

# Visual representations and their pedagogical effectiveness in teaching the Spanish verbal system: a pre-experimental validation study

SANDRA GUERRERO GARCÍA (CORRESPONDING AUTHOR)

*University College Dublin*

FRANCISCO JOSÉ MANJÓN POZAS

*Universidad de Granada*

Received: 2025-05-21 / Accepted: 2025-10-13

DOI: <https://doi.org/10.30827/portalin.viXV.33794>

*Porta Linguarum* ISSN paper edition: 1697-7467, ISSN digital edition: 2695-8244

**ABSTRACT:** This study, part of the IMAGINE project (PID2021-128771OB-I00), explores the pedagogical effectiveness of iconic-figurative images in teaching Spanish verb tenses as a foreign language. Grounded in cognitive grammar, it examines how visual representations—ranging from figurative to schematic—support native speakers in accurately identifying tense, mood, and aspect, especially depending on discourse contextualization. Visuals are presented as symbolic tools that bridge intuitive knowledge with explicit grammatical concepts. The study compares two test types: a narratively contextualized one (CONTEST) and a metalinguistic schematic version (METATEST), administered with and without explicit instruction. Sixty-one native speakers participated, divided into two groups. Results show the narrative context significantly aids in verb form selection by reducing cognitive load and enhancing inference. In contrast, more abstract or decontextualized tasks, like the METATEST without instruction, posed greater challenges, particularly for complex tenses such as the conditional perfect. The findings suggest that combining visuals with narrative frameworks and well-designed explicit instruction can improve L2 grammar learning. Additionally, the study emphasizes the need to rigorously validate visual materials before applying them in educational settings.

**Keywords:** aspect and manner, cognitive grammar, iconic-figurative images, Spanish as L2, tense, verbal system, iconic-figurative images.

## **Representaciones visuales y su eficacia pedagógica en la enseñanza del sistema verbal español: un estudio de validación pre-experimental**

**RESUMEN:** Este estudio, enmarcado en el proyecto IMAGINE (PID2021-128771OB-I00), explora la eficacia pedagógica de las imágenes icono-figurativas en la enseñanza de los tiempos verbales del español como lengua extranjera. Basado en la gramática cognitiva, examina cómo las representaciones visuales —desde las más figurativas hasta las más esquemáticas— ayudan a los hablantes nativos a identificar con precisión el tiempo, modo y aspecto verbal, especialmente en función del grado de contextualización discursiva. Las imágenes se presentan como herramientas simbólicas que vinculan el conocimiento intuitivo con conceptos gramaticales explícitos. El estudio compara dos tipos de prueba: una narrativamente contextualizada (CONTEST) y otra de carácter esquemático metalingüístico (METATEST), administradas con y sin instrucción explícita. Participaron sesenta y un hablantes nativos, distribuidos en dos grupos. Los resultados muestran que el contexto narrativo facilita sig-

nificativamente la selección de formas verbales al reducir la carga cognitiva y favorecer la inferencia. En cambio, las tareas más abstractas o descontextualizadas, como el METATEST sin instrucción, presentaron mayores dificultades, especialmente con tiempos complejos como el condicional compuesto. Los hallazgos sugieren que la combinación de imágenes con marcos narrativos y una instrucción explícita bien diseñada puede mejorar el aprendizaje gramatical en L2. Además, el estudio subraya la necesidad de validar rigurosamente los materiales visuales antes de su aplicación en entornos educativos

**Palabras clave:** gramática cognitiva, español L2, sistema verbal, tiempo, aspecto y modo, imágenes icono-figurativas

## 1. INTRODUCTION

This study corresponds to a pre-experimental validation phase aimed at assessing the pedagogical potential of metalinguistic images related to Spanish verb tenses (Castañeda 2023). The main objective is to explore to what extent native speakers recognize figurative and iconic resources as appropriate representations of tense values (present, past, and future), mood (indicative and conditional —Alarcos Llorach 1994—), and aspect (such as the imperfective/perfective or perfect/aorist oppositions). In addition, the study examines how the presence or absence of a narrative context influences participants' performance, with special attention to the degree of abstraction of the images—more figurative or more schematic—depending on the level of detail in the discursive framework in which they are embedded.

This exploratory phase functions as a pilot study, seeking to refine the test materials and procedures prior to more extensive validation. In this context, regarding the descriptive part of this work, we consider the recognition of the intended tense, mood, and aspectual values by most participants (i.e., over 50%) as a preliminary indicator of pedagogical potential. While this threshold does not establish statistical validity in a strict sense, it serves as a working benchmark to assess whether the visual representations are intuitively accessible to native speakers. Results falling significantly below this mark would prompt a revision of the images or the contextual framing, whereas consistently high agreement would justify advancing to further testing stages.

To this end, two evaluation instruments were used: a narratively contextualized test (CONTEST) and another based on metalinguistic iconography (METATEST). These instruments aim, on the one hand, to determine whether the presence of a narrative-discursive framework influences the selection of the appropriate verb tense and, on the other hand, to analyze whether explicit instruction on metalinguistic elements facilitates this choice compared to a test without instruction.

The degree of correlation between the tested images and native speakers' linguistic intuition is a key element for the design of more effective teaching materials, which can later be validated in quasi-experimental studies with L2 Spanish learners, thus enabling evaluation of their real impact on grammatical learning.

## 2. THEORETICAL FRAMEWORK

This study justifies the choice of cognitive grammar (within cognitive linguistics) as the theoretical framework (Geeraerts 2006; Langacker 1987, 1990, 2019; Ibarretxe-Antuñano

et al. 2019) to figuratively and contrastively describe Spanish verb tenses for pedagogical purposes, highlighting its advantages over other approaches. In particular, we wish to emphasize that this model acknowledges the symbolic value of morphosyntactic elements and allows grammatical differences to be represented as variations in the way the same scene is conceptualized. This facilitates both cross-linguistic comparison and the creation of teaching materials based on perceptual and iconic principles accessible to learners of different languages.

Another key feature of cognitive linguistics and grammar—especially useful for this work—is the understanding of linguistic categories as complex and accessible at different levels of abstraction, thanks to various cognitive processes. Thus, linguistic knowledge involves managing both general schemas and routines tied to specific contexts. One of the goals of this study is to establish the most appropriate level of abstraction to create an effective and useful didactic description for teaching verb tenses through iconic-figurative representations.

This proposal recognizes that figurative representation of grammatical distinctions varies depending on the case: some allow for clear and general images, while others require more specific descriptions. In the test validation phase, one aim is to determine the most suitable level at which these representations are comprehensible to native speakers.

An additional reason why cognitive grammar is the most appropriate model for our pedagogical proposal is its usage-based approach, in which linguistic units are understood as signs whose meaning only materializes in specific discursive and social contexts. For this reason, the study presents results based on the combined use of two types of iconic-figurative tests, which differ primarily in the degree of contextualization and the presence of metalinguistic information.

### **2.1. Iconic-figurative images in L2 teaching**

The grammatical categories of tense, mood, and aspect represent one of the greatest challenges in acquiring a foreign language due to their abstract nature, their dependence on the discursive context, and typological differences between languages. Studies such as Andersen and Shirai (1994) indicate that learners tend to acquire prototypical uses of verb tenses first, which can lead to overgeneralization of forms. Bardovi-Harlig (2000) and Bardovi-Harlig et al. (2024) for her part, highlights the verbal aspect, particularly in Spanish, presents additional challenges due to its close relationship with the internal structure of the event and its narrative interpretation, making it one of the most complex areas even for advanced learners.

According to several studies on cognitive linguistics applied to language teaching, the use of iconic-figurative images in foreign language learning seems to simplify the understanding and memorization of complex grammatical meanings (Littlemore 2009; Tyler 2012; Niemeier 2017). This assertion is supported by numerous studies, especially regarding tense, mood, and aspect. According to Paivio's dual coding theory (1986), the combination of verbal and visual stimuli significantly improves the comprehension and retention of abstract concepts. This idea is reflected in proposals such as those by De Knop and Dirven (2008), who emphasize that images allow for the visual representation of dynamic action features such as duration or repetition. Similarly, Boers et al. (2008) demonstrate how visual metaphors help illustrate temporal and aspectual differences, for example, contrasting "open" and "closed"

actions. Finally, applied studies such as those by Castañeda (2023) and Salaberry (2000, 2018) show that the use of metalinguistic images, such as diagrams or pictograms, helps L2 learners understand and conceptualize verb tense values, reducing interference from their mother tongue and improving access to verb forms in use.

Building on this background, and considering the inherent difficulties of these complex categories in the L2 acquisition process, the use of iconic-figurative images not only facilitates the visualization of abstract concepts but also allows students to more easily integrate grammatical values into discursive situations, promoting deeper understanding and greater long-term retention.

## 2.2. Content validity

In the field of foreign language teaching, content validity, construct validity, and concurrent validity are essential notions for evaluating the relevance and reliability of assessment instruments, such as exams, diagnostic tests, or performance tasks. These concepts, originating in psychometrics and educational evaluation theory, are used to ensure that tests actually measure what they are intended to assess.

Authors such as Anastasi and Urbina (1997), Messick (1989), Weir, C. J. (2005) or Weir C. J. and O'Sullivan, B. (2011) have focused on determining whether the included items adequately and sufficiently represent the content universe that defines the domain of interest. In this study, this principle is applied to the validation of didactic materials, specifically images designed to illustrate verb tenses in foreign language teaching.

Ensuring content validity in language assessments is essential for the instruments used to appropriately represent the educational objectives, language competencies, and grammatical structures to be assessed. A test with solid content validity enables accurate and coherent evaluation of student performance, as its items directly correspond with the content taught and the expected skills. Consequently, this type of validity not only ensures a fair evaluation process but also informs key pedagogical decisions, such as feedback and curriculum design.

In this context, a pre-validation procedure for teaching resources is proposed. This procedure involves submitting materials for evaluation by native speakers before their final implementation to confirm that the representations offered by the images align with the speakers' linguistic intuition. This prior verification ensures that the materials are consistent, relevant, and intuitive for the end users—non-native speakers of Spanish as a second language—thus reinforcing their content validity in a rigorous and substantiated manner. Accordingly, this study focuses on the use of two image-based validation instruments: a test with a context and another test with metalinguistic elements, to analyze native speakers' performance in a controlled context.

## 3. METHODOLOGY

This study falls within a pre-experimental validation phase of images designed to illustrate verb tenses in foreign language teaching, with the aim of analyzing the impact of visual support and narrative context on the identification and selection of Spanish verb forms by native speakers. This means confirming that visual representations (scenes, actions, situations) are relevant and faithful to the linguistic construct (the use and meaning of verb

tenses) intended to be taught or assessed, and that they align with the linguistic intuition of native speakers. This validation helps identify the strengths and weaknesses of the materials, optimize their design, and understand how context or instructions influence the interpretation of abstract representations. Native speakers play a key role in verifying whether the images align with their linguistic intuition.

The primary objectives of the study were to evaluate whether native speakers correctly recognize the temporal, modal, and aspectual values through the images and to determine if the presence of a narrative context influences the correct selection of verb tenses. The hypothesis posits that explicit instruction combined with a narrative context significantly improves the accurate selection of verb tenses.

To achieve these objectives, a mixed-methods design was adopted, combining qualitative and quantitative approaches using the SciPy package for Python. The goal was to explore both the performance and the perceptions of participants in response to two different evaluation instruments: a test contextualized with a narrative framework (CONTEST) and a test based on metalinguistic iconography without discursive context (METATEST), the latter administered both with and without explicit instruction.

The study was conducted online via the Google Forms platform, where participants completed the tests individually under standardized instructions. Responses were automatically collected and stored in an Excel spreadsheet for subsequent qualitative and quantitative analysis.

Regarding the sample, there were 21 participants in the condition without explicit instruction (CONTEST + METATEST) and 40 participants in the condition with explicit instruction (CONTEST + METATEST, with a metalinguistic guide). The total sample comprised 61 adult participants with university-level education, ranging in age from 18 to 52 years. All participants were native speakers and were distributed across two experimental conditions: 21 participants took part in the condition without explicit instruction (CONTEST + METATEST), and 40 participants in the condition with explicit instruction (CONTEST + METATEST accompanied by a metalinguistic guide). Gender distribution was approximately balanced in both groups, with a slight male predominance in the first group (11 men and 10 women) and a mild female predominance in the second (21 women and 19 men). This distribution was not intended to achieve exact parity, as the study does not follow a stratified sociolinguistic design. Consequently, no further demographic segmentation was carried out.

The methodology used in this study has made it possible to explore how visual support, discursive context, and metalinguistic instruction influence native speakers' choices and understanding of verb forms. Although the data present certain limitations, they provide a valuable basis for gaining deeper insight into how these factors affect participants' linguistic behavior in relation to the study's objectives.

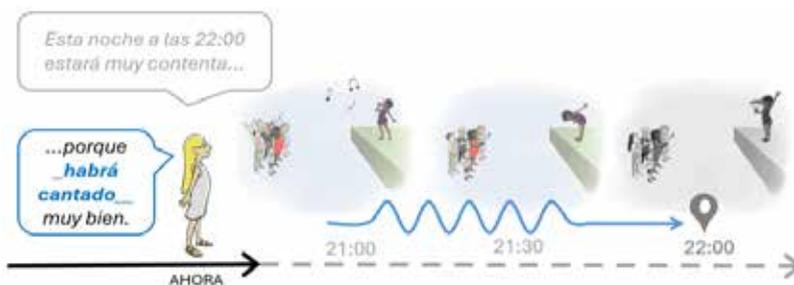
### 3.1. Assessment tests

This study evaluated two main types of visual sequences. The first corresponds to a figurative sequence with a fully narrative structure, referred to as *CONTEST*, in which a series of events takes place at a train station while Laura, the protagonist, talks to her mother on the phone. This type of visual stimulus facilitates a contextualized interpretation of the verb tense by presenting a progression of events that naturally refers to a coherent narrative framework (Figure 1).



**Figure 1.** *Figurative representation of the present perfect tense in the CONTEST test*

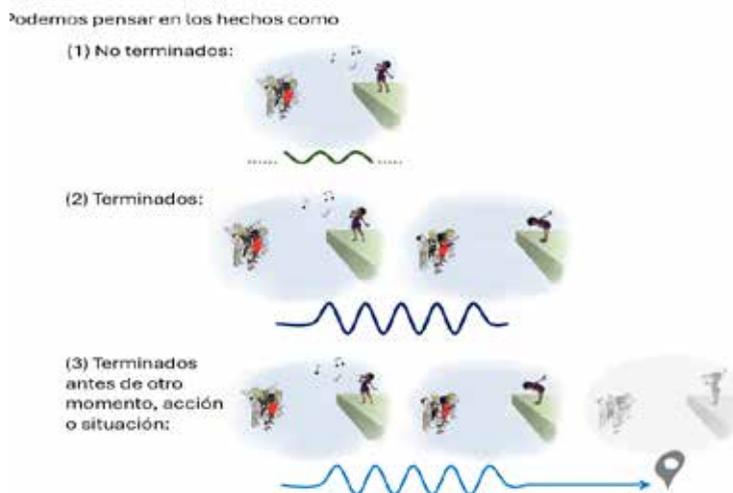
The second sequence consists of schematic images incorporating metalinguistic iconography, such as timelines, reference points, and visual markers of the temporal axis, as shown in Figure 2. These representations were specifically designed to depict abstract temporal concepts in both static and animated formats. The purpose of this test is to illustrate different verb tenses using the verb *cantar* (“to sing”) as the base form.



**Figure 2.** *Figurative representation of the future perfect tense in the METATEST*

Although the native speakers who completed the tests possess an intuitive and functional command of the verbal system, they have generally not received formal instruction in metalinguistic content—neither in the context of foreign language teaching (ELE) nor from cognitively based approaches. In other words, while they have an intuitive and functional grasp of verb usage, they do not always have the metalinguistic tools needed to analyze or justify that usage from a reflective or explicit perspective.

For this reason, an initial experimental condition is proposed in which explicit metalinguistic instruction is provided only in the METATEST. This instruction focuses on the interpretation of the iconography incorporated into the image sequences—both static and animated. The iconography is designed as a visual resource that functions as a form of cognitive scaffolding, aimed at facilitating the processing and conceptualization of temporal relationships. In this context, visual cues replace traditional verbal explanations while still constituting a metalinguistic treatment of the verbal structure, intended to support the explicit understanding of grammatical tenses, as illustrated in Figure 3.



**Figure 3.** Example of explicit metalinguistic instruction in the METATEST

In the second experimental condition, no instruction was provided regarding the visual representation system, with the aim of observing possible differences in participants' performance in terms of the selection and understanding of verb tenses. This comparison will allow for evaluating the effectiveness of visual metalinguistic instruction as a means of facilitating reflective access to the Spanish tense system.

In both conditions, the use of the following verb tenses is assessed, exemplified with the verb *cantar* ("to sing"), selected for their relevance to the linguistic proficiency level under study: present indicative (*canta*), present perfect indicative (*ha cantado*), preterite (*cantó*), imperfect indicative (*cantaba*), pluperfect (*había cantado*), simple future (*cantará*), simple conditional (*cantaría*), future perfect (*habrá cantado*), and conditional perfect (*habría cantado*).

#### 4. STATISTICAL ANALYSIS AND RESULTS

The introduction and adoption of any pedagogical tool—regardless of its apparent innovative nature or the theoretical strengths that support it—requires a rigorous and thorough process of scientific validation. This process is essential not only to ensure the tool's effectiveness in achieving its intended learning objectives but also its reliability—that is, the consistency and stability of its measurements—and its fairness, ensuring that it does not introduce biases that could disadvantage certain groups of students. In this regard, validation goes beyond merely fulfilling a methodological requirement; it becomes an ethical safeguard and a *sine qua non* condition for the responsible adoption of innovations in the field of education. For iconic-figurative tests, which are proposed as a potential improvement over more conventional approaches to teaching verb tenses, this imperative for validation is particularly pressing and fully justifies the statistical analyses applied to the tests examined in this study.

**4.1. Global results and discussion**

To begin the data analysis, descriptive calculations were performed to provide an overview of the participants’ performance. The data analyzed corresponds to the responses of the 61 native speakers distributed into two groups: Group A (N = 21), who completed the tests without explicit instruction, and Group B (N = 40), who received a brief metalinguistic guide prior to testing. Each response was coded dichotomously—1 for correct answers and 0 for incorrect ones—for each item in the two instruments analyzed (CONTEST and METATEST).

The results obtained in this exploratory study offer preliminary support for the pedagogical potential of metalinguistic images in illustrating tense, mood, and aspect distinctions in Spanish. As hypothesized, Group B outperformed Group A in terms of accurate verb form selection across both test formats. The trend became especially clear when the participants received contextual narrative cues, suggesting a synergistic effect between explicit instruction and discursive embedding.

Table 1 below summarizes the item-level performance of both groups across the two instruments. It highlights the influence of instructional condition and contextualization on the correct interpretation of grammatical content.

**Table 1.** *Distribution of items by group and test based on the number of correct answers*

| ITEMS                                     | CONTEST<br>GROUP A | CONTEST<br>GROUP B | METATEST<br>GROUP A | METATEST<br>GROUP B |
|---|--------------------|--------------------|---------------------|---------------------|
| Ítem 1<br>Presente                        | 100%               | 95%                | 95.2%               | 100%                |
| Ítem 2<br>Pretérito imperfecto            | 66.7%              | 55%                | 33.3%               | 55%                 |
| Ítem 3<br>Pretérito indefinido            | 100%               | 92.5%              | 66.7%               | 62.5%               |
| Ítem 4<br>Pretérito pluscuamperfecto      | 92.5%              | 97.5%              | 66.7%               | 72.5%               |
| Ítem 5<br>Pretérito perfecto<br>compuesto | 57.1%              | 65%                | 95.2%               | 60%                 |
| Ítem 6<br>Condicional compuesto           | 100%               | 100%               | 61.9%               | 65%                 |
| Ítem 7<br>Condicional simple              | 95.2%              | 97.5%              | 33.3%               | 22.5%               |
| Ítem 8<br>Futuro simple                   | 95.2%              | 100%               | 81%                 | 80%                 |
| Ítem 9<br>Futuro compuesto                | X <sup>1</sup>     | X                  | 71.4%               | 72.5%               |

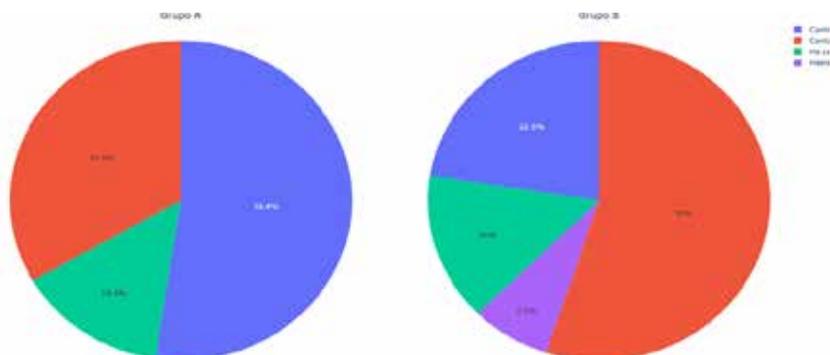
<sup>1</sup> This verb tense was not included in the narrative that contextualizes the events presented in the *CONTEST* test.

When analyzing the results obtained in the *CONTEST* test, it is observed that there is no significant difference in performance between the two groups. Group A achieved a score of 88.34%, while Group B obtained 87.81%. It is worth noting that both groups surpassed 55% correct answers across all evaluated items.

In the case of Group B, the group with the larger number of participants, the average number of correct responses was 0.8125, which equals 81.25% correct answers on average. The standard deviation was 0.081981, indicating low variability in performance, with relatively homogeneous scores among participants.

In the case of the *METATEST*, Group B, which received explicit instruction, showed an improvement in the number of correct answers on items 1, 2, 4, 6, and 9 compared to Group A, maintaining the same number of correct answers on item 8, with a minimal difference of 1%. This difference in the number of correct answers could be attributed to the sample size, as Group A had half the number of participants as Group B. However, in general, and except for items 3, 5, and 7 highlighted in bold, an increase in performance is observed when explicit instruction is present.

It is important to note that several items related to the verb *cantar* in the *METATEST* show a significant improvement in performance when comparing the responses from Group B (with explicit instruction) to those from Group A (without instruction). This pattern is especially evident in the case of the imperfect preterit (*cantaba*), whose interpretation improves significantly with explicit instruction, increasing from 33.3% correct answers in Group A to 55% in Group B, suggesting a positive effect of metalinguistic support on the understanding of this verb tense (Figure 4).

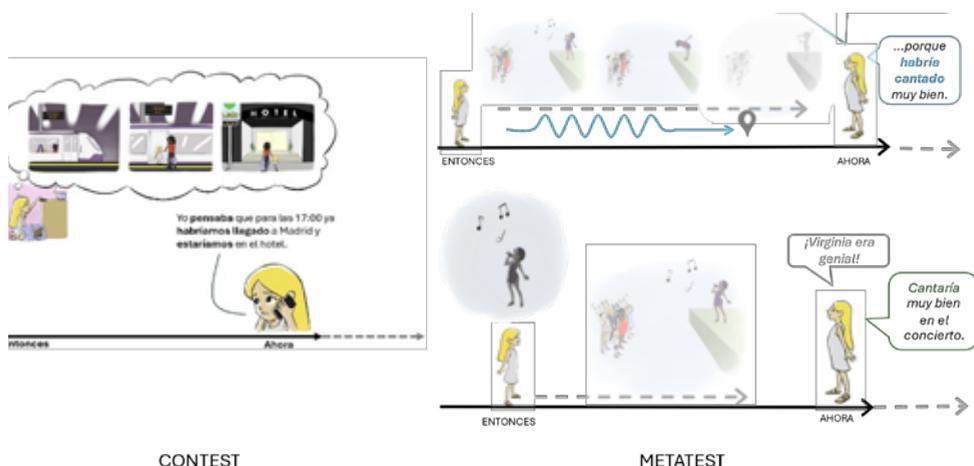


**Figure 4.** Comparison of correct answers for the “pretérito imperfecto” between Group A and Group B in the *METATEST*

If we compare the results of the *CONTEST* and *METATEST* assessments shown in the table, we observe that in the case of item 6—the past conditional (e.g., *habría cantado*), which is the most complex tense from a deictic perspective—both groups, A and B, show a similar percentage of correct responses, around 65%, compared to the 100% accuracy recorded in the *CONTEST*. This suggests that when verb forms are presented in isolation or in decontextualized sentences, as is the case in the *METATEST*, the task becomes significantly

more complex. The absence of a discursive framework increases the level of abstraction and hinders intuitive access to meaning by removing the contextual support that typically guides verbal comprehension and usage in real communicative situations.

Even so, the data require a more nuanced interpretation when we also consider the results for the simple conditional (*cantaría*), whose deixis is somewhat more accessible than that of the past conditional (*habría cantado*), although still complex. This is due to the temporal shift from the present to the past that it entails, particularly when used to express the so-called “future in the past,” which is the meaning targeted in the tests. The contrasting representations of both tenses are shown in Figure 5.

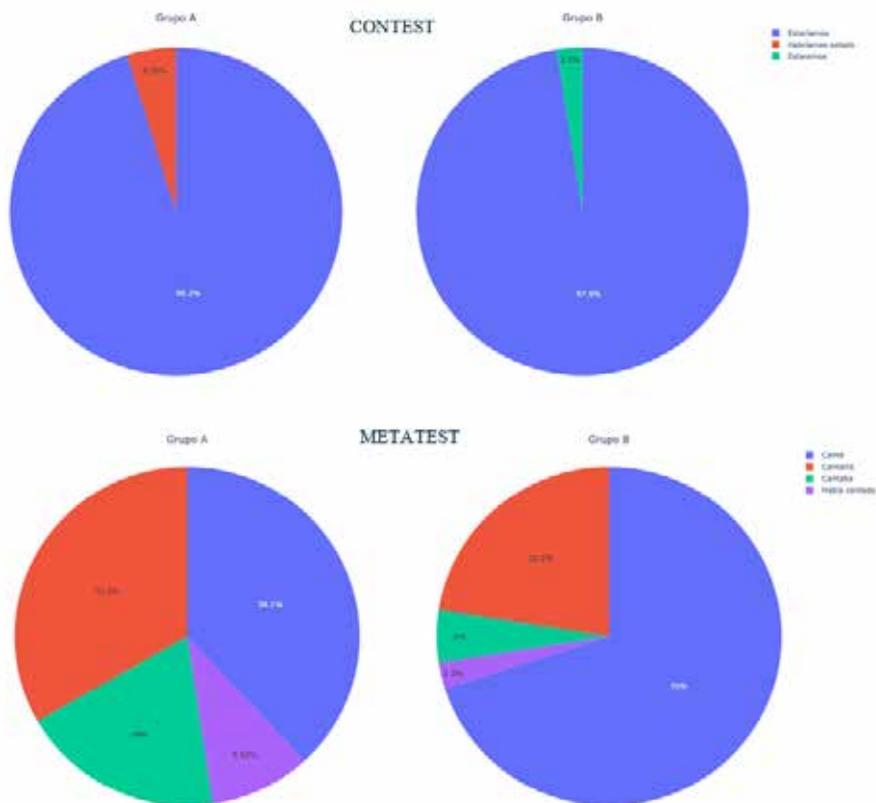


**Figure 5.** Representation of the past conditional and simple conditional in CONTEST and METATEST

In Figure 6, regarding the simple conditional (shown in blue), we observe high variability in responses and a low level of accuracy in both groups in the METATEST, with 33.3% correct answers in Group A and 22.5% in Group B—significantly lower than the results obtained for the past conditional. This difficulty may be attributed to the absence of a discursive context that would disambiguate its interpretation. In the CONTEST, where a narrative framework is provided, the accuracy rates increase substantially to 95.2% and 97.5%, respectively.

The reason why the past conditional—despite its greater conceptual complexity (referring to a completed hypothetical or predicted event prior to a moment posterior to “then”)—yields much higher scores than the simple conditional in the METATEST may lie in the presence of a disambiguating discursive context also included in the corresponding representation for the past conditional in the METATEST (see Figure 6). In contrast, the limited discursive context offered for the simple conditional fails to exclude alternative interpretations, such as the preterit, which was one of the preferred options selected by the participants. In this

case, respondents prioritized the speaker's present perspective (i.e., viewing events from “now”) over the “future in the past” interpretation (i.e., how events would unfold from a “then” perspective).



**Figure 6.** Comparison of correct answers for the simple conditional between Group A and Group B in both CONTEST and METATEST

From the data presented in Table 1, two relevant observations can be drawn. First, the number of correct answers obtained from both groups in the *CONTEST* test exceeds that recorded in the *METATEST*. This result highlights that the presence of a discourse context facilitates the correct selection of the verb form. Second, in the latter test, Group B, which received explicit instruction, increased the number of correct answers on five items. This result suggests that explicit instruction aimed at interpreting iconography with metalinguistic content helps native speakers select the correct verb form more appropriately, except in cases of greater grammatical complexity, such as using the future-in-the-past, where the discourse context becomes more relevant.

These findings align closely with the theoretical assumptions underpinning this study, particularly those of cognitive grammar and dual coding theory. The improved performance observed in Group B supports the idea that combining verbal explanations with visual metaphors enhances the interpretation of abstract grammatical categories, such as tense and aspect. The synergy between discursive context and explicit instruction confirms that linguistic categories are more readily understood when grounded in usage-based, context-sensitive frameworks—as posited by Langacker (1987, 2019). Furthermore, the differential performance between narrative and decontextualized conditions reinforces the notion that grammatical meaning is not only symbolic but also deeply dependent on the speaker's conceptualization of events, an idea central to cognitive linguistics.

Finally, the results provide preliminary empirical support for the content validation approach proposed in this study. They indicate that metalinguistic images—especially when accompanied by contextual and instructional scaffolding—can serve as effective tools for illustrating complex verbal phenomena. Since this is a pilot phase, we intend to refine the materials based on these initial findings and subsequently assess their impact in controlled quasi-experimental settings with L2 learners, allowing systematically and rigorously evaluate how these materials influence grammatical acquisition.

#### 4.2. Statistical analysis

The statistical analysis conducted aims to determine whether there are significant differences in the selection of verb tenses between the first sequence (CONTEST) and the second (METATEST). Given the small sample size of Group A (N=21), the analysis focused only on the scores obtained by the participants in Group B (N=40) in both tasks, in order to assess whether the observed difference in their performance was statistically significant.

The statistical analysis conducted aims to determine whether there are significant differences in the selection of verb tenses between the first sequence (CONTEST) and the second (METATEST). Given the small sample size of Group A (N=21), inferential statistical analysis focused solely on the scores obtained by participants in Group B (N=40) across both tasks to assess whether observed performance differences were statistically significant. Nevertheless, although no inferential statistics were applied to Group A due to its limited size, a descriptive analysis of their responses was performed. This exploratory examination provides valuable preliminary insights into participants' verb tense selection patterns in the absence of explicit instruction. As this study corresponds to a pilot phase, such descriptive data help identify trends and inform the refinement of the pedagogical materials under investigation. Moreover, comparing both groups—one with and one without explicit guidance—supports the broader objective of evaluating the relative effectiveness of instructional approaches, even when statistical comparisons are constrained by sample size limitations.

First, the normality of the data from Group B corresponding to CONTEST and METATEST was evaluated using the Shapiro-Wilk test, designed to detect deviations from a normal distribution in small samples. The significant values obtained were 0.011 for CONTEST and 0.007 for METATEST. Since both were below the threshold of 0.05, the null hypothesis of normality was rejected, indicating that the data did not follow a normal distribution.

Once it was confirmed that the data did not follow a normal distribution, the Wilcoxon signed-rank test was used, a non-parametric test that serves as an alternative when a normal

distribution cannot be assumed. In this case, the scores obtained by each participant from Group B in the train test and the singing test were compared, thus constituting pairs of dependent data. The analysis was conducted using the Python Scipy library (<https://docs.scipy.org/doc/>), a standard tool for statistical analysis.

The statistical result of the Wilcoxon test indicated a W statistic of 150.0 and a p-value  $< 10^{-3}$ , suggesting that the difference observed in the mean number of correct responses between the CONTEST and METATEST was statistically significant. This implies that the probability of this difference occurring by chance is extremely low ( $p < 0.05$ ), supporting the existence of a real difference between the conditions, and with high confidence, the performance of Group B was better in CONTEST than in METATEST.

This is particularly relevant as it indicates that tasks presenting verb forms within a discourse context, as in the CONTEST test, make recognizing and selecting the correct options considerably easier for participants (p-value  $< 10^{-3}$ ). The discourse framework acts as a scaffolding that facilitates interpretation, reducing the cognitive load associated with verbal selection. Thus, instead of relying solely on the retrieval and application of abstract grammatical rules—a process that can be demanding, especially for non-native speakers—participants can use the coherence of the story and discourse as their main guide.

In the METATEST, however, the narrative and discourse environments tend to be more limited; the contexts accompanying each verb tense are, in many cases, partial or underdeveloped, and no cohesive narrative framework is constructed like the one in the CONTEST sequence. Nevertheless, in those items that include a more complete discourse context, the proportion of correct responses tends to be higher.

This process, being closer to the natural use of language, requires less deliberate cognitive effort, thus explaining why the task feels “easier” compared to performing the METATEST. While this facilitation benefits all participants, it operates in subtly different ways: native speakers can rely more on their linguistic intuition, now guided by the context, while non-native speakers can use the context to fill in gaps in their explicit knowledge of grammatical rules.

## 5. CONCLUSIONS

Although a larger number of informants in the sample is always desirable, given the context of data collection and the inherent limitations of the content, the data obtained so far, while tentative, clearly allow us to identify certain relevant clues.

The comparative analysis of the data from the CONTEST and METATEST tests clearly demonstrates the fundamental role played by the narrative and discursive context in facilitating the processing and proper selection of Spanish verb forms. The presence of a coherent contextual framework, as seen in CONTEST, acts as a cognitive support that reduces processing load and allows participants to rely on discourse-based inferences to determine appropriate verb usage. In contrast, the absence of such a context, typical of METATEST, significantly increases the difficulty by requiring greater metalinguistic effort and relying more on explicit grammatical knowledge, revealing different limitations in both native and non-native speakers.

It has also been observed that participants' performance improves significantly when provided with an initial clue, a guide that helps them focus their attention on the elements

making up the representations. These elements, it should be noted, exhibit some degree of arbitrariness, making this initial aid especially helpful for comprehension.

Key findings highlight the following: a) the narrative context reduces cognitive load and facilitates inference; b) decontextualized tasks reveal the gap between implicit and explicit linguistic knowledge; c) even a partial discursive context can be sufficient to disambiguate complex forms, sometimes overcoming the inherent difficulty of the form, as exemplified by the verb form “habría cantado”; d) variability in responses persists even in rich contexts, reflecting linguistic flexibility and individual interpretation; e) explicit instruction can be beneficial, but its effectiveness varies according to the verb form and the quality of the explanation.

These results have direct and significant implications for pedagogical practice and assessment design. It is strongly recommended to prioritize the use of authentic narrative and discursive contexts in teaching and assessing grammar. Additionally, an integrated pedagogical approach that combines contextualized practice with focused explicit instruction, supported by multimodal resources and clear training on tasks, is suggested.

In the field of assessment, the construct validity of the instruments should be carefully considered, recognizing that different test formats measure different facets of linguistic competence. Research of this type is essential to inform the development of evidence-based practices that optimize the learning and assessment of the complex verbal grammar of Spanish.

The present findings carry direct and significant implications for pedagogical practice and the design of assessment tools. It is therefore strongly recommended to prioritize the incorporation of authentic narrative and discursive contexts in the teaching and evaluation of grammar. Furthermore, an integrated pedagogical approach that combines contextualized practice with targeted explicit instruction—supported by multimodal resources and comprehensive task training—is advocated. In the field of assessment, the construct validity of instruments should be carefully considered, recognizing that different test formats measure distinct facets of linguistic competence. Research of this type is essential to inform the development of evidence-based practices that optimize the learning and assessment of the complex verbal grammar of Spanish.

These encouraging results lay a solid foundation for future research, although certain aspects warrant further development. The relatively small sample size, particularly in Group A, suggests that future studies with larger and more diverse participants will be essential to strengthen the generalizability and statistical robustness of the findings. Additionally, as this is a pilot study, it provides valuable insights to refine both the instructional materials and the procedures, which can be further enhanced and validated through more extensive research, incorporating complementary measures such as qualitative feedback.

## ACKNOWLEDGEMENTS

We would like to express our special gratitude to Lucía Castañeda González, Alejandro Castañeda Castro (Department of General Linguistics and Theory of Literature, UGR), and Adolfo Sánchez Cuadrado (Department of General Linguistics and Theory of Literature, UGR) for their invaluable collaboration and contributions to this work.

## FUNDING INFORMATION

This study is part of the work carried out within the framework of the project “Gramática figurativo-constructiva del tiempo y aspectos verbales en inglés y español. Aplicación didáctica cognitivo-contrastiva para el desarrollo de la competencia plurilingüe. (IMAGINE: Imágenes Gramaticales del Inglés y el Español)” REF.: PID2021-128771OB-I00, funded by *Ministerio de Ciencia e Innovación. FEDER Una forma de hacer Europa*.

## 6. REFERENCES

- Alarcos Llorach, E. (1994). *Gramática de la lengua española*. Espasa Calpe.
- Anastasi, A., & Urbina, S. (1997). *Psychological testing*. Prentice Hall.
- Andersen, R. W., & Shirai, Y. (1994). Discourse motivations for some cognitive acquisition principles. *Studies in second language acquisition*, 16(2), 133-156.
- Ayoun, D., & Salaberry, M. R. (2005). *Acquisition of Spanish aspectual distinctions by L2 learners*.
- Bardovi-Harlig, K. (2000). *Tense and aspect in second language acquisition: Form, meaning, and use*. Oxford: Blackwell.
- Bardovi-Harlig, K., & Comajoan-Colomé, L. (2024). A critical review of instructional-effect studies of past tense-aspect: language development and pedagogical implications. *Journal of Spanish Language Teaching*, 1–16. <https://doi.org/10.1080/23247797.2024.2434334>
- Boers, F., Lindstromberg, S., Littlemore, J., Stengers, H., & Eyckmans, J. (2008). Variables in the mnemonic effectiveness of pictorial elucidation. *Applications of Cognitive Linguistics*, 6, 189.
- Castañeda Castro, A. (2019). Lingüística cognitiva. En J. Muñoz Bassols, E. Gironzetti y M. Lacorte (eds.), *The Routledge Handbook of Spanish Language Teaching: metodología, recursos y contextos para la enseñanza del español L2*. (p.p.)261-278. Londres/Nueva York: Routledge.
- Castañeda Castro, A. (2023). Aplicación de los modelos temporalista y epistémico en la enseñanza de tiempo y aspecto en la clase de ELE. El potencial de la gramática cognitiva como opción integradora. En M. Martínez Atienza y A. Zamorano Aguilar (eds.). *Teoría de la lengua y enseñanza-aprendizaje de ELE* (P.P. 15-45). Tirant Lo Blanch.
- De Knop, S., & Dirven, R. (2008). Motion and location events in German, French and English: A typological, contrastive and pedagogical approach. *Applications of cognitive linguistics*, 9, 295.
- Geeraerts, D. (ed.) (2006). *Cognitive Linguistics. Basic Readings*. Berlín/Nueva York: Mouton de Gruyter.
- Ibarretxe-Antuñano, Iraide, Teresa Cadierno y Alejandro Castañeda (eds.) (2019). *Lingüística cognitiva y español L2/LE*. Londres: Routledge.
- Kerlinger, F. N. (1986). *Foundations of behavioral research*. Holt, Rinehart and Winston.
- Langacker, R. (1987): *Theoretical Prerequisites. Vol. 1 de Foundations of Cognitive Grammar*. Stanford Univ. Press.
- Littlemore J. (2009). *Applying cognitive linguistics to second language learning and teaching*. Springer.
- Messick, S. (1989). Validity. En R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). American Council on Education/Macmillan.

- Niemeier, S. (2017). *Task-based grammar teaching of English*. Narr Francke Attempto.
- Paivio, A. (1986). *Mental representations: A dual-coding approach*. New York: Oxford University Press.
- Salaberry, M. R. (2000). The acquisition of English past tense in an instructional setting. *System*, 28(1), 135-152.
- Salaberry, M.R. (2018). Advanced Conceptualizations of Tense and Aspect in L2 Acquisition. En *The Handbook of Advanced Proficiency in Second Language Acquisition*. <https://doi.org/10.1002/9781119261650.ch19>
- Tyler, A., y Evans, V. (2012). *Cognitive linguistics and second language learning: Theoretical basics and experimental evidence*. Routledge.
- Weir, C. J. (2005). *Language testing and validation: An evidence-based approach*. Palgrave Macmillan.
- Weir C. J. and O'Sullivan, B. (2011). Test development and validation. En O'Sullivan, B. (ed.) *Language Testing: Theories and Practices*. Palgrave Advances in Language and Linguistics. Palgrave MacMillan.
- Zhang, Q., & Lu, X. (2015). *The effect of context on second language vocabulary learning: A meta-analysis*. Review of Educational Research, 85(4), 643-679.