi comparatiu entre entrevistes lliures i entrevistes estructurades.** L'entrevista sociolingüística és un dels mètodes més freqüentment utilitzats per obtenir mostres destinades a l'estudi de la llengua oral. Tot i això, diferents autors han posat de manifest que les entrevistes no proporcionen un material adequat per a l'anàlisi de certs fenòmens de la llengua conversacional. La cortesia mitigadora ha estat assenyalada com un, mentre que les seves característiques situacionals determinarien que a l'entrevista amb prou feines es produeixin amenaces a la imatge dels parlants.

Partint d'una mostra d'entrevistes, el treball posa de manifest com les diferents estratègies desplegades pels entrevistadors condicionen en gran mesura el tipus de fenòmens lingüístics que apareixen als materials resultants. Específicament, es mostra com a les entrevistes més flexibles, informals i desestructurades, es localitzen abundants i variats exemples d'amenaça a la imatge, cosa que les converteix en un material apte per a l'estudi de certs fenòmens relacionats amb la cortesia.

**Paraules clau:** entrevista sociolingüística, cortesia, imatge, amenaça, metodologia lingüística.

**Abstract. Face-threatening acts in the sociolinguistic interview: a comparative study between free and structured interviews.** The sociolinguistic interview is one of the most commonly used methods for obtaining samples intended for the study of spoken language. However, various

authors have pointed out that interviews do not provide adequate material for the analysis of certain conversational language phenomena. Mitigating politeness would be one of these phenomena, as, due to their situational characteristics, interviews rarely produce face-threatening acts.

Based on a sample of interviews, this work highlights how the different strategies employed by interviewers largely determine the type of linguistic phenomena that appear in the resulting materials. Specifically, it shows how in more flexible, informal, and unstructured interviews, there are numerous and varied examples of face-threatening acts, making this kind of interview suitable material for the study of certain phenomena related to politeness.

**Keywords:** sociolinguistic interview, courtesy, face-threatening acts, linguistic methodology.

## 1. Introducción

La entrevista sociolingüística es uno de los métodos más frecuentemente utilizados para la obtención de muestras destinadas al estudio de la lengua oral. Las razones de que ocupe ese lugar central en los estudios lingüísticos son muchas y variadas. Para empezar, suele proporcionar un sonido de alta calidad (Vida-Castro et al., 2016), determinado por la posibilidad del investigador para escoger el lugar de la entrevista y, dentro de este, crear el escenario adecuado para una buena grabación. Además, al ser los interlocutores solamente dos, no pueden desarrollarse conversaciones paralelas, y los solapamientos, que dificultan la transcripción y la comprensión, son menos frecuentes que en las interacciones con más participantes. Por otra parte, el método de entrevista permite seleccionar a los informantes en virtud de las necesidades de estratificación de la muestra, y proporciona la práctica garantía de obtener material suficiente de cada uno de ellos (Briz, 2012; Silva-Corvalán y Enrique-Arias, 2017), lo cual resulta muy difícil de lograr cuando se registran eventos de habla naturales, entendiéndose como tales, aquellos que se desarrollan con unos objetivos que están al margen de la investigación. Por si esto fuera poco, frente a otros métodos como la grabación secreta de conversación, no presenta ningún tipo de conflicto ético. No es de extrañar, que Schilling (2013b) afirme que le resulta difícil concebir “a better instrument than the sociolinguistic interview in terms of efficiently obtaining large quantities of high quality recorded speech that closely approximates everyday speech” (p. 111).

Sin embargo, prácticamente desde su nacimiento, la entrevista sociolingüística también ha sido objeto de fuertes críticas, que han puesto en entredicho su eficacia como método de investigación para llegar a conocer la lengua de una comunidad de habla (De Fina y Perrino, 2011; Schilling-Estes, 2008). Entre otros aspectos, se ha criticado la circunstancia de que, debido a su propia estructura y a la situación que

crea, presenta serias limitaciones para el estudio de ciertos fenómenos de la lengua conversacional, como pueden ser las estructuras interrogativas (Meyerhoff et al., 2012; Schilling, 2013<sup>a</sup>), los “significados expresivos” (Vázquez Veiga, 2003), la toma de turnos (Briz, 2010; Briz, 2012), las *tag questions* (Serrano, 2011) o el *code-switching* (Meyerhoff et al., 2012).

En la misma línea de los autores anteriores, Briz (2010, 2012) pone el acento en la cortesía. Según este autor, la entrevista puede no ser el lugar más adecuado para el estudio de ciertos fenómenos corteses ya que, en las entrevistas, las imágenes de entrevistador y entrevistado no suelen sufrir amenazas. Esta posición parece estar respaldada por los datos que proporciona Albelda Marco (2004), resultado de un análisis comparativo de los fenómenos corteses localizados en una muestra de entrevistas sociolingüísticas y conversaciones coloquiales. Esta autora encuentra que la frecuencia de actos de amenaza a la imagen en la entrevista se reduce a 2,15 por cada 100 minutos de grabación, mientras que en las conversaciones la cifra asciende a 96,29 por cada 100 minutos.

De todos modos, antes de aceptar que los materiales obtenidos mediante entrevistas no son, en ningún caso, apropiados para el estudio de la cortesía relacionada con las amenazas a la imagen, es necesario tener en cuenta que no todas las interacciones calificadas como *entrevistas* dentro de la disciplina lingüística tienen las mismas características. Así, bajo las denominaciones de *entrevista sociolingüística* o *entrevista semidirigida* se incluye prácticamente cualquier interacción desarrollada entre un entrevistador y un informante que se desarrolle con el fin de recoger muestras de habla para el estudio lingüístico, desde los intercambios más estructurados de preguntas y respuestas, a los más flexibles y relajados, en los que los entrevistadores tratan deliberadamente de acercar la entrevista a una conversación coloquial.

Partiendo de esta base, el objetivo de este trabajo consiste en determinar si las diferencias en la metodología de la entrevista pueden tener alguna influencia en el tipo de fenómenos que se pueden estudiar en ella, poniendo el foco específicamente en comprobar si las entrevistas desestructuradas podrían proporcionar un material válido para el estudio de las amenazas a la imagen, en tanto presentan unas características cercanas a la conversación coloquial, en donde dichos fenómenos son mucho más frecuentes. Para ello, se tomó una muestra de entrevistas dividida en dos partes en función de sus características estructurales, y se aplicó sobre ella un análisis cualitativo y cuantitativo de los actos de amenaza a la imagen.

Con respecto a la estructura del trabajo, en el apartado 2 se presentan la muestra utilizada, así como el método de análisis. En el 3 se muestran los resultados del análisis cuantitativo, y en el 4 se expone el análisis cualitativo dividido en dos partes: primero se analizan los ejemplos localizados en la muestra de ESCEGRA (4.1) y, a continuación, una selección de casos de amenazas a la imagen tomados de la muestra del corpus ESLORA (4.2). El trabajo termina con un apartado final de recapitulación y presentación de conclusiones.

## 2. Muestra, metodología y marco teórico

Para llevar a cabo el análisis se seleccionó una muestra de 145223 palabras y más de doce horas de grabación, dividida en dos partes, una formada por ocho entrevistas procedentes del corpus ESCEGRA, y otra con el mismo número de entrevistas tomadas del corpus ESLORA. Las entrevistas del corpus ESCEGRA tienen una estructura clásica de preguntas y respuestas, mientras que las de ESLORA se desarrollan de manera más libre y desestructurada. En la tabla 1 se presentan las características generales de los dos tipos de entrevistas, siempre teniendo en cuenta que, dentro de cada muestra, existe un cierto nivel de heterogeneidad.

TABLA 1. CARACTERÍSTICAS DE LOS DISTINTOS TIPOS DE ENTREVISTAS DE LA MUESTRA

Entrevista estructurada (ESCEGRA)	Entrevista libre (ESLORA)
Se marca claramente el inicio de la entrevista.	El inicio de la entrevista se funde con la presentación y charla previa.
Empleo de un guion previo, genérico y formado por preguntas sin relación.	Coconstrucción libre turno a turno.
El entrevistador mantiene el control absoluto sobre la interacción.	El entrevistador cede parte del control de la interacción.
El entrevistador se limita a preguntar y emitir marcadores de recepción.	El entrevistador emite todo tipo de intervenciones.
El entrevistador se mantiene neutral.	El entrevistador se implica personalmente en la conversación.

En cuanto a las características de los hablantes, las interacciones fueron seleccionadas tratando de mantener una cierta homogeneidad y equilibrio entre los dos bloques, de manera que pudiesen considerarse comparables. Así, ambas partes cuentan con entrevistadoras de género femenino (seis mujeres en el caso de ESLORA, y dos en el caso de ESCEGRA) e igual número de informantes femeninos y masculinos. Además, todos los participantes tienen edades comprendidas entre los 21 y los 43 años, y estudios medios o superiores.

En la Tabla 2 se presentan las características de la muestra de manera más detallada.

Sobre esta muestra se aplicó un análisis cuantitativo-cualitativo de los actos de amenaza a la imagen (en adelante FTA, acrónimo de *Face Threatening Act*), que consistió en localizar, contabilizar y analizar todos los FTA presentes en las entrevistas. Con el fin de que los resultados obtenidos de cada una de las dos muestras no solo pudiesen ser comparados entre sí, sino también con los del ya citado estudio de Albelda Marco (2004), se partió de la misma base teórica que esta y se trató de reproducir su metodología de análisis.

TABLA 2. CARACTERÍSTICAS DE LA MUESTRA

<b>Entrevista</b>	<b>Corpus</b>	<b>Minutos</b>	<b>Palabras</b>	<b>Entrevistadores (género, edad, estudios)</b>	<b>Informantes (género, edad, estudios)</b>
SCOM_H13_012	ESLORA	66	13284	Mujer, 28, superiores	Hombre, 21, superiores
SCOM_H13_014	ESLORA	63	13620	Mujer, 22, superiores	Hombre, 28, superiores
SCOM_M13_010	ESLORA	55	10854	Mujer, 40, superiores	Mujer, 26, superiores
SCOM_H12_027	ESLORA	78	15787	Mujer, 23, superiores	Hombre, 29, medios
SCOM_H12_028	ESLORA	56	11020	Mujer, 23, superiores	Hombre, 27, medios
SCOM_M12_030	ESLORA	53	10870	Mujer, 22, superiores	Mujer, 30, medios
SCOM_M11_050	ESLORA	61	12095	Mujer, 24, superiores	Mujer, 24, medios
SCOM_M13_008	ESLORA	79	14181	Mujer, 24, superiores	Mujer, 29, superiores
Trans 1	ESCEGRA	42	7260	Mujer, 40, superiores	Hombre, 29, superiores
Trans 2	ESCEGRA	31	4865	Mujer, 21, superiores	Hombre, 25, superiores
Trans 3	ESCEGRA	24	3628	Mujer, 21, superiores	Hombre, 25, superiores
Trans 4	ESCEGRA	27	5831	Mujer, 21, superiores	Mujer, 23, superiores
Trans 5	ESCEGRA	23	4127	Mujer, 21, superiores	Mujer, 29, superiores
Trans 9	ESCEGRA	34	6100	Mujer, 40, superiores	Hombre, 43, superiores
Trans 11	ESCEGRA	35	6418	Mujer, 40, superiores	Mujer, 41, superiores
Trans 12	ESCEGRA	32	5283	Mujer, 40, superiores	Mujer, 41, superiores

Albelda Marco (2004), que sigue a su vez los postulados clásicos sobre cortesía de Brown y Levinson (1978) y Kerbrat-Orecchioni (1996), define los FTA como aquellos actos que “dañan el prestigio público de la persona a la que van dirigidos, que le hacen quedar mal, transmiten críticas u ofensas, en definitiva, que le humillan o le empujan a desempeñar alguna tarea sin contar con la voluntad del otro” (Albelda Marco, 2004, p. 117). Según esta autora, los FTA pueden manifestarse lingüísticamente de maneras muy diversas, entre otras, por medio de amenazas, órdenes, críticas, insultos, recriminaciones, reprimendas, burlas, quejas, reproches, manifestaciones de desprecio, réplicas o rechazos. Cabe enfatizar el hecho de que se habla en todo momento de amenazas a la imagen y no necesariamente de ejemplos de descortesía ya que, tal como señala Albelda Marco (2004), no existe una relación biunívoca entre estos dos fenómenos. Teniendo en cuenta que la interpretación de una determinada manifestación lingüística como cortés o descortés no tiene un carácter universal, sino que es dependiente del contexto situacional en el que se produzca (Félix-Brasdefer, 2018; Spencer-Oatey y Kádár, 2021), en muchos casos se realizan actos que, de manera aislada, se pueden considerar amenazantes para la imagen pero que, observados en su contexto no resultan descorteses, pudiendo ser, por ejemplo, casos de “descortesía fingida” (*mock impoliteness*, Leech, 2014, p. 239) o descortesía “no auténtica” (Bernal, 2008; Félix-Brasdefer, 2018).<sup>1</sup>

### 3. Resultados del análisis cuantitativo

Dado que las dos partes de la muestra tienen el mismo número de entrevistas, pero diferente tamaño, en la tabla 3 se presentan tanto los números de frecuencia absoluta de FTA como las cifras de frecuencia relativa:

TABLA 3. FRECUENCIA DE FTA EN LAS DOS MUESTRAS DE ENTREVISTA

Entrevista	N.º FTA <sup>2</sup>	FTA x100 min	Entrevista	N.º FTA	FTA x100 min
<b>ESLORA</b>			<b>ESCEGRA</b>		
SCOM_H13_012	34	51,51	<b>Trans 1</b>	0	0
<b>SCOM_H13_014</b>	4	6,35	<b>Trans 2</b>	0	0
<b>SCOM_M13_010</b>	1	1,81	<b>Trans 3</b>	0	0
<b>SCOM_H12_027</b>	28	35,9	<b>Trans 4</b>	3	11,11
<b>SCOM_H12_028</b>	15	26,78	<b>Trans 5</b>	0	0

1. Resulta interesante a este respecto la diferencia entre *cortesía codificada* y *cortesía interpretada* establecida por Briz (2004).

2. Dado que la identificación de los FTA tiene un componente de interpretación subjetiva, es necesario contar con un margen de error en la localización y contabilización de los casos.



<b>SCOM_M12_030</b>	15	28,3	<b>Trans 9</b>	0	0
<b>SCOM_M11_050<sup>3</sup></b>	35	57,38	<b>Trans 11</b>	0	0
<b>SCOM_M13_008</b>	11	13,92	<b>Trans 12</b>	0	0
<b>TOTAL ESLORA</b>	143	27,98	<b>TOTAL ESCEGRA</b>	3	1,21

Tal como se puede observar en la tabla 3, los panoramas son radicalmente diferentes en una y otra parte de la muestra, comenzando por los números globales: los 27,98 FTA por cada cien minutos de grabación localizados en ESLORA se convierten en ESCEGRA en 1,21, una cifra diecisiete veces menor. Pero, además, en la muestra de ESLORA, se localizó al menos un FTA en cada una de las ocho entrevistas, mientras que en la muestra de ESCEGRA, los tres FTA identificados se encuentran en solamente una de ellas, entrevista que, como se verá en el apartado 4.1, está en gran medida condicionada por ciertas características de sus hablantes.

Por otra parte, el número de 1,21 FTA por cada 100 minutos hallado en las entrevistas más estructuradas de la muestra se sitúa muy cerca de los 2,15 para el mismo parámetro del estudio de Albelda Marco (2004), sobre todo teniendo en cuenta que varios de los FTA localizados por la autora en las entrevistas de su muestra (Gómez Molina, 2001) son reconocidos por esta como de asignación dudosa.

#### 4. Análisis cualitativo

##### 4.1. FTA en la muestra del corpus ESCEGRA

Los tres FTA localizados en la muestra del corpus ESCEGRA se concentran en una única entrevista, la n.º 4, que tiene la peculiaridad de que se desarrolla entre dos mujeres muy jóvenes (21 y 23 años) unidas, además, por una relación de amistad cercana. Este hecho puede, en gran medida, explicar lo que sucede en los ejemplos (1)-(3):<sup>4</sup>

3. La entrevista SCOM\_M11\_050 no está incluida en la versión actual del corpus ESLORA, la 2.2 de noviembre de 2023.

4. En los ejemplos reproducidos se emplea un sistema de transcripción ortográfica combinado con algunas marcas y etiquetas destinadas a reflejar las particularidades de la lengua oral: la barra simple (/) y la doble (//) se utilizan para marcar, respectivamente, pausas breves y pausas largas (de menos de un segundo), mientras que los corchetes ([]) señalan segmentos de habla solapada. Las etiquetas utilizadas son las siguientes: <risa/>, <alargamiento> (indica aumento de cantidad que afecta a algún sonido de la palabra marcada), <énfasis>/</énfasis>, <ininteligible>, <ruido>, <cita>/</cita> (para los segmentos en estilo directo), <palabra\_cortada> y <silencio> (para pausas de más de un segundo). Por último, <(m:)> y <(e:)> se emplean para marcar apoyos discursivos. Los hablantes se identifican con las iniciales “E” para las entrevistadoras e “I” para los informantes.

E: <(m:)> ¿No te ha pasado nada<alargamiento> así gracioso<alargamiento> en el autobús?

I: He visto accidentes/ cuando se estrella uno/ [eso]

E: [bueno]

I: No tiene mucha gracia vamos.

E: Ya ya y <alargamiento> bueno<(e:)> ¿sabes algún chiste sobre coches?

I: Me estás dejando todo el rato pillada <risas/>

E: No sabes/ bueno. (Trans 4)

E: Bueno y esta vez ¿sabes algún chiste sobre comidas?

I: No sé yo muchos chistes <risas/> hombre/ *tú hazme otras preguntas a ver*

E: <(e:)> Bueno vamos a pasar ahora a<alargamiento> hablar/ de<alargamiento> comidas típicas de los pueblos. (Trans 4)

En (1) se observa cómo, ante la pregunta *¿sabes algún chiste sobre coches?*, la informante formula una queja (*me estás dejando todo el rato pillada*), que constituye un FTA que parece cuestionar la capacidad de la entrevistadora para conducir la entrevista de manera eficaz. En (2), el acto amenazante se corresponde con una orden formulada a través de un imperativo, *hazme otras preguntas*. Los imperativos pronunciados por los entrevistadores no suelen ser interpretados como amenazantes en el marco de las entrevistas, ya que constituyen una parte definitoria del género. En (2) es la entrevistada y no la entrevistadora la que está formulando una orden, lo cual constituye una inversión de los roles comunicativos y una amenaza para la imagen de la entrevistadora, puesto que, de nuevo, se está poniendo en duda su capacidad para desempeñar adecuadamente la tarea que le corresponde.

En (3) las dos interlocutoras están hablando sobre las leyes que restringen los lugares en los que está permitido fumar, con las diferentes consecuencias que estas normas pueden tener:

E: ¿Lo ves como un problema<alargamiento>/ que no puedan fumar un problema que se pueda<alargamiento>/ que pueda deriva<alargamiento>r en consecuencia<alargamiento>s negativas?

I: Sí porque también <alargamiento>/ es una manera de relacionarse mucha gente <alargamiento> es tímida utiliza el tabaco como un pretexto para conocer gente para las discotecas el típico <alargamiento>/ la típica excusa de<alargamiento> <cita> ¿me das un cigarrillo? no sé qué </cita>/ y ya pues si ya no te dejan fumar en esos sitios/ también es una manera de aislar a toda esa gente/ entonces no<alargamiento>

E: ¿Cambiarán los métodos para ligar?

I: También pueden cambiar <risas> un poquito entonces/ *¿qué mala es!* <risas>

E: Bueno <(e:)> ahora<alargamiento> va otro tema distinto. (Trans 4)

En su última intervención, la informante, tras su escueta respuesta, hace una especie de aparte con una exclamación (*¡qué mala es!*) que, a pesar de estar formulada en tercera persona, solo puede ir dirigida a su interlocutora, ya que no hay nadie más presente. Aunque la exclamación no se interprete de manera literal, sí se puede interpretar como una protesta por cómo se está desarrollando la entrevista; la informante parece considerar que la pregunta que se le ha formulado no es adecuada, que su interlocutora trata de dejarla mal o de burlarse de ella.

Obviamente, estos tres FTA han de ser interpretados a la luz de la relación de amistad que une a las dos participantes; es casi seguro que la informante no se hubiera comportado de la misma manera con una entrevistadora desconocida. Cabe señalar, además, que los FTA van acompañados de risas por parte de la hablante que los emite, lo que descarta una intención abiertamente hostil. Sin embargo, esto no invalida su carácter amenazante, ya que, a pesar de que se trate de amigas cercanas, no se debe obviar el hecho de que no se encuentran en un acto social y privado, sino en una interacción transaccional, en un contexto académico, que tiene sus propias normas, para empezar, una distribución desigual del poder comunicativo. Además, la interacción está siendo grabada y va a ser escuchada por un número potencialmente ilimitado de personas, entre otras, los colegas y superiores de una entrevistadora que no está simplemente charlando, sino desempeñando una tarea académica o profesional, de manera que todo aquello que ponga en duda su capacidad para conducir la entrevista de manera adecuada, constituye una amenaza para su imagen. Resulta indicativo, en ese sentido, que la entrevistadora ni siquiera reacciona directamente a los FTA emitidos por su amiga, sino que, en todo momento, trata de continuar la entrevista en un tono formal, como si estuviese hablando con una desconocida.

Como conclusión al análisis de la muestra del corpus ESCEGRA, y teniendo en cuenta también el estudio de Albelda Marco (2004), se puede afirmar que los FTA constituyen un fenómeno muy poco frecuente en las entrevistas de características más clásicas. Su aparición parece estar condicionada por hablantes y circunstancias específicas. Ello no resulta sorprendente teniendo en cuenta que este tipo de entrevistas son intercambios planificados y estructurados en donde cada participante se limita a cumplir su rol, es decir, el entrevistador formula preguntas genéricas, normalmente procedentes de un guion común, sin apenas espacio para la espontaneidad, y el entrevistado las responde de la manera más colaborativa posible con excepciones como la que se acaba de ver.

## **4.2. FTA en las entrevistas de ESLORA**

### *4.2.1. Correcciones*

Como señala Bayraktaroglu (1991) “to correct someone is to mark him with inefficiency or incompetence and therefore is an FTA/other just as an accusation is” (p. 26). En la muestra del corpus ESLORA se localizan varios casos en los que un

interlocutor corrige a otro respecto a diferentes aspectos, poniendo en riesgo su imagen y obligándolo casi siempre a rectificar.

En algunos casos, los FTA nacen del afán de las entrevistadoras por mostrarse participativas, colaborativas y adaptarse a los gustos y circunstancias de los informantes, con el fin de que estos se sientan cómodos. Para lograr ese objetivo, las entrevistadoras han de estar constantemente haciendo deducciones a partir de lo que los informantes les van contando, y ello conlleva un riesgo considerable de hacer asunciones equivocadas.

En el ejemplo (4), las dos mujeres están hablando de la vida profesional de la informante, que, tal como ella misma ha explicado, se reparte entre la práctica odontológica en una clínica y la docencia universitaria. Dicha información lleva a la entrevistadora a deducir que está hablando con una profesora, y a formularle una pregunta al respecto:

I: <ruido=boca> en cambio/ encerrarte en una clínica/ por vida fff/ me costaría [mucho]

E: [ya]

I: que me gusta/ pero para trabajar/ para poder/

E: variar un poco

I: combinarlo con la<alargamiento> facultad/

E: y pasaste muy pronto de alumna a profesora ¿no? así

I: no soy profesora

E: buah <risas>a dar clase </risas>

I: no soy profesora

E: a dar clases/ bueno

I: eeh y doy clase en la/ en<alargamiento>/ prácticas ¿sabes? yo/ eeh yo no tengo ningún tipo de contrato/ [con la universidad] (SCOM\_M13\_008)

En cuanto la entrevistadora le otorga a la informante el título de profesora, esta la corrige inmediatamente, negando su afirmación de manera tajante, sin ningún tipo de preparación o atenuación, y lo hace en dos ocasiones, a pesar de que, después de la primera corrección, la entrevistadora emite una rectificación (*buah a dar clases*). En este caso, la informante parece no estar cómoda cuando se le atribuye una posición profesional que, en sentido estricto, no le corresponde y, aunque las risas rebajan el tono, no duda en contradecir abiertamente a su interlocutora, lo que constituye un acto amenazante al poner en entredicho su capacidad para hacer inferencias acertadas.

Otro ejemplo de corrección de un informante a una entrevistadora aparece en (5):

I: (...) y de repente de ~Huesca nos vinimos a Galicia/ yo cuando llegué aquí <alargamiento>/ yo tenía acento mañico y claro a mí ahora me resulta/ <énfasis>extrañísimo</énfasis> porque yo soy de Santiago de toda la vida *yo soy [gallego / no me]*

E: [te sientes gallego]

I: *no es que me sienta/ soy gallego* <risas> *no me siento soy gallego* </risas>/ y <alargamiento>/ y claro me pienso que tuve acento mañico/ en algún momento de mi vida y era <cita> ¡joder! ¡qué cosa más rara! </cita> (SCOM\_H13\_014)

En este caso, la entrevistadora, en un afán colaborador y cortés, trata de completar el enunciado del entrevistado basándose en los datos que él mismo le ha brindado minutos antes: ha nacido en Madrid y pasado su primera infancia en Huesca. Así pues, su intervención es resultado de una deducción según la cual el entrevistado debe de *sentirse* gallego ya que, desde su punto de vista, no lo *es*. Sin embargo, dicha deducción resulta ser errónea, y así se lo hace saber el informante, que rechaza tajante y repetidamente su afirmación y la sustituye por la correcta: *no es que me sienta/ soy gallego (...) no me siento soy gallego*. La entrevistada, en este caso, no rectifica, sino que, para tratar de cerrar el conflicto sin ver más dañada su imagen, decide introducir un cambio radical de tema en su siguiente intervención: *¡menudo juego de luces!*

Las correcciones realizadas por entrevistadoras a informantes tienen un carácter diferente, ya que no se aplican a deducciones o asunciones falsas, sino simplemente a afirmaciones de cualquier tipo que las entrevistadoras consideran erróneas o poco exactas, y que deciden puntualizar. Por ejemplo, durante la entrevista SCOM\_M11\_050 la informante ha señalado el merengue como su estilo musical favorito, lo que da lugar a que la entrevistadora le pregunte si dicho estilo musical también es del agrado de sus amigos:

E: [a<alargamiento> tus amigos también les gusta eso?]

I: [sí<alargamiento> sí<alargamiento> sí<alargamiento>] no <alargamiento> no les gu <palabra cortada> o sea hay varios tipos de música ¿no? hay otros que tal pero después al final siempre acaban pidiendo el pachanguero <risas><ininteligible></risas>

E: pero/ espera un momento a ver/ una cosa es pachanguero y otra cosa es salsa y merengue de verdad// [<ruido=chasquido boca>]

I: [no<alargamiento>] no bueno/ sí escuchamos un poco de <risas>todo</risas> (SCOM\_M11\_050)

En la respuesta de la informante parece incluirse el merengue dentro de la categoría musical de lo que denomina *pachanguero*, algo con lo que la entrevistadora no está de acuerdo; de ahí que la corrija con una intervención que no solo resulta amenazante por su contenido proposicional, sino por su forma, antes de la corrección en sí misma aparece la conjunción adversativa *pero* y, a continuación, una construcción con un imperativo (*espera un momento a ver*). Este acto amenazante provoca en la informante un momento de vacilación (*no no bueno sí*) y, posteriormente, una afirmación general y poco comprometedor, *escuchamos un poco de todo*, que le permite salir del paso.

#### 4.2.2. Advertencias

En un momento dado de la entrevista SCOM\_H13\_012, el entrevistado, un estudiante de Periodismo, comenta cómo sus compañeros licenciados suelen ver frustradas sus expectativas de ser columnistas en periódicos nacionales o reporteros de guerra, y han de conformarse con cubrir noticias sin apenas repercusión en publicaciones regionales, como por ejemplo *El Progreso* de Lugo. La respuesta de la entrevistadora a este razonamiento es un acto claramente amenazante para la imagen de su interlocutor:

E: el Progreso/ no trates mal al Progreso/ que los dueños son amigos míos / ¿eh? (SCOM\_H13\_012)

Igual que sucedía en (6), esta intervención resulta amenazante no solo por su contenido proposicional, que constituye una advertencia, sino por su forma, que incluye el empleo del imperativo en negativo y el apéndice interrogativo final. Ante eso, el informante no tiene otra opción que iniciar una secuencia de turnos de justificación y reparación, llegando incluso a amenazar su propia imagen pidiendo disculpas. Al mismo tiempo, la entrevistadora trata de reparar el daño diciendo que se trata de una broma aunque específica que su afirmación es cierta, y admitiendo que el ejemplo propuesto por el informante es acertado. Finalmente, y ante la imposibilidad de recuperar el acuerdo, la entrevistadora propone un cambio de tema para cerrar el asunto como se aprecia en (8):

I: bueno y digo [El Progreso/ como digo El Correo Gallego y<alargamiento>]  
 E: [<risa/>/] que no/ que es de b<palabra cortada> es verdad/ pero es de broma <risa/>  
 I: bueno no/ bueno El Correo Gallego no/ perdonen <risa/> digo El Progreso de Lugo porque es un periódico regional  
 E: claro<alargamiento>/ [totalmente]  
 I: [totalmente]  
 E: es paradigma de [lo que estás contando/]  
 I: [es<alargamiento>/ decir] <transcripción dudosa>en vez de ser</transcripción dudosa> la región de Ourense/ mala suerte que tenías amigos en/ ya es mala [suerte ¿no?]  
 E: [sí/ sí sí]  
 I: pero bueno / ¿qué se le va a hacer? [en cualquier caso/ *pido disculpas si se han ofendido*]  
 E: no<alargamiento> <ininteligible> [y ¿qué qué]  
 I: [pero si no/ pues nada <ininteligible>]  
 E: *piensas hacer cuando termines?* / ¿tienes alguna idea<alargamiento>? (SCOM\_H13\_012)

#### 4.2.3. Censuras y críticas

En los ejemplos (9)-(11) las entrevistadoras, de manera más o menos directa, censuran algún aspecto relacionado con la vida de los informantes:

- I: hay unos baños públicos/ que son [los únicos]  
 E: <ininteligible>  
 I: baños públicos que conozco y tal/ y están/ ahí colocados// y cuando voy// cuando es de noche/ porque<alargamiento>/ tengo ganas de ir al baño/  
 E: ah pero/ ¿tú vas?/  
 I: ¿eh?  
 E: yo no he [ido nunca]  
 I: [un día hice] una prueba [en plan de]  
 E: [¿qué miedo!] ¿no?  
 I: ¿eh?/ no no / miedo ninguno [en plan de]  
 E: [quiero decir] *miedo por sucios*<alargamiento>/ ¿no?/ ¿no? (SCOM\_H13\_012)

En este caso, la secuencia *ah pero/ ¿tú vas?*, encabezada por una interjección que, según Vázquez Veiga (2003), puede actuar como una marca de “información no esperada”, seguida de la conjunción adversativa, y con énfasis en el pronombre de segunda persona, lleva implícita una crítica al comportamiento del informante, que se pone todavía más de relieve al utilizar la entrevistadora de manera enfática el pronombre tónico de primera persona en su siguiente turno, oponiendo de ese modo su conducta a la del entrevistado: *yo no he ido nunca*. Además, más adelante, continúa la censura con una exclamación que hace referencia al rechazo que le provoca el comportamiento del informante: *¿qué miedo!* [...] *miedo por sucios*.

En (10), se formula una crítica de manera indirecta:

- E: ¿dónde vives?  
 I: en Fontiñas/ en casa de mis padres  
 E: ah ¿vives con tus padres?  
 I: por ahora sí  
 E: <risas>por ahora sí</risas> [con proyectos] (SCOM\_H12\_027)

La entrevistadora deja claro, al encabezar su pregunta con la interjección *ah*, con valor idéntico al visto en el ejemplo anterior, que el hecho de que el informante viva con sus padres le resulta inesperado y, de manera directa, transmite que, por sus características sociodemográficas (sobre todo la edad), debería estar ya independizado. El entrevistado percibe perfectamente la censura implícita, y por ello se ve obligado a proteger su imagen señalando que es cuestión de tiempo que dicha situación cambie, lo cual es tomado a burla por la entrevistada, que repite entre risas su intervención.

En algunos casos, como (11), la censura contiene un punto de ridiculización, lo que eleva el tono de la amenaza:

- E: y algo equivalente o algo que <ruido=chasquido boca> no sé  
 I: vigilante de seguridad//

E: ah pero eso// ¿qué tipo de vigilante de seguridad por Dios?/ [los que se pasan horas de pie ahí durmiendo]

I: <risas>los de <transcripción dudosa>los de Aída</transcripción dudosa></risas>

E: no eso eso n<palabra cortada>/ l<palabra cortada> a ver/ <risas>tú misma pero/ a mí no me parece equivalente</risas> <risa/>

I: no hombre no como el policía no hay ¡eh!

Antes del fragmento presentado en (11), la informante ha explicado a la entrevistadora que su deseo de ser policía se vio truncado por no cumplir con los requisitos de altura mínima para acceder a las pruebas de selección. Al ser preguntada por una posible alternativa profesional, señala la posibilidad de hacerse vigilante privada de seguridad. Sin embargo, la entrevistadora parece considerar que ese no es un oficio deseable, y lo deja claro en su intervención: de nuevo la interjección *ah* y la conjunción adversativa introduciendo una pregunta terminada con la interjección *por Dios*, que ridiculiza ese oficio y, por extensión, los deseos de la informante y a ella misma. Luego remata la amenaza con una fórmula equivalente a una advertencia (*tú misma*), y especifica que, a ella, el oficio de vigilante de seguridad no le parece en absoluto equivalente al de policía.

#### 4. Conclusiones

Se ha mostrado de manera empírica cómo las características específicas de las entrevistas en cuanto a estructura y comportamiento interactivo de los entrevistadores tienen una gran influencia sobre el tipo de fenómenos lingüísticos que se pueden encontrar en las muestras correspondientes, específicamente en la frecuencia y carácter de FTA. Así, mientras que en las entrevistas estructuradas la presencia de FTA es mínima, en las entrevistas de estructura más libre y flexible, los FTA aparecen en un número relativamente abundante y presentan un carácter variado.

Así pues, los resultados de este estudio demuestran que, si bien no todas las entrevistas constituyen un buen material para el estudio de los FTA, las entrevistas realizadas de manera flexible, como las de ESLORA, sí pueden ser una posibilidad aceptable para el estudio de este tipo de fenómenos, siempre teniendo en cuenta que los resultados de estudios realizados en base a cualquier tipo de entrevista serán una aproximación, nunca un reflejo exacto de lo que sucede en una interacción totalmente espontánea. Este hecho no carece de importancia, ya que los materiales auténticamente coloquiales son, debido a las dificultades éticas y técnicas que implica su recolección, mucho más difíciles de conseguir, en un volumen aceptable y con una distribución equilibrada y representativa de una comunidad de habla, que los de entrevista.



## Referencias

- Albelda Marco, Marta (2004). "Cortesía en diferentes situaciones comunicativas. La conversación coloquial y la entrevista sociológica semiformal". En Diana Bravo y Antonio Briz (Eds.), *Pragmática sociocultural. Estudios sobre cortesía en español* (pp. 109-134). Barcelona: Ariel.
- Bayraktaroglu, Arin (1991). Politeness and interactional imbalance. *International Journal of the Sociology of Language*, 92, 5-34. doi: <https://doi.org/10.1515/ijsl.1991.92.5>
- Bernal, María (2008). ¿Insultan los insultos? Descortesía auténtica vs. descortesía no auténtica en el español coloquial. *Pragmatics*, 18(4), 775-802.
- Briz, Antonio (2004). "Cortesía verbal codificada y cortesía verbal interpretada en la conversación". En Diana Bravo y Antonio Briz (Eds.), *Pragmática sociocultural. Estudios sobre cortesía en español* (pp. 67-94). Barcelona: Ariel.
- Briz, Antonio (2010). "El registro como centro de la variedad situacional. Esbozo de la propuesta del grupo Val.Es.Co. sobre las variedades diafásicas". En Irene Fonte y Lidia Rodríguez Alfano (Comps.), *Perspectivas dialógicas en estudios del lenguaje* (pp. 21-56). México: Universidad Autónoma de Nuevo León.
- Briz, Antonio (2012). "Los déficits de los corpus orales del español (y de algunos análisis)". En Tomás Jiménez Juliá, Belén López Meirama, Victoria Vázquez Rozas y Alejandro Veiga (Eds.), *Cum corde et in nova grammatica. Estudios ofrecidos a Guillermo Rojo* (pp. 115-137). Santiago de Compostela: Universidade de Santiago de Compostela, Servizo de Publicacións e Intercambio Científico.
- Brown, Penelope, y Levinson, Stephen (1978). "Universals in Language Usage: Politeness Phenomena". In Esther N. Goody (Ed.), *Questions and Politeness: Strategies in Social Interaction* (pp. 56-310). Cambridge: Cambridge University Press.
- De Fina, Anna, y Perrino, Sabina (2011). Introduction: Interviews vs. 'natural' contexts: A false dilemma. *Language in Society* 40, 1-11. doi: <https://doi.org/10.1017/S0047404510000849>
- ESLORA: Corpus para el estudio del español oral, versión 2.0 de septiembre de 2020, ISSN: 2444-1430. Recuperado de <http://eslora.usc.es>. Actualmente el corpus está en su versión 2.2, de noviembre de 2023, lo que implica que en los ejemplos reproducidos en el trabajo podría haber ligeras diferencias respecto a la última versión.
- Félix-Brasdefer, J. César (2018). *Pragmática del español: contexto, uso y variación*. Londres: Routledge.
- Gómez Molina, José R. (Coord.) (2001). *El español hablado de Valencia. Materiales para su estudio. Nivel sociocultural alto*. Anejo XLVI de *Cuadernos de Filología*. Valencia: Universitat de Valencia.
- Kerbrat-Orecchioni, Catherine (1996). *La conversation*. Paris: Seuil.
- Leech, Geoffrey N. (2014). *The pragmatics of politeness*. Oxford: Oxford University Press.

- Meyerhoff, Miriam, Adachi, Chie, Nanbakhsh, Golnaz, y Strycharz, Anna (2012). "Sociolinguistic Fieldwork". In Nicholas Thieberger (Ed.), *The Oxford handbook of linguistic fieldwork* (pp. 121-146). Oxford: Oxford University Press.
- Moya Corral, Juan Antonio (Coord.) (2007). *El español hablado en Granada. Corpus oral para su estudio sociolingüístico. I Nivel de estudios alto*. Granada: EUG.
- Schilling, Natalie (2013a). *Sociolinguistic Fieldwork. Key Topics in Sociolinguistics*. Cambridge: Cambridge University Press.
- Schilling, Natalie (2013b). "Surveys and interviews". In R. J. Podesva y D. Sharma (Ed.), *Research methods in linguistics* (pp. 96-115). Cambridge; New York: Cambridge University Press.
- Schilling-Estes, Natalie (2008). "Stylistic variation and the sociolinguistic interview: a reconsideration". In R. Monroy Casas y A. Sánchez Pérez (Coords.), *25 años de lingüística en España: hitos y retos = 25 years of applied linguistics in Spain: milestones and challenges* (pp. 971-986). Murcia: Servicio de Publicaciones de la Universidad de Murcia.
- Serrano, María José (2011). *Sociolingüística*. Barcelona: Ediciones del Serbal.
- Silva-Corvalán, Carmen, y Enrique-Arias, Andrés (2017). *Sociolingüística y pragmática del español*. Washington D.C.: Georgetown University Press.
- Spencer-Oatey, Helen, y Kádár Daniel Z. (2021). *Intercultural Politeness: Managing Relations across Cultures*. Cambridge: Cambridge University Press.
- Vázquez Veiga, Nancy (2003). *Marcadores discursivos de recepción*. Santiago de Compostela: Universidade de Santiago de Compostela.
- Vida-Castro, Matilde, Ávila Muñoz, Antonio, y Carriscondo-Esquivel, Francisco (2016). *Manual práctico de sociolingüística*. Madrid: Síntesis.

# Patrones metafóricos y proceso judicial: conceptualización y comunicación

**Mariana Cucatto** - Universidad Nacional de La Plata (UNLP) y  
Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET)  
marianacucatto@yahoo.com.ar

Rebut / Received: 07-03-24

Aacceptat / Accepted: 27-05-24

**Resumen. Patrones metafóricos y proceso judicial: conceptualización y comunicación.** Las metáforas en los textos jurídicos no son un mero ornamento, sino que constituyen un cotidiano, abundante y poderoso instrumento de pensamiento y comunicación. El propósito de este artículo consiste en estudiar el valor que adoptan las metáforas en el ámbito de la práctica judicial y, en particular, el lugar que ocupan en las sentencias. A partir de una muestra de resoluciones emitidas por diferentes órganos jurisdiccionales de la provincia de Buenos Aires (Argentina) nos interesa: (a) analizar algunos patrones metafóricos —en tanto formas o esquemas de conceptualización recurrentes— destinados a conceptualizar el proceso judicial; (b) mostrar cómo el empleo de estos patrones metafóricos obedece a determinadas motivaciones pragmáticas que ponen en evidencia un modo de pensar metafórico, mediante el cual se instauran ciertas formas de percibir, de pensar, de actuar y de comunicar. Consideramos, además, que este pensamiento metafórico se vincula con rutinas comunicativas altamente normativizadas y normalizadas.

**Palabras clave:** sentencias, patrones metafóricos, proceso judicial, pensamiento metafórico, rutinas comunicativas.

**Resum. Patrons metafòrics i procés judicial: conceptualització i comunicació.** Les metàfores als textos jurídics no són un mer ornament, sinó que constitueixen un quotidià, abundant i poderós instrument de pensament i comunicació. El propòsit d'aquest article és estudiar el valor que adopten les metàfores en l'àmbit de la pràctica judicial i, en particular, el lloc que ocupen en les sentències. A partir d'una mostra de resolucions emeses per diferents òrgans jurisdiccionals de la província de Buenos Aires (Argentina) ens interessa: (a) analitzar alguns patrons metafòrics —en tant que formes o esquemes de conceptualització recurrents— destinats a conceptualitzar

el procés judicial; (b) mostrar com l'ús d'aquests patrons metafòrics obeeix a determinades motivacions pragmàtiques que posen en evidència una manera de pensar metafòrica, mitjançant la qual s'instauren certes maneres de percebre, pensar, actuar i comunicar. Considerem, a més, que aquest pensament metafòric es vincula amb rutines comunicatives altament normativitzades i normalitzades.

**Paraules clau:** sentències, patrons metafòrics, procés judicial, pensament metafòric, rutines comunicatives.

**Abstract. Metaphorical patterns and judicial process: conceptualization and communication.**

Metaphors in legal texts are not a mere ornament, but constitute a daily, abundant and powerful instrument of thought and communication. The purpose of this article is to study the value of metaphors in judicial practice and, in particular, their place in sentences. Based on a sample of resolutions issued by different courts in the province of Buenos Aires (Argentina) we are interested in: (a) analyzing some metaphorical patterns –as recurring forms or schemes of conceptualization– aimed at conceptualizing the judicial process; (b) show how the use of these metaphorical patterns obeys certain pragmatic motivations that reveal a metaphorical way of thinking, through which certain ways of perceiving, thinking, acting and communicating are established. We also consider that this metaphorical thought is linked to communicative routines, highly normativized and normalized.

**Keywords:** sentences, metaphorical patterns, judicial process, metaphorical thinking, communicative routines.

## 1. Introducció

Este artículo forma parte de una investigación mayor<sup>1</sup> en la que nos hemos propuesto estudiar, desde el marco teórico-metodológico de la Lingüística Cognitiva, más especialmente, a partir de la teoría de la metáfora conceptual (Evans y Green, 2004; Gibbs, 2017; Kövecses, 2010, 2015; Lakoff y Johnson, 1980; Lakoff, 1993, 2008; Soriano, 2012), el rol que adquieren las metáforas en el ámbito de la práctica

---

1. Este trabajo se enmarca en dos proyectos de investigación: (1) “En torno a las metáforas conceptuales en el lenguaje jurisdiccional: una mirada desde la lingüística cognitiva”, dirigido por la Dra. Mariana Cucatto, Proyecto de Investigación y Desarrollo, financiado por la Universidad Nacional de La Plata (UNLP); (2) “Las metáforas en las sentencias judiciales: lenguaje, conceptualización y comunicación. Una aproximación desde la Lingüística Cognitiva”, financiado por el Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina. Por otra parte, deseamos aclarar que una primera versión de este trabajo fue presentada en la conferencia inaugural del XII Encuentro Nacional de la Asociación Latinoamericana de Estudios del Discurso (Delegación Chile), celebrado en la Universidad de Las Américas, Santiago, 24 y 25 de agosto de 2023.

jurisdiccional y, específicamente, en las sentencias. Se toma este género como un espacio de conceptualización en el que se representan términos propios del derecho, escenas, sujetos, relaciones sociales conflictivas y en el que se fundamenta una decisión que dirime intereses en disputa.

Desde esa perspectiva, entendemos la metáfora como un mecanismo cognitivo presente en el lenguaje, el pensamiento y la experiencia, que interviene activa y creativamente en la construcción dinámica del significado (Cucatto y Rojas, 2024; Cuenca y Hilferty, 1999; Ibarretxe-Antuñano y Valenzuela, 2021; Lakoff y Johnson, 1980; Lakoff, 1993, 2008; Ibarretxe-Antuñano y Cadierno, 2019). Además, estimamos que resultan fundamentales los aspectos pragmáticos que subyacen a la producción e interpretación de las metáforas, recuperando de la tradición iniciada por la teoría de la metáfora conceptual el potencial comunicativo que aquellas poseen y su relación con contextos socio-culturales de uso, esto es, su carácter “situado”.

Nuestra propuesta se encuadra en una estrategia metodológica de carácter cualitativo, basada en el análisis de datos lingüísticos reales, extraídos de un corpus conformado por una muestra intencional de sentencias proferidas por diferentes órganos jurisdiccionales de la provincia de Buenos Aires (Argentina), durante el período 2000-2023.

En esta oportunidad nos interesa abordar aquellos patrones metafóricos —en tanto formas o esquemas de conceptualización recurrentes que responden a motivaciones pragmáticas específicas (Cucatto, 2023)— destinados a conceptualizar el proceso judicial. En efecto, buscamos mostrar de qué manera el empleo de estos patrones metafóricos desplegados en las sentencias por parte de los operadores judiciales pone de manifiesto un modo de pensar metafórico, vinculado con rutinas comunicativas que nos hablan de una “mentalidad o cultura jurídica” (Cucatto, 2018, 2021) altamente normativizada —regida por la ley— y normalizada —reproducida por tradición—.

Para poder alcanzar nuestro propósito, este artículo se estructura de la siguiente manera: después de esta introducción, se desarrolla un breve marco teórico, se exponen algunos presupuestos metodológicos y el corpus de trabajo; posteriormente, se analizan algunos ejemplos extraídos de dicho corpus, a fin de establecer algunos patrones metafóricos empleados por los magistrados para conceptualizar un proceso judicial; por último, se ofrecen algunas consideraciones finales.

## **2. Marco conceptual. Sobre las metáforas y el discurso judicial**

Son muchos los autores que reconocen que la metáfora es un instrumento presente en los textos jurídicos (Bazzanella, Morra y Rossi, 2008; Berger, 2004; González Zurro, 2015; Hibbitts, 1994; Jumanca, 2013; Malem, 2006; Melo de Souza y Fontenelle Carneiro, 2020; Murray, 1984; Peerenboom, 2007; Santos, 2021; Twardzisz, 2008; Vegara Fabregat y Mansilla, 2017; Winter, 2003). Sin embargo, en algunas ocasiones, se recomienda evitar el uso de metáforas, en especial, en los decisorios (Malem, 2006), a

pesar de que estas puedan “arrojar luz sobre el desarrollo jurídico o un aspecto particular del desarrollo jurídico” (Peerenboom, 2007, p. 341).

Si bien el estudio de las metáforas en la discursividad jurídica ha comenzado a problematizarse desde las últimas décadas del siglo pasado, no ha sido hasta la emergencia de la teoría de la metáfora conceptual que contamos con un conjunto de herramientas potentes para abordarlas en sus dimensiones lingüísticas, de pensamiento, sociales y comunicacionales. En efecto, es a principios de los ochenta cuando se produce el llamado giro cognitivista (Steen, 2011) y la metáfora pasa a ser explorada en sus múltiples dimensiones, atendiendo a su relevancia en lo que respecta a la estructuración y simbolización del sistema conceptual.

Muy brevemente diremos que esta teoría fue formulada originalmente por Lakoff en colaboración con Johnson (1980) y con Turner (1989), y su hipótesis de partida consistió en asumir que “nuestro sistema conceptual es en gran medida metafórico, la manera en que pensamos, lo que experimentamos y lo que hacemos cada día también es en gran medida cosa de metáforas” (Lakoff y Johnson, 1986, p. 39). En este marco, la metáfora no es un simple adorno u ornamento, sino que es concebida como un mecanismo cognitivo que permite acceder a conceptos abstractos a partir de conceptos más concretos, simples y familiares. Asimismo, la metáfora hace posible que podamos entender y experimentar un tipo de cosa en términos de otra (Lakoff y Johnson, 1980); no es solo una cuestión de lenguaje, sino también cognitiva, ya que permite establecer vínculos entre dominios del sistema conceptual, de tal forma que una persona pueda conceptualizar el mundo al otorgar un sentido metafórico a su experiencia social (Lakoff, 1993). De este modo, la metáfora se basa en una proyección, correlación o *mapping* entre diferentes dominios conceptuales, desde el dominio conceptual de origen —fuente—, accesible a nuestra experiencia física o social, más concreto, hacia el dominio conceptual de destino —meta, diana—, más abstracto, que requiere ser explicado (Croft y Cruse, 2004; Dirven y Verspoor, 2004; Evans y Green, 2004; Johnson, 1987; Kövecses, 2010, 2015; Lakoff y Johnson, 1980, 1999; Lakoff, 1987, 1993; Soriano, 2012).

En nuestro caso, debemos destacar que durante el largo trayecto que representa —a veces— un proceso judicial, las personas que intervienen tienen que lidiar con situaciones, fenómenos y conceptos de gran nivel de abstracción y, en lo que se refiere a los magistrados, estos deben “motivar” —fundamentar— sus decisiones y hacerlas claras y comunicables. Igualmente, los operadores judiciales saben que “hacer” un proceso, comprender su naturaleza y dinámica, es una tarea ardua y compleja, no solo para los legos en cuestiones jurídicas, sino también para los profesionales del derecho. Todo esto, sin duda, convierte al proceso en un terreno fértil para el desarrollo de metáforas. Por ejemplo, en “los recurrentes vencidos” y “el carácter de vencedor del impugnante” es dable identificar la presencia de la metáfora *el proceso judicial es una guerra*, dado que algunos elementos convencionales del dominio origen *guerra* —como el ataque, la defensa, la victoria o la derrota, vencedores y vencidos (Lakoff y Johnson, 1980)— sirven para comprender un dominio meta, el proceso *judicial*. Más adelante podremos apreciar

de qué modo esta metáfora se erige como un patrón metafórico —o esquema metafórico recurrente— a la hora de conceptualizar la intrincada dinámica que se establece en un proceso judicial.

Por otra parte, resulta relevante considerar el lugar que ocupan las metáforas en la vida social y cultural. Al respecto, es factible identificar cuáles prevalecen en determinados entornos sociales o en ciertas culturas. Esta prevalencia de ciertas metáforas en determinados contextos demuestra, además, en qué medida estas impregnan el pensamiento y la acción en la vida cotidiana y cómo estas se caracterizan por su nivel de convencionalidad, sobre todo, en determinados entornos (Gibbs, 2017; Kövecses, 2015; Lakoff, 1993; Lakoff y Johnson, 1999; Steen, 2011), como puede ser, en nuestro caso, la Administración de Justicia. En este ámbito especializado puede verificarse lo que Kövecses llama “presión de coherencia”, esto es, las personas que participan en un intercambio comunicativo intentan ser coherentes con distintos aspectos contextuales, tales como el entorno físico y el contexto cultural y social (2009, p. 18). Esta “presión de coherencia” hace que los operadores judiciales recurran reiteradamente a ciertas formas de comunicar, en tanto modelos legitimados por una tradición judicial, en busca de cierto éxito comunicativo.

En este sentido, tal como hemos mostrado en investigaciones previas, el discurso judicial posee un alto nivel de “ritualismo” que se manifiesta en la estricta sujeción a las normas jurídicas procedimentales y en el empleo de ciertas formas de expresión o “patrones”, los cuales se repiten una y otra vez en las distintas actuaciones que se desarrollan dentro de un proceso judicial. En esta ocasión, hacemos hincapié en el valor que adquieren las metáforas en el ámbito de la práctica jurisdiccional y, particularmente, en el lugar que ocupan en las sentencias, entendiéndolo que su utilización responde a determinadas motivaciones pragmáticas que ponen en evidencia rutinas comunicativas que nos hablan de lo que hemos denominado una “mentalidad o cultura jurídica” (Cucatto, 2018, 2021).

### **3. Metodología y corpus de investigación**

A partir de los resultados producidos en investigaciones anteriores y por lo expuesto en el segundo apartado, hemos constatado que la Lingüística Cognitiva, especialmente la teoría de la metáfora conceptual, provee un marco teórico-metodológico apropiado para estudiar las metáforas presentes en las sentencias, uno de los géneros más representativos del proceso judicial en su conjunto. En este artículo procuramos abordar aquellos patrones metafóricos —en tanto formas o esquemas de conceptualización recurrentes que responden a motivaciones pragmáticas específicas— destinados a conceptualizar un proceso judicial. Asimismo, como expresamos al cierre de la sección previa, sostenemos que el empleo de dichos patrones metafóricos pone de manifiesto un modo de pensar metafórico, vinculado con rutinas comunicativas puestas en práctica cotidianamente por los operadores judiciales.

En esta dirección, esta propuesta se encuadra en una estrategia metodológica de carácter cualitativo, ya que apreciamos que el enfoque cognitivo cumple con los criterios propios de este tipo de investigación en ciencias sociales, es decir, la consideración de las personas y sus puntos de vista, los significados que estas construyen social y culturalmente en relación con sus prácticas (Valles, 1997; Vasilachis, 2006).

El corpus de referencia de nuestra investigación está conformado por una muestra intencional<sup>2</sup> de cien sentencias emitidas por diferentes órganos jurisdiccionales de la provincia de Buenos Aires (Argentina), durante el período 2000-2023. Estos decisivos se encuentran disponibles en la base de datos JUBA —<https://juba.scba.gov.ar/Busquedas.aspx>—, que reúne sentencias de distintas instancias, fueros y departamentos judiciales de esa provincia argentina.

Cabe aclarar que se ha seleccionado el género “sentencia” porque: 1) constituye uno de los géneros más característicos del discurso jurisdiccional; 2) es el acto procesal con el que normalmente finaliza un proceso judicial.

En cuanto al procedimiento empleado en nuestra investigación, se procedió del siguiente modo:

- Lectura exploratoria de las sentencias, a fin de detectar aquellas metáforas orientadas a conceptualizar un proceso judicial.
- Reconocimiento de las expresiones lingüísticas a través de las que se materializan esas metáforas.
- Determinación de los patrones metafóricos mediante los cuales los magistrados configuran el material lingüístico, generan diferentes esquemas recurrentes y le imprimen una forma metafórica/proyectiva.
- Identificación de las proyecciones conceptuales encontradas en esos patrones. Se distinguen los dominios conceptuales y se analiza la red de correspondencias que se establecen entre los dominios fuente/origen y meta/diana.
- Función/efecto de estos patrones metafóricos en relación con la manera de conceptualizar un proceso judicial.
- Correlación entre los patrones metafóricos, las motivaciones pragmáticas puestas de manifiesto en esos “usos metafóricos”, en vinculación con un modo de pensamiento metafórico y las rutinas comunicativas involucradas.

Por último, señalamos que los ejemplos fueron trabajados en su formato original, incluso con sus errores gramaticales, ortográficos o de tipeo, si existieran. Aclaramos que

---

2. En líneas generales se entiende por muestreo intencional a un tipo de muestreo determinístico, “aquel en el que todos los elementos muestrales de la población son seleccionados bajo estricto juicio personal del investigador” (Namakforoosh, 2002, p. 189). Como afirma Izcará Palacios “el método cualitativo, haciendo uso del muestreo intencional (no probabilístico) pretende profundizar en aspectos concretos de la realidad social, a través de la interpretación y análisis de los discursos de los actores sociales, sin perseguir la generalización de los resultados a la población general en términos de error calculable estadísticamente” (2007, p. 21).



las sentencias utilizadas son documentos disponibles públicamente en la base de datos JUBA, algunas anonimizadas —por afectar la intimidad de las partes o de terceros— y otras con indicación de la identidad de las personas implicadas.

#### 4. Patrones metafóricos y proceso judicial

En este apartado presentamos solo algunos patrones metafóricos a través de los cuales se conceptualizan —y comunican— aspectos relevantes de un proceso judicial. Estos patrones, que se repiten una y otra vez en las sentencias, nos permiten conocer cómo conciben el proceso judicial las personas que intervienen en él, esto es, hacen posible conocer lo que llamamos su pensamiento jurídico y, asimismo, ponen en evidencia rutinas comunicativas propias de la Administración de justicia.

##### 4.1. *El proceso judicial es un camino*

Se suele definir el proceso judicial como un conjunto de actos realizados por los jueces, por las partes y por los auxiliares del juez/a, en una serie gradual y progresiva, cuya finalidad es la emisión de una sentencia que resuelva un conflicto de intereses mediante la aplicación del ordenamiento jurídico vigente (Sosa, 2018).

En este sentido, es posible observar que el proceso judicial es conceptualizado tomando, como dominio fuente, el esquema de imagen de un camino.<sup>3</sup> Lakoff (1987) señala que nuestra concepción del movimiento está basada en este esquema de imagen y que es uno de los más comunes, dado que emerge de nuestro constante funcionamiento corporal, de nuestra experiencia y, por esta razón, sirve como dominio fuente de muchas metáforas. El autor afirma que sus componentes estructurales básicos son: un punto de partida, un punto de llegada o meta y, entre ambos, un trayecto o recorrido, además de elementos que facilitan y obstaculizan el recorrido. En el caso del proceso judicial, el punto de partida suele ser una demanda y el punto de llegada una sentencia definitiva, y, entre ambos puntos, pueden existir una serie de puntos intermedios, por ejemplo, otros actos procesales, contestación de demanda, recursos, entre otros. Por otra parte, es dable aclarar que las leyes procesales<sup>4</sup> son

---

3. Por una cuestión de espacio no pretendemos profundizar en esta temática, solo agregar que las imágenes esquemáticas se definen como esquemas básicos de base física y corporal que se incorporan a la cognición desde las primeras experiencias del niño en el mundo y resultan sumamente productivas para conceptualizar situaciones y procesos complejos a lo largo de la vida (Cuenca y Hilferty, 1999; Ibarretxe-Antuñano y Cadierno, 2019; Johnson, 1987; Lakoff, 1993, 2008).

4. Argentina es un país federal, en el que la nación, por un lado, y cada una de las provincias, por otro, tienen códigos de procedimientos propios y diferentes. Ver arts. 1 y 5 de la Constitución de la Nación Argentina, disponible en <http://servicios.infoleg.gob.ar/infolegInternet/anexos/0-4999/804/norma.htm>.

aquellas que regulan el tránsito generalmente desde la demanda hasta la sentencia definitiva; son las que indican el “*camino procesal apto*”:<sup>5</sup>

(1) Para que proceda la acción meramente declarativa debe existir, en primer lugar y siempre que no haya otro *camino* procesal apto, alguna incertidumbre respecto una relación de derecho entre partes y, luego, el peligro de que ello provoque un daño o perjuicio a quien la promueve.

CC0001 SI 20815, 17/10/2019, “Blandino Victorina Francisca José s/ acción declarativa”, p.1.

Si bien el *proceso judicial es un camino* “normativizado” por la ley, y, como consecuencia, “normalizado” en la práctica profesional, difícilmente el recorrido sea igual en todos los procesos judiciales. De este modo, en el transcurso de un proceso pueden surgir obstáculos dilatorios o de otro tipo que impidan o bloqueen su avance; es posible que surjan desvíos, encrucijadas, bifurcaciones, así como también pueden darse (o no) situaciones eliminadoras de barreras o facilitadoras de un tránsito más fluido:

(2) En tal sentido, debe rechazarse toda hermenéutica que *limite y/o cierre el camino* a la jurisdicción, por tratarse de una garantía que se erige en uno de los pilares básicos del estado de derecho [...].

SCBA LP A 73905, 10/07/2019, “Merlo Ocampo, María Rosaura c/ Municipalidad de La Plata sobre Amparo por Mora”, p. 8.

(3) Este vicio consiste en un notorio *desvío* de las reglas de la lógica que derive en una conclusión incoherente y contradictoria en el orden lógico formal o es incompatible con las constancias objetivas de la causa [...].

SCBA LP A 73881, 03/05/2018, “Gómez, Jorge Alejandro c/ Municipalidad de La Plata. Pretensión Declarativa de Certeza”, p. 4.

(4) Agrega que aun cuando ello no ocurriera, estas no deben ser objeto de condena al IOMA. De lo contrario, sería la propia ley la que pondría al referido organismo en una *encrucijada* insoluble [...].

SCBA LP A 73295, 30/03/2016, “R., S. M. y o. c/ M. d. S., s/ Amparo”, p. 5.

(5) La impugnación *se bifurca* en los siguientes agravios [...].

SCBA LP L 108252, 06/06/2012, “Attene Ramos, Edgardo y otros c/Astillero Río Santiago s/Diferencias salariales”, p. 2.

---

5. Deseamos aclarar que en los ejemplos solo abordaremos el patrón bajo análisis, más allá de la presencia de otras metáforas correspondientes a otros patrones metafóricos, como en (1) “daño”.

(6) Debe el recurrente realizar un esfuerzo adicional para explicar su interés en recurrir que permita *remover el obstáculo procesal para habilitar* la competencia de esta Corte en mejora de la situación ya obtenida [...]

SCBA LP P 134585, 24/10/2022, “Altuve, Carlos Arturo s/ recurso extraordinario de inaplicabilidad de ley”, p. 6.

Un “debido proceso” —un proceso judicial justo— es un derecho que tienen las personas a obtener una respuesta jurisdiccional válida y eficaz. Tienen derecho a iniciar y proseguir un proceso hasta su finalización, en cuyo recorrido han de tener derecho de defender lo que consideran sus derechos. Sin embargo, las personas que transitan ese camino suelen encontrarse con obstáculos que “*limite[n] y/o cierre[n] el camino a la jurisdicción*” (2): deben evitar “*desvío[s]*” (3), por ejemplo, cuando fundamentan sus pretensiones; tienen que alejarse de las “*encrucijada[s] insoluble[s]*” (4) que las aparten de la ley, así como también deben optar por trazar *bifurcaciones* (5) o evadirlas, a fin de “*remover obstáculo[s] procesal[es] para habilitar*” (6) la llegada a destino. En el caso del proceso judicial, el destino/meta es una sentencia definitiva justa.

#### 4.2. *El proceso judicial es una guerra*

Como ya mencionamos, un proceso judicial habitualmente surge ante intereses en conflicto. Los operadores judiciales conceptualizan metafóricamente el proceso judicial como una guerra, un verdadero campo de batalla en el que hay que defenderse o atacar, en el que se diseñan estrategias, tácticas, se apela a estocadas, se es herido o se hiere, se resiste, se pierde o se gana:

(7) En tal situación, el *ganador* no está potenciado para incoar un *embate contra* el fallo [...]. No obstante si el *vencido* lo *ataca*, todas las *defensas* planteadas por aquel quedan sometidas a la Cámara, y deben ser tratadas inexcusablemente por ella, como si hubiera habido una “adhesión” de quien resultó *ganancioso*.

CC0202 LP 127482, 01/09/2020, “Locio Sonia Beatriz C/ B., M. N. y otros s/ acción de despojo”, p. 15.

(8) Buena parte de la *estrategia* argumental de la apelante ha tenido como propósito evitar, a como dé lugar, que el reproche que se le formula sea analizado a la luz de las normas [...].

CC0102 MP 171893, 31/05/2022, “Maglione Paulo y otro c/ Farma Works SCS s/ daños y perjuicios por afectación a la dignidad”, p. 11.

(9) El agravio en su primera *estocada* intenta *dar embate* a la agravante del encuadre típico, utilizando como base argumentativa la inexistencia de “banda”, y el *ataque* se plantea en dos *flancos*.

TC0003 LP 18318, 17/04/2008, “A., M. M. y V., J. C. s/Recurso de casación”, p. 5.

(10) Se presenta una situación excepcional de incompatibilidad con el debido proceso derivada de la violación de un trámite esencial que *hiere* irremediamente al mismo.

SCBA LP P 75687, 20/08/2008, “R., C. s/Violación”, p. 4.

Si bien el fin último de un proceso judicial para las partes es llegar a una meta, no lo es a costa de cualquier desenlace. Las personas que intervienen en un proceso aspiran a lograr una conclusión favorable, a resultar *ganancios*[as] (7); para ello desarrollan *estrategia*[s] (8), se defienden, *ataca*[n] (7) y si es necesario *hiere*[n] (10), recurren a *embate*[s] (7), (9) o a *estocada*[s] (9). EL PROCESO JUDICIAL ES UNA GUERRA, una contienda en la que la jurisdicción debe, ante un conflicto de intereses, “*custodiar y proteger un bien jurídico*” (11):

(11) Como se ha afirmado, el respeto por el derecho de incidencia colectiva a la debida participación ciudadana previa a la decisión de asuntos medioambientales, no se reduce a la observancia de una mera formalidad. Se trata de *custodiar y proteger un bien jurídico relevante*.

SCBA LP A 75276, 24/06/2020, “Mancuso, Antonio c/ Organismo Provincial para el Desarrollo Sostenible OPDS s/ amparo”, p. 9.

En línea con lo expresado por Santibáñez y Osorio, “la metáfora es una construcción lingüística abstracta que tiene un uso también sobresaliente en contextos de conflicto o disputa” (Santibáñez y Osorio, 2011, p. 14); sin lugar a dudas, este patrón metafórico responde a determinadas motivaciones pragmáticas que ponen de manifiesto un modo de pensar metafórico, en el cual se conceptualiza el proceso judicial como una disputa procesal en la que se busca la paz por medio de una decisión judicial no arbitraria, justa.

### **4.3. El proceso judicial es una curación**

El proceso judicial es concebido como un acto terapéutico, curativo. Un proceso judicial que no constituye un tratamiento correcto, aquel indicado por la ley, puede ser entendido como *infectado*, y, como consecuencia, algún bien jurídico puede sufrir una *lesión*:

(12) Erróneo trastocamiento que *infecta* el estadio previo de admisibilidad con el final de juzgamiento sobre el mérito.

SCBA LP Rc 122757, 26/02/2020, “Cardano, Gustavo Fabián c/ Gigli, María Luisa Liliana Margarita s/ Daños y perjuicios”, p. 4.

(13) De todas formas, dado que la pena debe guardar proporcionalidad con la magnitud del contenido ilícito del hecho, o sea, con la gravedad de la *lesión* al bien jurídico concretamente afectado, el reclamo no trasciende.

SCBA LP P 131234, 19/02/2020, “Gallardo Neuch, Cristian Enrique; Iglesias, Gastón Rafael y Pino, Javier Alejandro s/ recurso extraordinario de inaplicabilidad de ley”, p. 8.

Asimismo, frente a un proceso judicial patológico existen actos procesales, por ejemplo, los recursos, que funcionan como un *remedio* al cual las partes acuden ante una sentencia adversa:

(14) En el *remedio* extraordinario interpuesto a favor de Lucía Derlis Mancuso, la defensa oficial planteó la nulidad absoluta de la sentencia de la Cámara por afectación de la garantía de defensa en juicio.

SCBA LP Rc 122757, 26/02/2020, “Cardano, Gustavo Fabián c/ Gigli, María Luisa Liliana Margarita s/ Daños y perjuicios”, p. 5.

Ahora bien, como indicamos, un proceso judicial por lo general tiene como objetivo resolver un conflicto de intereses. En este sentido, una decisión jurisdiccional representa no solo una meta a alcanzar, la paz tan ansiada, sino también EL PROCESO JUDICIAL ES LA CURACIÓN, puesto que es concebido como un mecanismo terapéutico cuyo resultado anhelado es un estado de sanación. Desde esta perspectiva, algunos actos procesales, según la ley o la jurisprudencia, son un medio idóneo para lograr tal sanación —*sanatoria*—, para corregir *defectos* o *vicios*:

(15) [...] el llamamiento de “autos para sentencia”, por aplicación de la directriz genérica que sienta el art. 133 del Código Procesal Civil y Comercial, se notifica ministerio legis; y [...] es efecto de ese llamamiento [...] que [...] todos los eventuales *defectos* o *vicios* de actividad anteriores pierden virtualidad. Que tal *sanatoria* reposa en [...] los principios básicos que campean en esta materia.

SCBA LP C 99748, 09/12/2010, “P. A., G. A. c/Y. D. H. s/Impugnación de paternidad-Reclamación de filiación”, p. 10.

#### **4.4. El proceso judicial es un ritual**

El proceso judicial puede ser conceptualizado como un ritual, un acto religioso en el que se sigue una serie de reglas establecidas por la ley. Así, un código procesal se denomina *código ritual*, dado que indica cómo se deben desarrollar cada uno de los actos/ceremonias que conforman dicho proceso. En este marco, las normas son consideradas casi sagradas, especialmente, las que configuran el “*sagrado* derecho de defensa” en juicio:

(16) El código *ritual* faculta al juez a dictar medidas ordenatorias e instructorias (art. 36 del CPCBA).

SCBA LP A 74496, 16/12/2020, “Malacalza, Carlos María y otros contra Caja de Retiros, Jubilaciones y Pensiones de la Policía de la Provincia de Buenos Aires sobre pretensión anulatoria”, p. 2.

(17) Es en función de esta clara alusión de la acusación privada al delito previsto y penado por el inciso 7.º del artículo 80 del Código Penal que creo que la violación incurrida por el *a quo* deviene formal y no cristaliza el daño concreto al *sagrado* derecho de defensa que denuncia la recurrente.

SCBA LP P 126435, 16/08/2017, “Silva Miguel Ángel, Moreno Carlos Fabián, López Luciano Mario Leonel, Juárez Carlos Jordán s/ recurso extraordinario de inaplicabilidad de ley”, p. 6.

EL PROCESO JUDICIAL ES UN RITUAL en el que un juez es un ser “*ungido* para impartir justicia”, cuya misión, entre otras, es velar para que se evite “*sacrificar* la tutela efectiva [...] en el *altar* del formulismo *ritual*”:

(18) Así, si bien no cabe avalar el desplazamiento del juez designado por la ley [...] quien fue *ungido* para impartir justicia.

SCBA LP 94657 I, 22/02/2006, “López, Fernando y otros c/Russo, Miguel Ángel y otros s/Daños y perjuicios”, p. 2.

(19) Este Tribunal no puede hacer caso omiso al error manifiesto y patente, advertido a partir de un detenido estudio de las actuaciones, y *sacrificar* la tutela efectiva [...] en el *altar* del formulismo *ritual*.

CC0202 LP, 121117, 07/09/2017, “De Micheli, Lucas Ricardo c/ Sánchez de Del Bueno Beatriz Cristina s/ resolución de contrato compra/venta inmuebles”, p. 6.

En efecto, un proceso judicial debe huir de un innecesario *exceso ritual*, el cual lo torne una *ceremonia cabalística* que lleve a las personas al “*altar* del formulismo *ritual*” (19):

(20) El tiempo transcurrido y los pasos procesales dados resultan un *exceso ritual* y un dispendio jurisdiccional, máxime que dichas medidas también se hallan previstas dentro del proceso de amparo.

SCBA LP, 73886, 29/11/2017, “Martín, Noelia L. y otros contra Municipalidad de la Costa. Acción de amparo”, p. 5.

(21) Todo el proceso debe volverse más humano, en el sentido de que se manifieste al pueblo no ya como una *ceremonia cabalística*, sino como un refugio accesible que el Estado ha puesto a disposición de todos los que creen en la justicia.

SCBA LP, 105958, 23/09/2015, “Rochman, Samuel y otros contra Valsecchi, Leonardo José. Incidente”, p. 6.

## 5. Consideraciones finales

Como hemos podido apreciar en el apartado anterior, los patrones metafóricos que hemos analizado nos han permitido mostrar cómo estos obedecen a determinadas motivaciones pragmáticas que ponen de manifiesto un modo de pensar metafórico de los operadores judiciales, esto es, el modo de conceptualizar (experimentar, entender y representar metafóricamente) un proceso judicial. Además, pudimos evidenciar que estos patrones, a su vez, generan y reflejan rutinas comunicativas altamente normativizadas —regidas por la ley— y normalizadas —pues se reiteran y se reproducen—.

De esta manera, un proceso judicial es concebido como: un *camino* complejo, repleto de obstáculos que deben ser sorteados para llegar a buen fin —una sentencia definitiva—; una *guerra* que surge ante un conflicto de intereses, una contienda en la que las personas involucradas no quieren perder; una sucesión de actos terapéuticos en los que se intenta, a través de remedios procesales, encontrar la *curación*; y un *ritual*, guiado por la jurisdicción, cuyas reglas están determinadas en los códigos procesales.

Es necesario aclarar que este trabajo no pretende agotar esta temática y que en nuestra investigación también hemos analizado otros patrones metafóricos, los que, en esta oportunidad, por una cuestión de espacio, no hemos podido desarrollar. Por ejemplo, el proceso judicial es conceptualizado como un espectáculo, como un juego, como un viaje, como un cálculo, como una negociación, como un tejido, como un diagrama. Los actos que componen un proceso, los principios y las normas son concebidos como personas, seres vivos, fenómenos de la naturaleza, terrenos u objetos.

Para finalizar, consideramos que no solo podrían formularse otros patrones, sino también que existiría cierta jerarquía entre estos —subpatrones metafóricos—.<sup>6</sup> En este sentido, creemos que cada una de estas cuestiones podría servir de disparador para nuevas y enriquecedoras investigaciones.

## Referencias

- Bazzanella, Carla, Morra, Lucía y Rossi, Piercarlo (2008). “L’approccio cognitivo alla metafora nel linguaggio giuridico”. En Caterina Raffaele (Ed.), *Fondamenti cognitivi del diritto* (p. 157-175). Pearson Paravia Bruno Mondadori.
- Berger, Linda (2004). What is the sound of a corporation speaking? How the cognitive theory of metaphor can help lawyers shape law. *Journal of the Association of Legal Writing Directors*, 2, 169-208.
- Croft, William y Cruse, David (2004). *Cognitive Linguistics*. Cambridge University Press.

---

6. Por ejemplo, del patrón metafórico *el proceso judicial es un camino*, dependería un subpatrón relacionado con la navegación; este subpatrón se observa en: “*llevar la causa a su puerto de destino*”.

- Cucatto, Mariana (2018): “Argumentación jurídica, obiter dicta y técnica recursiva: lo que abunda no daña”. En Helga Lell (Ed.), *Lenguaje y Derecho. Abordajes epistemológicos de una relación compleja* (p. 42-55). Marcial Pons-CONICET.
- Cucatto, Mariana (2021). Para mayor satisfacción (de): expresión conectiva plurifuncional, fundamentación y disuasión. *Círculo de Lingüística Aplicada a la Comunicación*, 88, 107-121. Disponible en <https://dx.doi.org/10.5209/clac.71982>
- Cucatto, Mariana (2023). *Sentencias judiciales, patrones metafóricos y comunicabilidad*. Conferencia presentada en el XII Encuentro Nacional de la Asociación Latinoamericana de Estudios del Discurso (Delegación Chile), Universidad de Las Américas, Santiago, Chile.
- Cucatto, Mariana y Rojas, Edgardo (2024): “Principales directrices metodológicas en Lingüística Cognitiva”. En Maria das Graças Soares, Luis Passeggi, Sulemi Fabiano Campos, Paulo Henrique Duque, Jose Romerito Silva (Eds.), *Metodologia da pesquisa em linguística teórica, descritiva e experimental* (pp.466-508). EDUFERN: Natal.
- Cuenca, Maria Josep y Hilferty, Joseph (1999). *Introducción a la lingüística cognitiva*. Ariel.
- Dirven, René y Verspoor, Marjolyn (2004). *Cognitive Exploration of Language and Linguistics*. John Benjamins.
- Evans, Vyvyan y Green, Melanie (2004). *Cognitive Linguistics. An Introduction*. Edinburgh University Press.
- Gibbs, Raymond (2017). *Metaphor Wars. Conceptual Metaphors in Human Life*. Cambridge University Press.
- González Zurro, Guillermo (2015). *Otra mirada a las decisiones de la Corte Suprema. Estudio metafórico desde la literatura, la corriente cognitiva y la imaginación*. New York University Press.
- Hibbitts, Bernard (1994). Making Sense of Metaphors: Visuality, Aurality and the Reconfiguration of the American Legal Discourse. *Cardozo Law Review*, 16, 229-356.
- Ibarretxe-Antuñano, Iraide y Cadierno, Teresa (2019). “La lingüística cognitiva y la adquisición de segundas lenguas ASL”. En Iraide Ibarretxe-Antuñano, Teresa Cadierno y Alejandro Castañeda Castro (Eds.), *Lingüística cognitiva y español LE/L2* (p. 19-51). Routledge.
- Ibarretxe-Antuñano, Iraide y Valenzuela, Javier (2021). *Lenguaje y Cognición*. Síntesis.
- Izcara Palacios, Simón (2007). *Introducción al muestreo*. Porrúa.
- Johnson, Mark (1987). *The body in the mind. The bodily basis of meaning, imagination and reason*. University of Chicago Press.
- Jumanca, Romanita (2013). Types of metaphors in the English Legal Discourse. *Romanian Journal of English Studies*, 9(1), 366-372.
- Kövecses, Zoltan (2009). “Metaphor, Culture, and Discourse: The Pressure of Coherence”. En Andreas Musolff y Jörg Zinken (Eds.), *Metaphor and Discourse* (p. 11-24). Nueva York: Palgrave Macmillan.



- Kövecses, Zoltan (2010). *Metaphor: A Practical introduction* (2.<sup>a</sup> Ed.). Oxford University Press.
- Kövecses, Zoltan (2015). *Where Metaphors Come From. Reconsidering Context in Metaphor*. Oxford University Press.
- Lakoff, George (1987). *Women, Fire and Dangerous Things*. University of Chicago Press.
- Lakoff, George (1993). "The contemporary theory of metaphor". En Andrew Ortony (Ed.), *Metaphor and thought* (p. 202-251). Cambridge University Press.
- Lakoff, George (2008). "The neural theory of metaphor". En Raymond Gibbs (Ed.), *The Cambridge handbook of metaphor and thought* (p.17-38). Cambridge University Press.
- Lakoff, George y Johnson, Mark (1980). *Metaphors We Live By*. University of Chicago Press.
- Lakoff, George y Johnson, Mark (1986). *Metáforas de la vida cotidiana* (Trad. Carmen González Marín). Cátedra.
- Lakoff, George y Johnson, Mark (1999). *Philosophy in the Flesh. The Embodied Mind and its Challenge to Western Thought*. New York: Perseus Books Group.
- Lakoff, George y Turner, Mark (1989). *More than Cool Reason: A Field Guide to Poetic Metaphor*. University of Chicago Press.
- Malem, Jorge (2006). El lenguaje de las sentencias. *Reforma judicial. Revista Mexicana de Justicia*, 7, 47-63.
- Melo de Souza, Lidiane y Fontenelle Carneiro, Mônica (2020). Metáforas conceptuais como ferramentas de argumentação e persuasão no discurso jurídico. *Revista de Direito, Arte e Literatura*, 6(2), 01-21.
- Murray, James (1984). Understanding law as metaphor. *Journal of Legal Education*, 34(4), 714-730.
- Namakforoosh, Mohammad (2002). *Metodología de la investigación*. Editorial Limusa.
- Peerenboom, Randall (2007). *China modernizes. Threat to the West or Model for the Rest?* Oxford University Press.
- Santibáñez, Cristián y Osorio, Jorge (2011). "Introducción: Lo metafórico". En Cristián Santibáñez y Jorge Osorio (Eds.), *Recorridos de la metáfora: mente, espacio y diálogo* (p. 11-29). Cosmigonon ediciones.
- Santos, Roberta (2021). *Argumentação é guerra: a metáfora no discurso jurídico*. Independently published: Amazon Kindle Edition.
- Soriano, Cristina (2012). "La metáfora conceptual". En Iraide Ibarretxe-Antuñano y Javier Valenzuela (Dirs.), *Lingüística Cognitiva* (p. 97-121). Siglo XXI Editores.
- Sosa, Toribio (2018). Composición del derecho al debido proceso. *Jurisprudencia Argentina*, 1(5), 147-163.
- Steen, Gerard (2011). The Contemporary Theory of Metaphor-now new and improved! *Review of Cognitive Linguistics*, 9(1), 26-64.
- Twardzisz, Piotr (2008). "Metaphors in Commercial Contracts". En Christina Alm-Arvius, Nils-Lennart Johannesson y David Minugh (Eds.), *Selected papers from the*

- Stockholm 2008 Metaphor Festival* (p. 237-253). Department of English, Stockholm University.
- Valles, Miguel (1997). *Técnicas cualitativas de investigación social: Reflexión metodológica y práctica profesional*. Síntesis.
- Vasilachis, Irene (2006). “La investigación cualitativa”. En Irene Vasilachis (Coord.), *Estrategias de investigación cualitativa* (p. 23-64). Gedisa.
- Vegara Fabregat, Laura y Mansilla, Ana (2017). Metáfora jurídica en la fraseología contrastiva en alemán y en español: el esquema de *camino*. *Alfinge*, 29, 223-241.
- Winter, Steven (2003). *A Clearing in the Forest: Law, Life and Mind*. University of Chicago Press.

# Las relaciones paradigmáticas de las colocaciones en los diccionarios monolingües de aprendizaje<sup>1</sup>

Yifan Zhang - Universidad de Changchun, Universidad de Sevilla  
zhangyifan5040@163.com

Rebut / Received: 10-03-24

Acceptat / Accepted: 21-07-24

**Resum. Les relacions paradigmàtiques de les col·locacions als diccionaris monolingües d'aprenentatge.** Les col·locacions tenen un paper fonamental en l'ensenyament i l'aprenentatge de l'espanyol com a llengua estrangera. Com una combinació restringida i característica, representa un desafiament significatiu per als estudiants no nadius. Tot i això, sovint els usuaris no troben les col·locacions que busquen, a causa de factors com la influència de les relacions paradigmàtiques, que poden portar a la substitució d'un dels components. En aquest treball, presentem les diverses relacions paradigmàtiques possibles (sinonímia, antonímia, hiperonímia, hiponímia, cohiponímia i meronímia) en les col·locacions. Mitjançant una anàlisi comparativa de les col·locacions formades per components amb aquestes relacions en diccionaris monolingües d'aprenentatge, s'observa que, a diferència dels diccionaris fraseològics, que registren explícitament aquesta informació, als diccionaris d'aprenentatge la seva inclusió varia segons l'amplitud i no té uniformitat. D'aquesta manera, el treball proposa procediments per optimitzar l'accés a les col·locacions en aquests repertoris lexicogràfics.

**Paraules clau:** relació paradigmàtica, relació semàntic, col·locació, diccionari d'aprenentatge.

**Abstract. Paradigmatic relations of collocations in monolingual learning dictionaries.** Collocations play a fundamental role in the teaching and learning of Spanish as a foreign

---

1. Esta investigación se enmarca en el proyecto "Estudio del tratamiento de colocaciones en diccionarios de aprendizaje de español" [JJKH20251086SK] financiado por el Departamento Provincial de Educación de Jilin.

language. As a restricted and characteristic combination, it represents a significant challenge for non-native learners. However, users frequently fail to find the collocations they are looking for, due to factors such as the influence of paradigmatic relations, which can lead to the substitution of one of the components. In this paper, we present the various possible paradigmatic relations (synonymy, antonymy, hyperonymy, hyponymy, cohyponymy and meronymy) in collocations. By means of a comparative analysis of collocations formed by components with these relations in monolingual learning dictionaries, it is observed that, unlike phraseological dictionaries, which explicitly record this information, in learning dictionaries their inclusion varies according to the extent and lacks uniformity. Thus, the paper proposes procedures to optimize access to collocations in these lexicographic repertoires.

**Keywords:** paradigmatic relationship, semantic relationship, collocation, learning dictionary.

## 1. Introducción

A pesar de que los estudios colocationales en el ámbito de la enseñanza de español como lengua extranjera (ELE) son relativamente recientes, siendo M. Seco (1987) el primer investigador en introducir el término *colocación*, esta unidad fraseológica es fundamental en la adquisición de lengua segunda o lengua extranjera. El conocimiento de las restricciones combinatorias de las palabras en una lengua separa las producciones del nativo de las del no nativo (Higuera García, 2006). Es común que los estudiantes realicen transferencias erróneas de patrones colocationales de su lengua materna al español. Además de los manuales, los diccionarios son una herramienta útil para conocer las colocaciones, las cuales también tienen una importancia vital en la dimensión lexicográfica (Cao, 2023; Ferrando, 2010; Handl, 2009; Kim, 2017; Koike, 2001; Laufer, 2011; Torner y Bernal, 2016).

Para los aprendices de nivel superior, en lugar de estudiar las colocaciones por separado, puede ser beneficioso familiarizarse con el concepto de las relaciones paradigmáticas aplicado a estas combinaciones. Esto contribuirá a enriquecer su conocimiento léxico y fraseológico, y además fomentará una mayor autonomía a la hora de crear combinaciones nuevas mediante sustituciones paradigmáticas. No obstante, como señala Corpas Pastor (1996), el tema de las relaciones paradigmáticas entre los elementos léxicos que pueden formar colocaciones no ha sido abordado de manera exhaustiva en la teoría colocational. Por eso, este artículo presenta un análisis cualitativo de la aparición y el tratamiento de las relaciones semánticas paradigmáticas de las colocaciones en los diccionarios monolingües de aprendizaje (DMA). El objetivo consiste en resumir las características propias y comunes del tratamiento de colocaciones con estas relaciones en estos repertorios lexicográficos con el motivo de ofrecer indicaciones a los usuarios y formular sugerencias para mejorarlo.

## 2. Consideraciones teóricas

Todavía no se ha llegado a un acuerdo consensuado del concepto de *colocación*. Generalmente se considera como una combinación compuesta por al menos dos componentes: una base *A* y un colocativo *B*, donde la base *A*, que mantiene su sentido literal, determina la selección del colocativo *B*, que adopta un sentido figurado o abstracto, pero el conjunto de ambos expresa un sentido unitario de carácter semitransparente (Corpas Pastor, 1996; Hausmann, 1989). En este sentido, los dos componentes de una colocación tienen posiciones desequilibradas, en las que un elemento determina el otro (Coseriu, 1977; Hausmann, 1979; Koike, 2001). Este análisis se enfoca en el nivel sintagmático del análisis, que corresponde al eje horizontal (Pazó y Travalía, 2007). Fue J. R. Firth (1957) quien contribuyó a la expansión del término y el que promovió el estudio de las relaciones sintagmáticas (Pérez Serrano, 2015). Aparte de este tipo de relaciones entre los componentes de las colocaciones, Bosque (2001) ha puesto de manifiesto que las relaciones colocacionales no se forman entre dos elementos aislados, sino entre clases léxicas, lo que implica a las relaciones paradigmáticas.

Las relaciones paradigmáticas de naturaleza semántica vinculan a cada uno de los constituyentes de la colocación con otras unidades léxicas de la lengua en el eje vertical. Estas relaciones han sido tratadas por otros autores, como Zuluaga (1980) o González y Ortega (2005), que describen variantes fraseológicas tales como series y sustituciones léxicas<sup>2</sup>. A su vez, en propuesta de Pazó y Travalía (2007), las series se consideran una categoría de sustitución léxica, al igual que cualquier variante dentro de una colocación. Las sustituciones se llevan a cabo en virtud de las relaciones semánticas paradigmáticas de los componentes de la colocación, que pueden incluir relaciones de sinonimia<sup>3</sup>, antonimia<sup>4</sup>, hiperonimia, hiponimia, cohiponimia<sup>5</sup>, holonimia<sup>6</sup>, meronimia<sup>7</sup> o comeronimia<sup>8</sup>.

---

2. Las series se producen al reemplazar un componente de una combinación por un antónimo, mientras que las sustituciones léxicas son producto del cambio de un componente por un sinónimo contextual o por un elemento con el que mantiene una relación jerárquica: hiperonimia, hiponimia, etc. (Pazó y Travalía, 2007).

3. La sinonimia se define como la relación existente entre dos o más unidades léxicas con distinto significante e igual significado.

4. La antonimia se produce entre dos palabras de significados opuestos.

5. El hiperónimo es el término que incluye a los demás; los términos incluidos son los hipónimos, mientras que la relación entre ellos es la cohiponimia.

6. Es la relación que una palabra tiene con otra cuando la primera denota el todo del cual la segunda es una parte. Por ejemplo, *coche* es un holónimo de *rueda*.

7. Es la relación inversa a la holonimia.

8. Es la relación entre dos o más palabras que son merónimas del mismo holónimo. En otras palabras, son partes del mismo todo. Por ejemplo, *rueda* y *volante* son comerónimos porque ambos son partes del *coche*.

Antes de adentrarnos en el análisis de las relaciones semánticas paradigmáticas dentro de las colocaciones en los diccionarios, es fundamental presentar primero la estructura formulada por Pazó y Travalía (2007). Según su propuesta, el primer elemento de la colocación, representado por  $X$ , está sintagmáticamente relacionado con el segundo elemento, representado por  $Y$ . Además, como afirman los dos investigadores,  $X$  también guarda relaciones paradigmáticas con  $[x1]$ ,  $[x2]$ , etc. Por otro lado,  $Y$  está paradigmáticamente relacionado con  $[y1]$ ,  $[y2]$ , etc. De esta manera, se puede deducir que, en una colocación determinada, es posible reemplazar tanto el componente base como el colocativo, obteniendo así en ambos casos combinaciones naturales de la lengua.

$$\begin{array}{l} X + Y \\ [x1] [y1] \\ [x2] [y2] \\ \cdot \quad \cdot \\ \cdot \quad \cdot \end{array}$$

En efecto, las colocaciones de distintas estructuras son susceptibles de aceptar sustituciones paradigmáticas, ya sea en casos de verbo + sustantivo o de sustantivo + adjetivo, así como en otros tipos menos productivos. A continuación, examinaremos un ejemplo de una colocación del primer tipo junto con algunas nuevas combinaciones formadas a partir de ella mediante sustituciones paradigmáticas.

En la combinación *cocinar el desayuno*, tanto *cocinar* como *desayuno* mantienen relaciones paradigmáticas con otras palabras, se puede sustituir tanto la base (*almuerzo*, *cena* ...) como el colocativo (*preparar*, *hacer* ...) para obtener combinaciones naturales de la lengua. Aquí se presentan algunas posibles variaciones:

cocinar el almuerzo  
cocinar la cena  
hacer el desayuno  
preparar el desayuno

Aunque autores como Coseriu (1977), Hausmann (1979) y Koike (2001) señalan que la base lleva al colocativo, observamos, en estos ejemplos, que el colocativo también lleva a otra base. Además de todo lo expuesto, se afirma que las relaciones semánticas paradigmáticas que posibilitan las sustituciones no se limitan a palabras específicas, sino que abarcan vocabulario de la misma índole. Esta idea se ve reflejada en la estructura del *Diccionario combinatorio del español contemporáneo (REDES)*. En esta obra, las palabras construyen unas redes léxicas complejas y múltiples entre sí, de modo que

---

9. Fórmula propuesta por Pazó y Travalía (2007).

un vocablo dado nos lleva a muchos otros y estos, a su vez, nos llevan a otras lexías, y así sucesivamente. Resumiendo lo dicho, estas redes complejas se rigen no solo por los vínculos léxicos de las palabras correspondientes al nivel sintagmático, sino también por los principios de la sinonimia, antonimia, hiponimia y meronimia, que corresponden al nivel semántico paradigmático.

### 3. Metodología

En este artículo, vamos a hacer un análisis comparativo de las colocaciones con estas relaciones semánticas paradigmáticas y sus presentaciones en los DMA. Para llevar a cabo esta tarea, primero construimos un corpus abierto de las colocaciones recopiladas en la obra *Colocaciones léxicas en el español actual: Estudio formal y léxico-semántico* de Koike (2001)<sup>10</sup>. Aunque contamos con más de 2500 colocaciones y una amplia variedad de unidades léxicas para cada tipo de relación mencionada, muchas de estas no son muy comunes. Por ello, hemos decidido recurrir al Corpus del Español del Siglo XXI (CORPES) para consultar la frecuencia de coaparición. Así, identificamos aquellas de alta frecuencia en nuestro corpus y las utilizamos para llevar a cabo una investigación descriptiva. En el futuro, una vez hayamos recopilado una cantidad considerable de colocaciones frecuentes de cada relación mencionada, procederemos con un análisis cuantitativo.

Nos basaremos en el análisis de dichas combinaciones en los cinco DMA publicados hasta la fecha, los cuales son representativos en el mercado español:

- *Diccionario de español para extranjeros para la enseñanza de la lengua española* (DIPELE, 2010).
- *Diccionario Salamanca de la lengua española* (DSLE, 1996).
- *Diccionario de español para extranjeros* (DEPE, 2002).
- *Diccionario de la lengua española para estudiantes de español* (DILEPEE, 2002).
- *Gran diccionario de uso del español actual* (GDUEA, 2001).

Procedemos a realizar búsquedas de su presencia de cada una de las colocaciones citadas en los artículos lexicográficos para comprobar si la colocación está presente en la microestructura de los diccionarios, ya sea en la definición, en los ejemplos o en la subentrada de la base o del colocativo.

---

10. Koike aborda en esta obra las combinaciones paradigmáticas según el tipo de palabra y relaciones semánticas paradigmáticas. En este artículo, debido a la limitación de espacio, solo hemos estudiado las colocaciones con el segundo caso.

#### 4. Las relaciones paradigmáticas de las colocaciones en los DMA

En este apartado, analizaremos respectivamente el tratamiento de las colocaciones con relaciones de sinonimia, antonimia, hiponimia y meronimia en los cinco DMA citados.

##### 4.1. Sinonimia

No se discuten aquí las colocaciones sinónimas que se forman con componentes completamente distintos, sino más bien aquellas que incluyen un componente sinónimo. Hemos identificado múltiples colocaciones que comparten un componente, mientras que el otro consiste en términos sinónimos. Cabe tener en cuenta que, en una estructura colocacional específica, es posible reemplazar tanto el elemento base como el colocativo, generando en ambos casos combinaciones lingüísticas naturales. Esta flexibilidad lleva a que la sinonimia se manifieste tanto en los elementos colocativos como en las bases utilizadas.

La mayoría de los sustantivos sinónimos o los que pertenecen al mismo campo semántico suelen colocarse con el mismo verbo (Írsula, 1992). Por ejemplo, en la colocación *apagar incendios*, podemos sustituir la base *incendio* por su sinónimo *fuego*. Además, dado que los verbos sinónimos tienden a combinarse con el mismo sustantivo, la sustitución paradigmática también se puede llevar a cabo en el componente colocativo. De este modo, podemos reemplazar *apagar* por *extinguir* en la colocación *apagar incendios*. A partir de estas relaciones, podemos crear las siguientes expresiones:

apagar incendios  
 apagar fuegos  
 extinguir incendios  
 extinguir fuegos

Existen otros ejemplos similares, los cuales hemos dividido en dos partes: sustitución de base o de colocativo.

##### 4.1.1. Sustitución de base

En este apartado, se seleccionan los ejemplos de sustitución de base en las colocaciones, que muestran alta frecuencia en el Corpus del Español del Siglo XXI (CORPES XXI). En las siguientes tablas, utilizamos las abreviaciones “C” (colocativo), “B” (base) y “A” (ambos) para indicar la ubicación de las colocaciones en los artículos correspondientes. A la derecha, se muestra la frecuencia total de aparición de cada colocación en los respectivos artículos de base y colocativo. Es importante señalar que una colocación puede aparecer tanto en la definición como en el ejemplo, en la subentrada o en otros elementos del artículo. En el artículo de *desempeñar* del DSLE, la colocación *desempeñar el papel* se registra en la definición, así como en el ejemplo:



**desempeñar** *v. tr.* 2. Realizar <una persona> [un papel, un trabajo o una función]:  
*Ha desempeñado el cargo de gobernador con dignidad. Desempeña muy bien el papel de caballero en la obra de Lope (desempeñar el papel, DSLE).*

En este caso, la marca “C/2” en la tabla indica la ubicación de esta colocación en el artículo del colocativo y se registra dos veces. En la tabla 1 se muestran más datos sobre los cinco DMA:

- (1a) desempeñar el papel
- (1b) desempeñar la función
  
- (2a) fruncir el ceño
- (2b) fruncir las cejas
  
- (3a) vencer un obstáculo
- (3b) vencer una dificultad

TABLA 1. PRESENTACIÓN DE VARIAS COLOCACIONES FORMADAS POR BASE SINÓNIMA EN LOS DMA

	Fre. <sup>11</sup>	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
desempeñar el papel	120	C/1	C/2	C/2	0	C/2
desempeñar la función	68	C1	B/1	B/2	A/2	0
fruncir el ceño	1456	A/3	A/3	B/2	A/2	A/3
fruncir la ceja	88	A/2	0	C/1	C/1	B/1
vencer un obstáculo	258	0	C/1	C/2	C/1	0
vencer una dificultad	179	0	C/1	C/1	C/1	0

Como se observa en esta tabla, los DMA muestran preferencia por recopilar las colocaciones más frecuentes, pero este no es el único criterio. Por ejemplo, en vez de optar por la colocación *desempeñar el papel*, el *DILEPEE* registra la alternativa tanto en el artículo de la base como en el del colocativo. Además, el *DSLE*, el *DEPE* y el *DILEPEE* presentan mayor posibilidad de incluir ambas colocaciones con significados sinónimos. El *DEPE* recoge todas las colocaciones listadas en la tabla 1, con un total de 10 registros, mostrando una frecuencia más elevada de estas unidades léxicas en comparación con

11. La frecuencia de las colocaciones es consultada en el Corpes XXI.

los otros diccionarios. Esto se debe a su uso de definición impropia, que indica el complemento argumental mediante fórmulas como *dicho de*, *referido a*, *aplicado a*, entre otras. Sin embargo, la combinación *fruncir el ceño*, que tiene una alta frecuencia en el CORPES XXI (1456 registros), se encuentra registrada en ambos artículos de todos los DMA citados, excepto el *DEPE*.

En lo que respecta a la asignación de estas unidades léxicas, las cinco obras lexicográficas muestran una tendencia a incluirlas en el artículo del colocativo. De esto se puede deducir que estos DMA realizan primordialmente una función descodificadora, tal como han señalado diversas investigaciones (Alonso Ramos, 2002; Castillo Carballo, 2002).

#### 4.1.2. Sustitución de colocativo

La sustitución de colocativo no es tan común como en el caso de la base, pero hemos hallado los siguientes ejemplos:

(4a) enarbolar la bandera

(4b) izar la bandera

TABLA 2. PRESENTACIÓN DE *ENARBOLAR BANDERA* E *IZAR BANDERA* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
enarbolar la bandera	219	C/1	A/3	C/2	C/1	C/2
izar la bandera	281	C/2	A/2	C/2	A/2	A/3

Basándonos en los datos presentados en la tabla 2, se evidencia que los cinco DMA tienden a favorecer la recuperación de colocaciones sinónimas cuando ambas son frecuentes. Además, el *DSLE* y el *GDUEA* prefieren recopilar esta unidad léxica tanto en el artículo del colocativo como en el de la base. Algunos de los sinónimos tienen una similitud extrema, mientras que otros comparten una acepción idéntica. Aquí distinguimos los sinónimos totales de los parciales. Los sinónimos totales, también conocidos como perfectos, son palabras que tienen significados idénticos independientemente del contexto en el que se usan. Pueden ser reemplazadas por otras en cualquier contexto sin modificar el sentido de la oración. Por ejemplo: *baile* y *danza*. Los sinónimos parciales o aproximados, en cambio, comparten el mismo significado en un cierto contexto, pero también tienen matices o connotaciones ligeramente diferentes. Por eso, no siempre son intercambiables de manera perfecta. Un ejemplo sería *tomar* y *beber*.

En el primer caso, no es necesario enumerar todas las colocaciones que contienen un componente sinónimo, ya que los estudiantes pueden realizar sustituciones de manera intuitiva. Por el contrario, en el segundo caso, es crucial resaltar todas las posibilidades de combinación subrayando la diversidad de matices y connotaciones que pueden surgir, ya que no se puede deducir uno directamente de otro.

De los ejemplos aducidos puede deducirse que no se ha alcanzado un acuerdo en lo que respecta al tratamiento de las colocaciones compuestas por sinónimos. Sin embargo, en el *DEPE*, hemos encontrado en la misma acepción la colocación *erizar el pelo* y *erizar el cabello*. En concreto, en la definición se emplea el sustantivo *pelo*, mientras en el ejemplo aparece su sinónimo *cabello*.

**erizar** v. Referido esp. al pelo, levantarlo o ponerlo rígido: *El pánico me erizó el cabello* (*erizar el pelo /erizar el cabello, DEPE*).

Esta práctica no solo evita el abuso de entradas léxicas en el diccionario, sino que también brinda a los estudiantes más oportunidades para enriquecer su vocabulario. Por lo tanto, este tratamiento del *DEPE* nos parece más adecuado. Además, Koike (2001) indica que los adjetivos sinónimos también presentan un fenómeno similar, ya que en algunos casos pueden modificar al mismo sustantivo y en otros no pueden hacerlo. No los presentamos aquí con más detalles debido a la limitación de espacio.

Según Gómez Molina (2004), resulta muy útil distinguir los casos en que varias unidades léxicas no sinonímicas llegan a indicar un valor sinonímico, lo cual da lugar a una neutralización semántica. Por ejemplo, los verbos *dar*, *pegar*, *emitir* o *soltar* no son sinónimos, pero *dar un grito*, *pegar un grito*, *emitir un grito*, *soltar un grito* actúan como sinónimos. En nuestro corpus existe un caso (4a) y (4b) que debemos tener en cuenta:

- (5a) error craso
- (5b) error garrafal

Como muestra la tabla siguiente, en este caso, solo el *DIPELE* no da cabida a ambas colocaciones. En efecto, las palabras *craso* y *garrafal* no son sinónimos. Según el *DRAE*<sup>12</sup>, tienen significados diferentes: *craso* significa ‘indisculpable’, mientras que *garrafal* significa ‘muy grande’. Por lo tanto, aunque pueden estar relacionados en ciertos contextos, no son términos intercambiables y no son sinónimos. De este modo, sería apropiado incluir ambas en un mismo diccionario, dado que los estudiantes extranjeros no pueden deducir una colocación directamente de la otra.

TABLA 3. PRESENTACIÓN DE *ERROR CRASO* Y *ERROR GARRAFAL* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
error craso	230	C/2	C/2	C/2	C/2	A/3
error garrafal	196	0	C/1	C/1	C/1	B/2

12. *Diccionario de la lengua española*, 23.<sup>a</sup> ed., [versión 23.7 en línea].

## 4.2. Antonimia

A la hora de introducir las colocaciones antónimas, los diccionarios tampoco presentan uniformidad, pues dentro de la misma obra podemos encontrar ejemplos en que aparecen bajo la misma subentrada y otros donde no. Este último procedimiento, además de aumentar el número de subentradas, no permite al usuario establecer referencias entre las diferentes expresiones relacionadas semánticamente por sinonimia o antonimia, puesto que suelen ocupar lugares distantes dentro del cuerpo del artículo al seguir el orden alfabético.

En el *GDUEA* se puede observar un ejemplo. En el primer caso, *rotunda victoria* y *rotundo fracaso* aparecen en la misma definición del artículo *rotundo*, usando una barra (/) para separar las dos opciones sinónimas. En el segundo caso, *conservar la calma* y *perder la calma* están registradas por separado en la subentrada ordenadas alfabéticamente.

**rotundo/a** 2. Se aplica a la acción, como *una victoria, una derrota* o algo similar, que supone un claro balance a favor o en contra de los protagonistas: *El gobierno de coalición del ministro japonés sufrió una rotunda derrota.* (*rotunda victoria/fracaso, GDUEA*).

*calma* Loc Conservar/mantener uno la calma, (en una situación difícil) no ponerse uno nervioso, mantener la capacidad de pensar y de actuar racionalmente: *Mantuvo la calma hasta que llegaron los bomberos y le rescataron.* [...] *Perder la calma*, ponerse uno nervioso en una situación difícil o no actuar como corresponde: *Suele mantenerse sereno cuando todos los demás han perdido la calma.* (*conservar la calma / perder la calma, GDUEA*).

### 4.2.1. Verbos antónimos

Koike (2001) sostiene que la manera en que los verbos antónimos se comportan también depende del significado del sustantivo con el que se combinan. Asimismo, los distintos niveles de uso metafórico del verbo ejercen influencia en la selección de los antónimos. De esta forma, se generan nuevas conexiones de antonimia. En consecuencia, la selección de los antónimos se convierte en uno de los mayores desafíos a vencer para los estudiantes extranjeros.

En el caso de *imponer sanción* y *levantar sanción*, solo la primera está registrada en el *DIPELE* sin indicar explícitamente la posibilidad de su opuesto. Sin embargo, este tipo de combinaciones de palabras no siempre permite deducir fácilmente su antónimo, especialmente para los hablantes no nativos del español. Por lo tanto, sería adecuado considerar la inclusión de ambas colocaciones en el diccionario.

(6a) imponer una sanción

(6b) levantar una sanción

TABLA 4. PRESENTACIÓN DE *LEVANTAR SANCIÓN* E *IMPONER SANCIÓN* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
levantar una sanción	331	0	0	0	0	0
imponer una sanción	126	B/1	0	0	0	0

Por el contrario, tanto la colocación *abrir ventana* como *cerrar ventana* están presentes en todos los diccionarios analizados, y el *DEPE* incluso cuenta con cinco registros en este aspecto. No obstante, para los estudiantes, puede resultar más fácil inferir una de la otra.

(7a) abrir la ventana

(7b) cerrar la ventana

TABLA 5. PRESENTACIÓN DE *ABRIR VENTANA* Y *CERRAR VENTANA* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
abrir la ventana	2111	A/2	B/1	A/3	C/1	A/2
cerrar la ventana	720	C/1	C/1	C/2	C/1	A/2

Ni la colocación *colgar teléfono* ni *descolgar teléfono* se registra en el *DIPELE*, que no registra ninguna acepción relativa, a diferencia de las otras cuatro obras. Por ejemplo, en el *DSLE* sí existe una acepción específica de este sentido.

**colgar** tr. 4. Interrumpir <una persona> la comunicación telefónica con [otra persona]:  
*Me colgó Inés el teléfono (colgar el teléfono, DSLE).*

En los otros tres DMA, se encuentran acepciones similares, y estas dos combinaciones de palabras se agrupan bajo la entrada del verbo *colgar* y su antónimo *descolgar*. Dado que ambos términos son antónimos completos y existe una referencia cruzada entre ellos, no es necesario incluir ambas por separado.

(8a) colgar el teléfono

(8b) descolgar el teléfono

TABLA 6. PRESENTACIÓN DE *COLGAR TELÉFONO* Y *DESCOLGAR TELÉFONO* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
colgar el teléfono	942	0	A/2	C/1	C/2	C/2
descolgar el teléfono	368	0	C/2	C/2	C/2	C/2

A partir de los ejemplos extraídos de nuestro corpus, observamos que los cinco DMA coinciden en recopilar más colocaciones formadas por verbos antónimos. Aparte de eso, también se registran sustantivos y adjetivos antónimos, aunque su presencia no es tan significativa en términos de cantidad.

#### 4.2.2. Sustantivos antónimos

Koike (2001) toma unos ejemplos de colocaciones formadas con *éxito* y *fracaso*, en las que el verbo o el adjetivo puede aparecer con sustantivos antónimos: *éxito/ fracaso clamoroso*, *rotundo éxito/ fracaso*, *éxito/ fracaso estrepitoso*, *completo éxito/ fracaso*, *tener un éxito/ fracaso*, etc. No obstante, no todos los sustantivos antónimos comparten el mismo colocativo al formar la colocación. Por ejemplo, el adjetivo *apoteósico* solo se coloca con *éxito* (*éxito/ \*fracaso apoteósico*) y solo *éxito* puede aparecer con el verbo *alcanzar* (*alcanzar el éxito/ \*fracaso*). Se puede deducir que algunos verbos y sustantivos solo pueden combinarse con sustantivos de significado o bien positivo o bien negativo.

De este modo, resulta crucial indicar en el artículo lexicográfico si el colocativo puede combinarse exclusivamente con bases de significado positivo o negativo. En el siguiente ejemplo, se explica de forma explícita su valor positivo.

**alcanzar** v. 5. Llegar a tener <una persona> [una cosa que busca o desea]: *Alcancé el objetivo que pretendía*. SIN. Lograr (*alcanzar el objetivo*, DSLE).

**alcanzar** 5. Conseguir una cosa que se desea: *al fin alcanzó la alcaldía* (*alcanzar un puesto*, DIPELE).

#### 4.2.3. Adjetivos antónimos

Los antónimos varían en función del sustantivo con el que están vinculados. Así, *voz blanca*, *ropa blanca*, *vino blanco* o *pan blanco* cuentan con antónimos diferentes (Aguilar-Amat, 1993). Los adjetivos *adulto* y *grave* (en *voz adulta* y *voz grave*, respectivamente) actúan como antónimos de *blanco* en el contexto de *voz blanca*. Por otro lado, *tinto* y *oscuro* (en *vino tinto* y *vino oscuro*) se oponen a *blanco* en la colocación *vino blanco*. Lo mismo ocurre con el adjetivo *claro*, cuyo significado varía dependiendo del objeto al que se refieran. Por ejemplo, se contrasta el *agua clara* con el *agua turbia*, reflejando una dicotomía entre la pureza y la opacidad. Asimismo, la distinción entre una *explicación clara* y una *explicación confusa* resalta la importancia de la claridad en la comunicación. Además, se observa la diferencia entre un *cielo claro* y un *cielo nublado*, sugiriendo un contraste entre la luminosidad y la falta de claridad atmosférica.

Debido a la autonomía semántica del sustantivo y a la dependencia semántica del adjetivo con respecto al sustantivo con el que se combina, no sería preciso hablar de *adjetivos antónimos absolutos*, sino de *antónimos contextuales*, para reflejar las múltiples relaciones de antonimia que adquieren los adjetivos (Bosque, 1985). Los adjetivos *alto* y

*bajo* no siempre establecen la relación de antonimia como se observa en ejemplos como *alta/baja calidad*, *alta/\*baja tecnología* y *altas/\*bajas horas de la noche*.

Koike (2001) señala que, en las colocaciones sustantivo-adjetivo, los adjetivos que magnifican o intensifican el significado del sustantivo producen más colocaciones sinónimas y antónimas que los demás. Por ejemplo: *frío extremo / acusado / crudo / punzante / penetrante / reinante*. Al contrario, los adjetivos que expresan el grado bajo de *frío* pueden ser *leve / moderado / ligero / delicado*, etc. La complejidad de esta cuestión, como hemos explicado, es una de las razones por las que debemos centrarnos en ella cuando compilamos un diccionario de aprendizaje. Sin embargo, según el siguiente ejemplo, ninguno de los DMA citados incluye ambas colocaciones antónimas.

(9a) herida mortal

(9b) herida leve

TABLA 7. RECOPIACIÓN *HERIDA MORTAL* Y *HERIDA LEVE* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
herida mortal	312	C/1	C/1	0	0	C/1
herida leve	2	0	0	A/2	0	0

### 4.3. Hiperonimia, hiponimia y cobiponimia

Los hiperónimos y los hipónimos pueden sustituir un constituyente de la colocación en ciertos contextos para evitar la repetición de una palabra. Es fundamental comprender la sustitución hiperhiponímica. De este modo, cuando el usuario consulta el diccionario, puede ampliar su léxico y formar más expresiones. En este apartado, investigamos qué prefieren usar los DMA: los hiperónimos o hipónimos, y en qué lugar aparecen.

#### 4.3.1. Sustitución hiperhiponímica

El término *colocaciones de sustitución hiperhiponímica* hace referencia a combinaciones de palabras en las cuales una palabra hiperónima, es decir, una palabra que abarca a otras en términos de significado más amplio, puede ser reemplazada por una palabra hipónima, que es más específica dentro de la misma categoría, sin que la estructura colocacional cambie significativamente. Del mismo modo, también es posible realizar la sustitución en sentido inverso, donde una palabra hipónima puede ser sustituida por su palabra hiperónima dentro de la misma estructura colocacional sin que se produzca un cambio sustancial en el significado.

Por ejemplo, en la colocación *romper relación* y *romper amistad*, *relación* actúa como un hiperónimo de *amistad*. Eso implica que *amistad* es una categoría más específica que está incluida en el término más general *relación*. En este ejemplo, ambos términos pueden

emplearse de manera intercambiable en la colocación sin que la estructura gramatical o el significado experimenten un cambio sustancial. La sustitución hiperhiponímica en este caso permite flexibilidad en el uso de la colocación, manteniendo la coherencia en el significado.

Una vez comprendida la sustitución hiperhiponímica, comprobamos su tratamiento en los DMA. Aunque en la definición lexicográfica hay una mención a un hiperónimo, generalmente prefieren emplear sus hipónimos en los ejemplos. Por ejemplo, el verbo *trazar* se aplica a una línea, pero, en realidad, en los ejemplos, en lugar de línea, suelen aparecer sus hipónimos, como puede verse en los casos siguientes:

**trazar** v. tr. 1. Dibujar <una persona> [líneas]: *La niña ya traza palotes en la escuela (dibujar líneas, DSLE).*

**trazar** 1. tr. [algo] Hacer líneas; dibujar: *la profesora trazó una parábola en la pizarra (trazar una parábola, DIPELE).*

**curva** f. 1. Línea que no es recta ni tiene ángulos: *Para dibujar una «U» trazó una curva en la pizarra (trazar una curva, DILEPEE).*

#### 4.3.2. Sustitución de cohipónimos

En nuestro corpus, se encuentran algunos casos de cohiponimia, tal como se evidencia en la tabla 8.

TABLA 8. FRECUENCIA Y APARICIÓN EN LOS DMA  
DE LAS COLOCACIONES FORMADAS CON COHIPONIMIA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
celebrar la misa	614	B/1	A/3	C/1	C/2	C/1
celebrar la asamblea	373	0	B/1	0	0	0
celebrar la junta	116	0	0	0	A/2	0
causar admiración	116	0	B/1	0	B/1	B/1
causar asombro	148	0	B/1	0	0	0
causar conmoción	256	0	B/1	0	B/1	B/1
causar molestia	346	A/2	C/1	0	C/1	B/1
coger el autobús	261	A/2	C/1	0	C/1	C/1
coger el taxi	382	B/1	0	0	0	0



sentir dolor	2852	A/2	0	0	0	B/1
sentir miedo	2640	0	0	0	B/1	B/1
sentir amor	1547	A/2	0	0	C/1	0
sentir vergüenza	1159	B/1	0	0	0	B/1
tomar el taxi	1058	0	0	0	C/1	0
tomar el autobús	350	A/2	B/1	C/1	B/1	B/1
tomar el avión	741	0	0	0	0	C/1
tomar el desayuno	563	A/2	B/1	0	A/2	A/2
tomar la cena	155	0	0	B/1	0	0

A partir de la información proporcionada en la tabla anterior, inferimos que la inclusión o exclusión de una combinación en el diccionario no se determina únicamente por su frecuencia. Por ello, no existe un criterio común de selección de estas unidades en los cinco diccionarios. Así que en ocasiones la combinación registrada en el artículo lexicográfico no corresponde a lo que suele aparecer en la práctica. Sobre esta cuestión, Koike (2001) argumenta que los usos reales se imponen a la descripción lexicográfica. Además, en este caso pocas unidades se registran en ambos artículos lexicográficos o en diferentes posiciones de la misma entrada, en comparación con las unidades de sustitución de sinónimos<sup>13</sup>. Asimismo, los datos de la tabla muestran una notable dispersión, reflejando una mayor diversidad en la selección de cohipónimos y una menor restricción entre los dos elementos de este tipo de colocaciones.

En las colocaciones del tipo verbo + sustantivo (en función de sujeto) también hay muchos casos de cohipónimos. Por ejemplo, *trinar* {*pájaro: ruiseñor/ jilguero/ canario*}, *cantar* {*aves: canario/ gallo/ pájaro; algunos insectos: cigarra*}, *balar* {*el carnero/ el cordero/ la oveja/ la cabra/ el gamo/ el ciervo*}. Cabe señalar que, en los dos primeros casos, *pájaro* y *aves* son hiperónimos de las palabras que les siguen. En cuanto a este caso de sustitución hiperhiponímica a la que hace referencia Koike (2001), el *DEPE* lo ha gestionado más adecuadamente, porque incluye el hiperónimo en la definición y los hipónimos en los ejemplos. Así, los usuarios de la obra conocerán con qué clase léxica se puede combinar y encontrarán ejemplos concretos, lo que evita la repetición innecesaria. En este diccionario, por ejemplo, hemos encontrado, en el artículo *marchitar*, la colocación *marchitar* {*la planta*} y *marchitar* {*la rosa*}. En concreto, en el ejemplo se usa el sustantivo *rosa*, mientras que en la definición aparece el hiperónimo de *rosa*: *planta*.

13. Véase 4.1.1.

**marchitar** v. 1. Referido esp. a una planta, hacerle perder la frescura, el verdor o la abundancia de hojas: *El sol ha marchitado las rosas. Las lechugas se marchitaron por falta de riego (marchitar las flores, DEPE).*

Al final, es importante destacar que no todas las combinaciones permiten sustituir uno de sus elementos por otra unidad léxica con rasgos semánticos similares. Es decir, a pesar de la frecuente similitud semántica entre los elementos de una serie léxica, los cohipónimos relacionados con el mismo hiperónimo no siempre comparten la misma extensión colocacional. Por ejemplo, *conferencia* se puede colocar con *pronunciar*, pero *ponencia* no (Corpas Pastor, 1996). Eso supone un desafío para el aprendiz extranjero. Por eso, los repertorios lexicográficos deben hacer mención especial a los casos donde no se puede sustituir el elemento por sus cohipónimos, ayudándole a comprender las diferencias en el uso de palabras similares y sus colocaciones específicas en el lenguaje.

#### 4.4. Meronimia

Para afirmar que, en cuanto a las colocaciones compuestas por hipónimos, sus hiperónimos pueden aparecer en su lugar (sustitución hiperhiponímica), Koike (2001) ofrece el ejemplo de *apacentar ovejas* y explica que la palabra *ovejas* se puede sustituir por su hiperónimo *rebaño* dentro de la misma colocación: *apacentar el rebaño*. Sin embargo, Travalía (2006) señala que *rebaño* no es el hiperónimo de *ovejas*, sino su holónimo. *Oveja* no es un tipo de *rebaño*, sino una parte del mismo. Los lexemas *rebaño* y *ovejas*, por tanto, no mantienen una relación de hiperonimia-hiponimia, sino de meronimia. Específicamente, se trata de una relación de *meronimia homogénea*, dado que cada miembro es igual. Pone de relieve Moreno Cabrera (1997, p.161): “en español tenemos construcciones como *un rebaño de ovejas*, mediante las cuales construimos una meronimia homogénea en la cual cada una de las ovejas es un *miembro* del rebaño de ovejas”.

Por lo tanto, las *relaciones de hiperonimia e hiponimia* que existen entre sustantivos colectivos e individuales, tales como *enjambre-abeja*, *clero-sacerdote*, *dentadura-diente*, *flota-barco*, en semántica constituyen construcciones de meronimia, dado que el segundo no es un tipo del primero, sino que tienen una relación entre los individuos y el colectivo. La palabra que señala la totalidad (*rebaño*, *racimo*, *comunidad*) se denomina *holónimo* y la palabra que corresponde a la parte (*ovejas*, *uvas*, *vecinos*, respectivamente) se llama *merónimo* (Travalía, 2006).

Eso nos recuerda una parte de las colocaciones del tipo de *sustantivo + de + sustantivo*, tales como *enjambre de abejas*, *ramo de flores*, en las que dos componentes reflejan la relación entre el individual y el colectivo.

TABLA 9. RECOPIACIÓN DE *RAMO DE FLORES* Y *ENJAMBRE DE ABEJAS* EN LOS DMA

	Fre.	DIPELE	DSLE	DEPE	DILEPEE	GDUEA
ramo de flores	966	C/1	A/2	C/1	C/2	C/1
enjambre de abejas	73	C/1	C/1	C/1	C/1	C/2

Con la tabla presentada, se evidencia que las colocaciones de este tipo están documentadas en el artículo del colocativo (solo en el caso del *DSLE* la colocación *ramos de flores* figura tanto en el artículo del colocativo como en el de la base).

**ramo** m. 1. Manojó de flores, hierbas y ramas, que se forma natural o artificialmente: *Le regalaron un - de rosas por su cumpleaños (DILEPEE).*

**ramo** 3. Conjunto natural o artificial de flores, ramas o hierbas: *le envió un - de felicitación (DIPELE).*

**ramo** s.m. 1. Conjunto de flores, ramas o hierbas cortadas y unidas: *un ramo de claveles, un ramo de rosas (DSLE).*

**ramo** s/m 1. Conjunto de ramas o flores cortadas, dispuestas de forma artística, que se coloca en algún lugar con fines ornamentales o que se entrega a alguien como obsequio: *Puso el ramo de flores en un florero y le invitó a pasar (GDUEA).*

**ramo** 1 Conjunto o manojó de flores, ramas o hierbas: *Me han regalado un ramo de rosas (DEPE).*

De estos registros encontrados en los DMA, podemos deducir que este tipo de colocaciones suelen estar ubicadas en la definición, ya que *ramo* se relaciona con un conjunto de flores, mientras que *flor* se refiere a una sola unidad de la planta. En el sentido semántico, el colocativo (*ramo*) es un conjunto de la palabra como base (*flor*). Así pues, en su explicación, no podemos pasar por alto este elemento fundamental.

Cabe mencionar que algunas sustituciones llevarán a combinaciones poco naturales en el idioma meta. Con el fin de prevenir que los aprendices realicen sustituciones que generen combinaciones poco usuales, es de suma importancia que los diccionarios y manuales ofrezcan orientación y tratamiento específico en este aspecto. Para lograr esto, debe incluir múltiples ejemplos de uso en contexto para cada palabra. Esto ayudará a los aprendices a ver cómo se usan las palabras en oraciones reales y qué combinaciones son comunes. Además, debe agregar notas de uso y advertencias sobre combinaciones incorrectas o poco comunes. Estas notas pueden explicar por qué ciertas palabras no se combinan de la manera que un aprendiz podría esperar. También se pueden incluir secciones dedicadas a las colocaciones comunes de cada palabra. Esto puede ir al

final de la entrada de cada palabra. Por último, hay que proveer sinónimos junto con explicaciones de las diferencias en su uso y colocación. Esto ayudará a los aprendices a entender cuándo y cómo usar cada palabra correctamente.

## 5. Conclusiones

Las colocaciones, unidades fraseológicas no prototípicas por el menor grado de fijación e idiomatización, a diferencia de otros tipos de unidades lingüísticas plurales, son más móviles en los artículos lexicográficos. En este artículo hemos analizado cómo se abordan las relaciones semánticas paradigmáticas de colocaciones en los DMA y hemos observado que, a diferencia de los diccionarios fraseológicos, en los que esta información se registra de forma explícita mediante ciertos procedimientos específicos, su inclusión en los diccionarios de aprendizaje depende de su amplitud y carece de uniformidad.

Por lo tanto, sería pertinente establecer los principios para su tratamiento. En concreto, en relación con los sinónimos y antónimos, consideramos que no es necesario incluir todos los totales o perfectos; en cambio, tiene sentido registrar aquellos parciales o aproximados en el diccionario. Esto se debe a que es fácil deducir otras frases en el primer caso, mientras que obviamente el segundo presenta dificultades para los estudiantes de lengua extranjera. En cuanto a los hiperónimos e hipónimos, si ambos términos son naturales en la lengua, podrían compilarse en secciones distintas dentro del artículo. En particular, proponemos ubicar el hiperónimo en la definición y los hipónimos en el ejemplo. De esta manera, se ofrecen más opciones a los usuarios y, al mismo tiempo, se ahorra espacio. Con respecto a las colocaciones formadas por componentes de relación de meronimia, deberían registrarse en la definición del colocativo. De esta manera, no solo define la colocación, sino que también explica cómo y por qué estas palabras se combinan, ayudando a los usuarios a entender que una palabra se refiere específicamente a un uso individual de la otra (el conjunto).

Con el motivo de incluir más opciones de colocaciones con relaciones semánticas paradigmáticas, es importante destacar que el *REDES* y el *PRÁCTICO (Diccionario combinatorio práctico del español contemporáneo)* pueden ser de gran utilidad para la incorporación de más colocaciones en los DMA de español. Estos dos diccionarios combinatorios son útiles en la medida que nos dan una lista de las posibilidades combinatorias. Aparte de estas sugerencias que se fijan más en el tratamiento de las colocaciones en la microestructura, sería más apropiado añadir indicaciones en los prolegómenos sobre el tratamiento de dichas relaciones para beneficiar a los usuarios.

Con todo eso, no solo ofrecería más posibilidades de combinación a quienes consultan el diccionario, sino que también se establecería un enfoque coherente para tratar las relaciones semánticas y las colocaciones, lo cual es útil para usuarios que buscan un entendimiento más profundo y sistemático del idioma. Dado que el corpus no es cerrado, podríamos realizar un análisis cuantitativo en el futuro con más datos añadidos para completar el estudio de las relaciones semánticas paradigmáticas de las colocaciones en los DMA.

## Referencias

- Aguilar-Amat, Anna (1993). *Las colocaciones de nombre y adjetivo. Un paso hacia una teoría léxico-semántica de la traducción*. (Tesis doctoral no publicada). Universidad Autónoma de Barcelona.
- Alonso Ramos, Margarita (2002). Colocaciones y contorno de la definición lexicográfica. *Lingüística Española Actual*, 24(1), 63-96.
- Alvar Ezquerro, Manuel (dir.). (2010). *Diccionario de español para extranjeros para la enseñanza de la lengua española*. Vox-Universidad de Alcalá de Henares.
- Bosque, Ignacio (1985). Usos figurados de los adjetivos que denotan dimensiones físicas. *Philologica hispaniensa: in honorem Manuel Alvar*, 63-80. Gredos.
- Bosque, Ignacio (2001). Sobre el concepto de “colocación” y sus límites. *Lingüística Española Actual*, 23(1), 9-40.
- Bosque, Ignacio (dir.). (2004). *Redes. Diccionario combinatorio del español contemporáneo*. SM.
- Bosque, Ignacio (dir.). (2006). *Diccionario combinatorio práctico del español contemporáneo*. SM.
- Cao, Dung Thi Phuong (2023). Online Collocation Dictionary in L2 Writing: How Learners Use and Perceive Its Effectiveness. *Journal of Language Teaching and Research*, 14(1), 108-120.
- Castillo Carballo, María Auxiliadora (2002). El aspecto colocacional en la producción lexicográfica. En Pérez Pascual, José Ignacio y Campos Souto, Mar (Eds.), *Cuestiones de lexicografía*, 97-105. Tris Tram.
- Corpas Pastor, Gloria (1996). *Manual de fraseología española*. Gredos.
- Coseriu, Eugenio (1977). *Principios de semántica estructural*. Gredos.
- Ferrando Aramo, Verónica (2010). Hacia un diccionario de colocaciones para estudiantes de ELE. En García Platero, Juan Manuel y Castillo Carballo, María Auxiliadora (Eds.), *Investigación lexicográfica para la enseñanza de las lenguas*, 165-184, Universidad de Málaga.
- Firth, John Rupert (1957). Modes of meaning. En John Rupert Firth, *Papers in Linguistics 1934- 1951*, 190-215. Londres: Oxford University Press.
- Gómez Molina, José Ramón (2004). Las unidades léxicas del español. *La enseñanza del léxico en español como segunda lengua/lengua extranjera*, 27-50. SGEL.
- González Aguiar, María Isabel y Ortega Ojeda, Gonzalo Damián (2005). En torno a la variación de las unidades fraseológicas. En Almela, Ramón, et al. (Eds.), *Fraseología contrastiva: con ejemplos tomados del alemán, español, francés e italiano*, 91-110, Universidad de Murcia.
- Gutiérrez Cuadrado, Juan (dir.). (2006 [1996]). *Diccionario Salamanca de la lengua española*. Santillana.

- Handl, Susanne (2009). Towards collocational webs for presenting collocations in learners' dictionaries, en *Researching collocations in another language: Multiple interpretations*, 69-85. Palgrave Macmillan UK.
- Hausmann, Franz Josef (1989). Le dictionnaire de collocations. *Wörterbücher. Dictionaries. Dictionnaires. Ein internationales Handbuch zur Lexikographie*, 1010-1019. Walter de Gruyter.
- Hausmann, Franz Josef (1979). Un dictionnaire des collocations est-il possible? *Travaux de littérature et de linguistique de l'Université de Strasbourg*, 17(1), 187-195.
- Higueras García, Marta (2006). *Estudio de las colocaciones léxicas y su enseñanza en el español como lengua extranjera*, ASELE, Colección Monografías (9), Málaga.
- Írsula Peña, Jesús (1992). Colocaciones sustantivo-verbo. En Wotjak, Gerd (Ed.), *Estudios de lexicografía y metalexicografía del español actual*, 159-167. Max Niemeyer.
- Kim, Sugene (2017). Minimum input, maximum output, indeed! Teaching collocations through collocation dictionary skills development. *Lexikos*, 27, 265-286.
- Koike, Kazumi (2001). *Colocaciones léxicas en el español actual: estudio formal y léxico-semántico*. Servicio de publicaciones de la Universidad de Alcalá de Henares.
- Laufer, Batia (2011). The contribution of dictionary use to the production and retention of collocations in a second language. *International Journal of Lexicography*, 24(1), 29-49.
- Maldonado González, Concepción (dir.) (2002). *Diccionario de español para extranjeros*. SM.
- Moreno Cabrera, Juan Carlos (1997). *Introducción a la lingüística: enfoque tipológico y universalista*. Síntesis.
- Pazó, José y Travalía, Carolina (2007). Las relaciones paradigmáticas en la enseñanza de las colocaciones en ELE. En Monroy Casas, Rafael y Sánchez Pérez, Aquilino (Eds.), *Congreso Nacional de Lingüística Aplicada: 25 años de lingüística en España*, 395-402, Universidad de Murcia.
- Pérez Serrano, Mercedes (2015). *Un enfoque léxico a prueba: efectos de la instrucción en el aprendizaje de las colocaciones léxicas* (Tesis doctoral). Recuperado de <http://hdl.handle.net/10366/128255>
- Sánchez Pérez, Aquilino (dir.). (2001). *Gran diccionario de uso del español actual*. SGEL.
- Seco, Manuel (1987). Problemas formales de la definición. En *Estudios de lexicografía española*, Madrid: Paraninfo, 15-34.
- Torner, Sergi y Bernal, Elisenda (2016). Collocations in learner's dictionaries. En Torner, Sergi y Bernal, Elisenda (Eds.), *Collocations and Other Lexical Combinations in Spanish* (sec. 2, cap. 8, p. 157-172). Routledge.
- Travalía, Carolina (2006). Las colocaciones implícitas. *Estudios de Lingüística*, 20, 317-332.
- Villar, Celia (dir.). (2002). *Diccionario de la Lengua Española para Estudiantes de Español*. Espasa Calpe.
- Zuluaga, Alberto (1980). *Introducción al estudio de las expresiones fijas*. Peter Lang.

## Informació sobre *Sintagma* / Information about *Sintagma*

*Sintagma* és una revista de lingüística que pretén afavorir l'intercanvi de coneixements en aquest àmbit, independentment de la teoria en què s'emmarquin els treballs. La revista resta oberta a qualsevol contribució en els diversos camps de la lingüística, tant teòrica com aplicada. S'accepten estudis sobre qualsevol llengua o varietat lingüística, així com treballs d'anàlisi interlingüística o comparativa, sempre que es tracti de contribucions originals. *Sintagma* apareix regularment cada any des de 1989 en un sol número. A més de l'edició en paper, la revista també presenta una edició digital, disponible a <http://www.sintagma-online.udl.cat>. Les llengües en què es poden presentar els originals són: l'anglès, el castellà, el català i el francès.

*Sintagma* is a journal of linguistics that aims at promoting the exchange of knowledge in this area, regardless of any theoretical framework. *Sintagma* accepts studies in both theoretical and applied linguistics, as well as works on any language or language variety, and contrastive linguistics. Contributions must be original and unpublished. The journal is a yearly publication founded in 1989. The journal is published on paper and is also available online at <http://www.sintagma-online.udl.cat>. *Sintagma* accepts manuscripts written in Catalan, English, French and Spanish.

### *Sintagma* està indexada a / *Sintagma* is indexed in:

#### **Bases de dades / Databases:**

Arts & Humanities Citation Index  
Bibliography of Linguistic Literature (BLL)  
Cabell's International  
CAPES  
Communication & Mass Media  
Complete (EBSCO)  
CSA Linguistics and Language  
Behavior Abstracts (LLBA)  
DIALNET  
Directory of Open Access Journal (DOAJ)  
Francis  
ISOC (CSIC)  
Linguistic Abstracts  
Linguistic Bibliography/  
Bibliographie Linguistique  
MLA International Bibliography  
Periodicals Index Online  
RACO (Revistes Catalanes amb Accés Obert)  
Red Iberoamericana de Innovación y  
Conocimiento Científico (REDIB)  
SCOPUS  
Social Sciences Citation Index

#### **Recursos d'avaluació / Evaluation Resources:**

Carhus Plus  
CIRC  
DICE (CSIC)  
ERIH PLUS  
FECYT-546/2021 Excelente 7<sup>a</sup> convocatoria  
IN-RECH  
JCR (ISI)  
Latindex  
MIAR  
Scimago-SJR

#### **Cercadors científics / Research search engines:**

Google acadèmic / Google Scholar  
Scirus

## **Informació per als autors**

- Els autors han d'enviar els originals en format electrònic a l'adreça revista.sintagma@udl.cat.
- S'han d'ometre referències directes dels autors al llarg del text i cal indicar el nom i la filiació d'aquests en un document a part, ja que els manuscrits seran avaluats anònimament per dos revisors experts en la temàtica de cada article.
- Els autors poden demanar excloure possibles revisors per conflicte d'interessos.
- Al cap de quatre mesos com a màxim es comunica als autors la decisió motivada de l'acceptació o el rebuig del treball presentat i se'ls envien els informes.
- Els treballs han d'ocupar un màxim de 5.000 mots (inclosos la bibliografia, les taules i els gràfics).
- Cal seguir les normes d'estil que es publiquen en el web de la revista (<http://www.sintagma.udl.cat>).

## **Information for authors**

- Authors should submit the original manuscripts in electronic format to revista.sintagma@udl.cat.
- Direct references that may help identify the author should be omitted from the text and name and affiliation of authors will be indicated in a separate file to ensure anonymous reviewing by two experts in the topic of each manuscript.
- Authors can ask to exclude any possible reviewers in order to avoid any conflict of interest.
- Authors will be informed of the result of the anonymous peer review process within approximately four months after reception of the original manuscripts.
- Papers should not exceed 5,000 words (including references, tables and figures).
- Authors should follow the norms of publication available at the website of the journal (<http://www.sintagma.udl.cat>).



## Revisors / Reviewers 2025

*Sintagma* agraeix la col·laboració de les persones que van participar com a revisors en el volum 37.  
*Sintagma* thanks the help of all the people who acted as reviewers in volume 37.

Agustín-Llach, Pilar	Universidad de la Rioja
Ainciburu, M. Cecilia	Universidad Nebrija
Bernal, María	Universitat d'Estocolm
Bisiada, Mario	Universitat Pompeu Fabra
Boone, Griet	Ghent University
Cabello, Manuel	Universidad de Huelva
Cañete, Paola	Universidad de Concepción
Checa, Irene	University of Wyoming
Cuesta, Azahara	Universitat de Barcelona
Díez, Belén	Universidad de Jaén
Duque, Eladio	Universidad Complutense Madrid
Fernández-Montraveta, Ana	Universitat Autònoma de Barcelona
Freixa, Judit	Universitat Pompeu Fabra
Frumuselu, Anca	Universitat de Lleida
Gesa, Ferran	Universitat de Barcelona
González, Beatriz	Nottingham University
González, Daniela	Universidad Nacional Cuyo
Guerrero, Susana	UMA
Gutiérrez-Colón, Mar	Universitat Rovira i Virgili
Hernández, Patricia	Universidad de Buenos Aires
Hidalgo, Raquel	Universidad Complutense Madrid
José Fernández, Nicolás	Universidad de Jaén
López Samaniego, Ana	Universitat de Bardelona
López, Sonia	Universidad de la Laguna
Malicka, Alexandra	Universitat Oberta de Catalunya
Marimon, Mireia	Universitat Pompeu Fabra
Marqués-Pasqual, Laura	University of California
Martínez-Adrián, María	Universidad del País Vasco
Pena, Azucena	Universidad Autónoma de Madrid
Perián, Carlos	Universitat Politècnica de València
Plotnitova, Anastasia	University of Groeningen
Sáez, Patricia	Université Lyon 2
Vázquez, Gloria	Universitat de Lleida
Veliz, Leonardo	Universidad Católica Silva Henríquez

- 7-29 Marinov Vranješ, S., Kabalin Borenić, V.: *Methodological issues in adapting and applying a lexical error taxonomy* / Aspectes metodològics per a l'adaptació i aplicació d'una taxonomia d'errors lèxics
- 31-49 Lena, L., Liu, Y.: *When number is not neutral: Translating Chinese bare nouns into English* / Quan el número no és neutre: Traduir noms escuets xinesos a l'anglès
- 51-66 Matea Kovač, M.: *The importance of corrective feedback in foreign language learning* / La importància del feedback correctiu en l'aprenentatge de llengües estrangeres
- 67-82 Felices Lago, Á.: *Towards the characterization of WORK problem schemas in the DIAPASON Ontology* / Cap a la caracterització dels esquemes de problemes de TREBALL a l'ontologia DIAPASON
- 83-98 Fernández Sanmartín, A.: *Los actos de amenaza a la imagen en la entrevista sociolingüística: un estudio comparativo entre entrevistas libres y entrevistas estructuradas* / Face-threatening acts in the sociolinguistic interview: a comparative study between free and structured interviews
- 99-114 Cucatto, M.: *Patrones metafóricos y proceso judicial: conceptualización y comunicación* / Metaphorical patterns and judicial process: conceptualization and communication
- 115-134 Zhang, Y.: *Las relaciones paradigmáticas de las colocaciones en los diccionarios monolingües de aprendizaje* / Paradigmatic relations of collocations in monolingual learning dictionaries