

Exploring the link between metalinguistic awareness and reading skills in migrant learners of Spanish

M.^a PILAR NÚÑEZ DELGADO

MARÍA SANTAMARINA SANCHO (CORRESPONDING AUTHOR)

University of Granada

Received: 2025-04-26 / Accepted: 2025-10-07

DOI: <https://doi.org/10.30827/portalin.viXV.33634>

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

ABSTRACT: This paper presents the results of a research study aimed at determining the levels of metalinguistic awareness among students of different nationalities in compulsory secondary education, specifically in the first year of ESO. It then explores the correlation between these levels and students' reading competence. The instruments used included the THAM-2 test of metalinguistic skills and a reading comprehension assessment, both administered to two groups from the same school (a total of 64 students), as well as a focus group composed of three students—each representing one of the three levels of metalinguistic awareness established by the test. These students, originally from Spain, Ukraine, and Morocco, also enabled a comparative analysis based on differences in L1. The data were analyzed using content analysis and the Constant Comparative Method (CCM) applied to participants' discourse. These findings, combined with the statistical results from the THAM-2, reveal a positive relationship between reading comprehension and students' metalinguistic skills, as well as the significant influence of multilingualism.

Keywords: metalinguistic awareness, reading comprehension, multilingualism, THAM-2, focus group.

Explorando la relación entre la consciencia metalingüística y las habilidades lectoras en estudiantes migrantes de español

RESUMEN: El estudio que se presenta examina la relación entre la consciencia metalingüística y la competencia lectora en estudiantes migrantes de primer curso de Educación Secundaria Obligatoria (ESO). En la investigación han participado 64 alumnos de diferentes nacionalidades, quienes realizaron el test para medir las habilidades metalingüísticas, THAM-2, junto con una prueba de comprensión lectora. Además, se ha llevado a cabo un grupo focal con tres estudiantes procedentes de España, Ucrania y Marruecos, pues estos representaban cada uno de los niveles de consciencia metalingüística identificados por el test, lo que ha permitido analizar diferencias según la lengua materna. Se ha utilizado un análisis de contenido y el Método de Comparación Constante (CCM) para procesar los datos cualitativos, complementados con análisis estadísticos del THAM-2. Los resultados muestran una correlación positiva entre consciencia metalingüística y comprensión lectora, y destacan la influencia significativa del multilingüismo en ambas competencias.

Palabras clave: consciencia metalingüística; comprensión lectora; multilingüismo; THAM-2; grupo focal.

1. INTRODUCTION

Language is a semantic code with specific characteristics that distinguish it from other communication systems. One of its most unique properties is its capacity for self-reflection—this is the metalinguistic function; a feature not found in other communicative codes. The role of this function in the teaching and learning process is undeniable, particularly in the case of reading comprehension. Numerous studies have demonstrated that the ability to consciously manipulate language and its structural elements facilitates the process of learning to read (Defior, 1990; Pinto, 1999; Milian & Camps, 2000; Fernández de Haro, Santamarina & Romero, 2009; Calero, 2011; González Manjón, 2015; Santamarina & Núñez, 2021; Floquet & Melogno, 2024).

2. LITERATURE REVIEW

2.1. Metalinguistic awareness: A conceptual approach

Metacognition can be defined, following Pratt and Grieve (1984), as the ability to reflect upon and regulate the outcomes of one's own thinking, arising from an increasing awareness of the nature of cognitive functions. This definition builds upon Flavell's (1976) foundational concept of metacognitive knowledge—that is, an individual's awareness and understanding of their own cognitive processes, including the monitoring, regulation, and organization of those processes in relation to the cognitive tasks, data, or information being engaged with.

Similarly, Weinstein and Mayer (1986) define metacognition as an individual's knowledge of their own mental processes and the degree of control they are able to exercise over them. In this sense, metacognition refers to the capacity to manage, that is, to organize, monitor, and adjust one's cognitive activity in response to the outcomes generated through its application.

If we apply the prefix “meta” to language, we understand metalanguage as the use of language to discuss language itself. In this sense, language—like other cognitive skills—operates within regulation and control mechanisms. These mechanisms are not limited to situations where the language is being learned as an object of study, but also in contexts where it serves functionally as a tool for communication (Oliva & Santamarina, 2015). This property of language entails the ability to reflect on linguistic forms by distinguishing elements that are internal to language from those that belong to broader, non-linguistic domains (Loureda, 2009). Metalanguage, in short, is one of the components of metacognition, alongside metamemory, metalearning, metaattention, and social metacognition (Tunmer, Pratt & Herriman, 1984). From a cognitive perspective, metalinguistic awareness involves higher-order executive functions such as cognitive inhibition, mental flexibility, and working memory. These processes are essential for monitoring and manipulating linguistic structures while suppressing automatic or intuitive responses. According to Bialystok (2004), this ability to consciously analyze and restructure language depends on the coordination of attentional control mechanisms, which allow individuals to shift between form and meaning. Additionally, the development of metalinguistic skills requires the integration of declarative linguistic knowledge with reflective reasoning processes, thereby engaging both analytic and executive domains of cognition.

In recent years, there has been a proliferation of research on metalanguage (Quispe-Morales, 2022; Carbajal-Carrera, 2025), approached from both qualitative and quantitative per-

spectives. As a result, the term “metalinguistic awareness” encompasses numerous definitions, stemming from the wealth of theoretical and empirical contributions on the subject (Pinto, Titone & González, 2000). Tunmer, Pratt, and Herriman (1984) argue that metalinguistic awareness refers to the ability to reflect upon and manipulate the elements that structure language (phonemes, words, propositional structures) as an object of knowledge in its own right. This distinguishes it both from the mere use of language to produce and understand utterances and from the knowledge of terms used to describe language, such as phoneme, word, sentence, etc. Mattingly (1972), for his part, highlighted that metalinguistic awareness pertains to an individual’s awareness of the sounds in their own language. For Quispe-Morales (2022), metalinguistic awareness or reflection is understood as the ability to think critically about language. He points out that language use, when not accompanied by reflection, is unlikely to be an effective mechanism for learning. Metalinguistic ability evolves gradually, meaning that the level of awareness exhibited by an individual is dependent on their mastery of the linguistic elements in question. In short, metalinguistic awareness is the ability to think about the linguistic nature of messages, manipulate their components, and justify their use (Pinto & Eleuch, 2015).

More recent research confirms the influence of metalinguistic reflection on text comprehension. Riffo et al. (2018), for example, conclude that the development of phonological and lexical skills is a crucial element in understanding written texts. Similarly, Bizama et al. (2017) present findings that demonstrate the importance of implementing both syntactic and metalinguistic skills to improve text comprehension. Rueda-Sánchez and López-Bastida (2016), following an extensive study, assert that morphological awareness enhances the effectiveness of text comprehension and writing in students with specific educational needs. Coloma et al. (2015) show that the development of decoding skills and vocabulary are key factors in the comprehension of written texts. Likewise, Quispe-Morales (2022) concludes that the development of metalinguistic awareness has a significant impact on literal, inferential, and critical reading comprehension in Spanish as a second language, particularly in Primary Education students.

In response to this body of work, our proposal aims to contribute to the field by emphasizing the benefits of fostering metalinguistic awareness in the classroom, particularly to improve reading comprehension in Spanish as a second language, among other outcomes.

2.2. Metalinguistic awareness and plurilingualism

Plurilingualism can be defined, according to Cenoz and Gorter (2015), as the use of two or more languages in an educational context, provided that the school promotes multilingualism and literacy. Authors such as Hammer et al. (2023) highlight that, due to the significant impact of immigration on classrooms over the past 25 years, multilingual pupils in countries like Sweden are now on par with refugee and immigrant children. In Spain, the presence of foreign students in the education system has continued to grow since the beginning of the 21st century (Fernández, 2024; INE, 2022).

One of the students who participated in the focus group for this study, the Moroccan student, was a native speaker of Arabic, with a strong command of both Arabic and a second language, Spanish. Additionally, he spoke a third language, French, albeit at a lower level. This student could be classified under what Fernández-Portero and Morera-Bañas (2025)

describe as a form of plurilingualism shaped by migratory circumstances, where multiple language experiences coexist in a fluid and overlapping manner. However, language policies and educational interventions often address these linguistic realities separately, according to the prevailing social context. As we will see in the presentation of the results, this multilingual background led to more accurate and elaborated responses from the student.

The literature reveals a wide range of neuroscientific studies that demonstrate the cognitive benefits for bilingual or multilingual individuals. For instance, Bialystok (2024) has extensively documented how managing multiple linguistic systems strengthens executive control in the brain. Rivera Mandarache et al. (2019) conclude that bilingualism positively impacts reading comprehension, as it promotes mental flexibility. Bilingual individuals, who are familiar with two cultures and their unique features, can navigate both effectively. In fact, metalinguistic reflection aims to ensure that students develop a coordinated bilingualism, equipping them with the ability to use both languages proficiently and understand texts written in different linguistic codes. In short, bilingualism is linked to enhanced executive control, enabling individuals to manage language interference more efficiently and select the appropriate language based on the context (Marsh et al., 2020).

2.3. Focus groups in educational research

Research in education plays a critical role within the social sciences, as it helps address the challenges and issues arising from educational practices. At the same time, it provides new instruments and tools for teachers and all stakeholders involved, ensuring an optimal teaching and learning process for students.

The goal of this research is to establish the levels of metalinguistic awareness among first-year ESO students of different nationalities, then correlate this awareness with their reading literacy, and analyze the possible influence—whether positive or negative—of multilingualism on the former. This influence has been identified as both real and beneficial in previous research, including studies by Cenoz (2003), Lasagabaster (2010, 2025), Pérez López (2022), and Alonso Rey (2023).

Once the metalinguistic skills test and the reading comprehension test were completed, a focus group discussion was conducted with three students, each representing one of the levels established by the test. This approach was chosen because our objective is not to obtain individual narratives—for which a case study would be more appropriate—but rather to contrast and analyze the responses of students with different levels of metalinguistic awareness. Furthermore, the focus group allows for triangulation of results, as Barbour (2013) confirms that the nature of focus group discussions means that stories are unlikely to unfold sequentially, as they might in a one-to-one interview. Therefore, the picture presented will be fragmented and attempts to analyze the data will be frustrated.

3. METHOD

3.1. Participants

As previously described, the students who participated in the study are of different nationalities, all of them attending compulsory secondary school in Spain, specifically in

the 1st year of ESO. Initially, the test of metalinguistic skills (THAM-2, validated for the Spanish language in 2015 by Santamarina and Pinto) was administered to 64 students in two groups of 1st year ESO students in the same school, a secondary school in a neighbourhood near the centre of a city in Andalusia (Spain). The area in which the school is located is characterised by being a working-class neighbourhood, built in the middle of the 20th century. It has a population of around 44,000 inhabitants, making it one of the most populous neighbourhoods in the city. Of this population, 11% are foreign residents with a wide national diversity. The school's website reports that it has experienced a notable increase in immigrant pupils, who account for 20% of the total.

The focus group was conducted with three students, each representing one of the three levels of metalinguistic awareness established by the test. These students are from different countries (Ukraine, Spain, and Morocco), which is also characteristic of the student population at the school and in the surrounding neighborhood.

3.2. Instrument: THAM-2 – An instrument to assess metalinguistic awareness

The THAM-2 is an instrument designed to measure metalinguistic skills in learners aged 9-13 years. The preamble of the THAM-2 provides a solid theoretical foundation for this age range. The current version of the THAM-2 (Santamarina and Pinto, 2015) is based on the original THAM-2 developed by Hakes in 1980, which consisted of four tests: Comprehension, Synonymy, Acceptability, and Phonemic Segmentation. Hakes conducted his research with children aged four to eight years. According to the authors of the current version, the starting point was the observation of the justifications provided by participants in Hakes' study. This led to the proposal to examine the reasoning behind the answers given by participants, drawing on Piaget's discoveries from the 1920s regarding children's reasoning. As Santamarina and Pinto (2015) explain, THAM-2 aims to promote the meta-distinction between what requires a type of elaboration that is primarily epilinguistic and what requires another type that may, on the other hand, be exclusively metalinguistic.

These innovations were extensively studied and developed within the European project MATEL (Metalinguistic Awareness Test in European Languages), which aimed to deepen the understanding of metalinguistic awareness and its role as a beneficial factor in bilingual development during students' educational growth (Bialystok, 2004, 2005; Pinto, 2002, 2004). The project enabled the creation and validation of various metalinguistic skills tests in several languages (Italian, French, German, and Spanish), making it possible to assess a range of skills considered fundamental for broad educational learning, particularly those related to language proficiency. One of these tests is the THAM-2, which consists of six components: Comprehension, Synonymy, Acceptability, Ambiguity, Grammatical Function, and Phonemic Segmentation, and includes a total of 96 items. A brief overview of the general features of the structure of each of these tests is provided below:

1. Comprehension

This subtest consists of six pairs of sentences. In each pair, the first sentence (A) presents a specific syntactic structure, while the second sentence (B) restates this structure with certain variations. For example, sentence 1.A might be a simple declarative in the active voice,

and sentence 1.B a simple declarative in the passive voice. According to the test manual, the participant is first asked to interpret the semantic-grammatical relationship expressed in each sentence of the pair (referred to as question L and its corresponding answer L). Subsequently, the participant must explain the reasoning or criteria used to arrive at that interpretation (question ML), thus eliciting their level of metalinguistic awareness.

2. *Synonymy*

The Synonymy subtest includes five pairs of sentences that differ syntactically. Four of these pairs are synonymous, meaning they convey the same message despite their structural differences. One pair, however, appears to differ only in structure, yet actually differs in meaning as well. This item functions as a control to assess the subject's ability to detect subtle semantic distinctions. In all cases, the L and ML components are symmetrical: participants first respond to whether the two sentences are equivalent in meaning (L) and then justify their response (ML).

3. *Acceptability*

This subtest includes six pairs of items whose sentences present some kind of anomaly. In the first five pairs, the anomalies are primarily semantic in nature, resulting from violations of lexematic compatibility rules. These violations are based on contrasts such as animate/inanimate, animal/human, or transitive/intransitive action, among others. The participant must identify whether the sentence is acceptable (L) and justify their response by identifying and explaining the type of anomaly detected (ML).

4. *Ambiguity*

This subtest is divided into two sections. The first present sentence that contain semantic ambiguities, while the second focuses on structural ambiguities. Semantic ambiguity arises from the polysemy of a single term whose different meanings can change the overall interpretation of the sentence, even though the grammatical structure remains the same. Participants are asked to identify how many and which meanings they can attribute to the ambiguous term (question P.L), and subsequently, to explain the meaning of the sentence in relation to each interpretation offered (ML). This task assesses the participant's ability to reorganize sentence meaning based on shifts in the meaning of a single lexeme—an ability considered metalinguistic in nature.

5. *Grammatical Function*

The Grammatical Function subtest is composed of six pairs of items. The first three aim to assess the understanding of key grammatical functions such as subject, object, and predicate. The L-level questions (Literal) guide the participant toward identifying the structure of actions and the roles distributed around them. The corresponding ML-level questions (Metalinguistic) explore the reasoning behind these identifications, prompting the participant to justify their answers through an explicit analysis of grammatical relations.

6. *Phonemic Segmentation*

The Phonemic Segmentation subtest differs from the previous ones in that it operates at a more fundamental level of linguistic analysis. Here, the sentence is no longer the unit of focus, and with its disappearance, the syntagmatic relations inherent to sentence structure are absent. Instead, the test focuses on the participant's ability to identify and manipulate phonemes, syllables, and free morphemes, engaging primarily with paradigmatic relations. This shift results in a redefinition of the L and ML dimensions, particularly regarding the cognitive demands placed on the participant. The test is divided into four parts, although the L/ML distinction is relevant only in two of them.

The coding and scoring procedure for L and ML responses are carried out independently, resulting in distinct scores and statistical values. The authors explain:

L responses are coded using a dichotomous true/false (or similar) format, corresponding to a score of 1 or 0. ML responses, on the other hand, are evaluated individually using three levels, based on psycholinguistic criteria [...] a qualitative tripartition which is quantitatively translated into a three-point scale: 0, 1, 2. Thus, each item can receive a score of 0, 1, or 2, and the final test score is calculated cumulatively by summing the individual scores for each item, as is also the case for L responses.

In the present study, we adhered to these scoring guidelines, which align with the three levels established in the test (ML0, ML1, and ML2). However, for the purposes of the content analysis and to better meet the study's objectives, we applied a different categorisation method to the text units from the focus group. Specifically, we categorised whether the linguistic (L) and metalinguistic (ML) answers were both correct for the same question. The THAM-2 manual, available electronically, provides details not only on the individual administration protocol, but also on the procedures to be followed for its implementation and the psychometric characteristics of the test, among other aspects.

3.3. **Data collection procedure**

The first analysis conducted was quantitative. In this phase, we obtained means, descriptive statistics, and the correlation between the two tests administered: the THAM-2 and the reading comprehension test. To generate these results and the corresponding graphs, we used SPSS software version 21. To strengthen the robustness of the qualitative analysis, this study incorporated source triangulation by comparing the results obtained from the THAM-2 tests and reading comprehension with the focus group data and the researcher's field notes. According to Navarro (2021), triangulation is an essential strategy in qualitative research, as it allows for the cross-validation of multiple methods or data sources to verify the validity of results and minimize inherent biases (proquest.com, researchgate.net). Specifically, data source triangulation—focusing on different types of data—and methodological triangulation were employed, which contributed to increasing the study's credibility and internal coherence, as proposed by Patton (1999) and Denzin (1978).

The second analysis was qualitative, based on content analysis through the categorisation of meaning. For this, we selected three participants from the study (two boys and one girl),

each classified into one of the three possible levels of response, who participated in the focus group (ML0, ML1, and ML2). This second phase was conducted primarily to examine the reasoning provided by the ML2 student, and—anticipating a later conclusion—to explore the positive influence of multilingualism on metalinguistic awareness.

The nature of the measurement used in this part of the research is qualitative. Grounded Theory—specifically, the *Constant Comparative Method* (CQM)—was adopted as the analytical framework for evidence reduction and interpretation. Grounded Theory was originally developed in 1967 by researchers Barney Glaser and Anselm Strauss.

Schettini and Cortazzo (2015) outline four key phases of the CQM method, which we followed during our analysis:

1. Comparison of events corresponding to each category.
2. Integration of categories and their properties.
3. Delimitation of the theory.
4. Theory writing.

Although the classical tradition upheld by Grounded Theory is the hypothetico-inductive model, this does not preclude the use of a deductive approach. In our study, we opted for a deductive categorisation of discourse. This choice was based on the predefined categories aligned with the structural blocks of the THAM-2 instrument for assessing metalinguistic awareness.

Acknowledging the active role of the researcher throughout the qualitative phase is fundamental for ensuring methodological transparency and rigour. According to Olmos-Vega, Stalmeijer, Varpio and Kahlke (2022), reflexivity involves ongoing self-examination of how the researcher's background, experiences, and positionality shape data collection and analysis. As reflexivity is a continuous, collaborative, and multifaceted practice, the researcher maintained a reflexivity journal documenting decisions, assumptions, and emotional responses during the focus group sessions. This process aligns with best practices in qualitative research, which recognize the researcher as a key instrument in knowledge construction. The reflexive log helped identify moments of positional bias—such as the tendency to guide responses—and to adjust moderation strategies, thereby enhancing interpretative credibility.

4. RESULTS

4.1. Statistical results

The graph below and table 1 show the mean scores of the two tests performed and the descriptive statistics respectively.

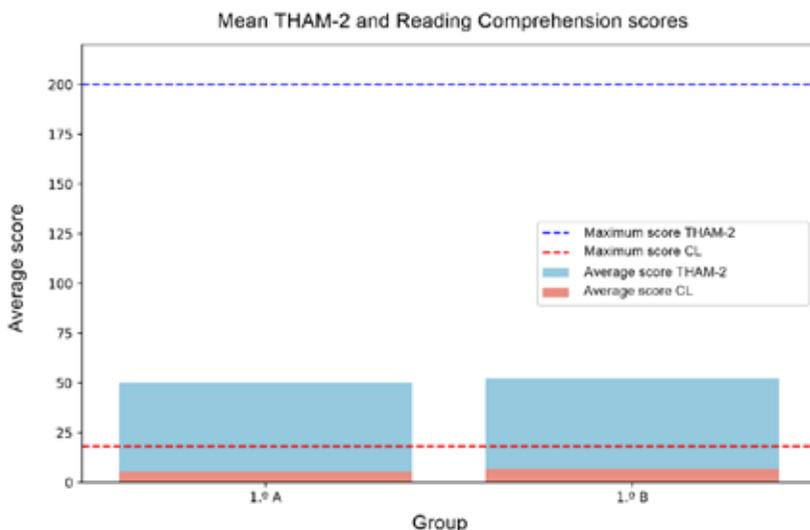


Figure 1. Mean THAM-2 and Reading Comprehension (CL) scores in the two grades: Grade 1A and Grade 1B. Source: SPSS v.21 and Julius AI.

Table 1. Descriptive statistics for metalinguistic awareness (MC) and reading comprehension (CL) totals.

	N	Minimum	Maximum	Mean	Standar dev.
Group	64	1	2	1,50	,504
Genre	64	1	2	1,50	,504
TOTALCM	64	14	84	51,23	15,049
TOTALCL	64	1	13	5,98	2,433
N Valid number (according to list)	64				

Note: SPSS v.21.

Both the graph and the table illustrate that the highest values in metalinguistic awareness in the sample are notably distant from the average test score. Only one student achieved a score of 84 (out of 200) on the THAM-2. This student is the representative of the ML2 level who later participated in the discussion group. In the reading comprehension test, this student also achieved the highest score of 13 out of a possible 18, which is the maximum score.

Table 2. Correlation between reading comprehension tests and THAM-2.

		TOTALCM	TOTALCL
TOTALCM	Correlación de Pearson	1	,796**
	Sig. (bilateral)		,000
	N	64	64
TOTALCL	Correlación de Pearson	,796**	1
	Sig. (bilateral)	,000	
	N	64	64

Note: SPSS v.21.

Furthermore, the correlation is significant at the 0.01 level (bilateral), indicating a statistically significant relationship between the two variables, with less than a 1% probability that this result is due to chance.

The following graph presents the results of the previous table in a more visual format. The blue dots on the scatter plot represent the students' THAM-2 (TOTALCM) and reading comprehension test (TOTALCL) scores. Each point corresponds to a student, reflecting their performance on both tests. The orange regression line illustrates the overall trend of the relationship between these two variables.

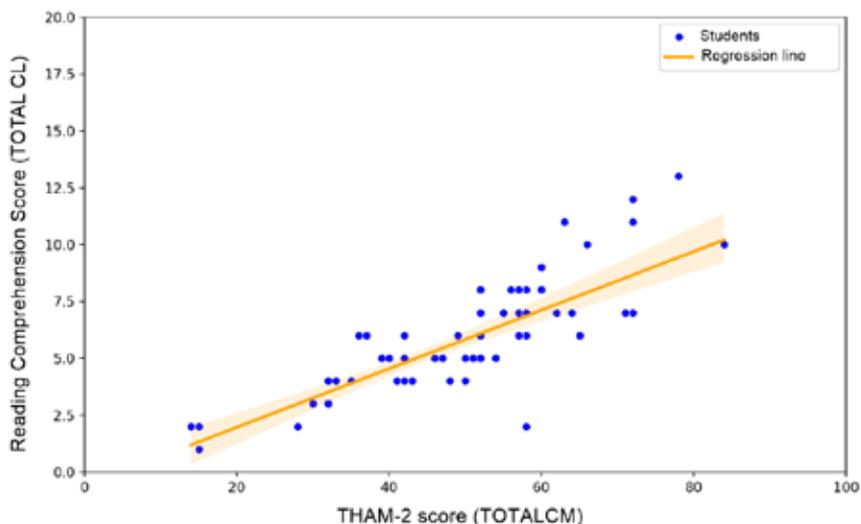


Figure 2. Correlation between metalinguistic awareness (THAM-2) and reading comprehension. Source: SPSS v.21.

The correlation coefficient of 0.796 reveals a strong positive relationship between the THAM-2 and reading comprehension test scores. As the THAM-2 score increases, so does the reading comprehension score. In other words, students who perform better on one test also tend to perform better on the other.

The overall performance of both groups suggests a generally low-to-moderate level of reading comprehension and metalinguistic awareness.

4.2. Results from the discussion group

Starting with the defining characteristics of discussion groups, ideally, a dialogue or conversation should unfold. However, when working with young learners, facilitating a meaningful dialogue that leads to a thorough discussion can be challenging. Despite this, some responses, particularly as the session progressed, showcased moments of discussion and elaboration.

Once the student participants had completed the THAM-2 and their results were corrected, three students were selected to represent each of the three levels of metalinguistic awareness established by the test: ML0 (boy), ML1 (girl), and ML2 (boy). These three students came from different countries: ML0 from Ukraine, ML1 from Spain, and ML2 from Morocco. This allowed for a comparative study based on differences in their L1 (first language). After the group session, a verbatim transcription of the discussion and the field notes taken by the researcher were made.

During the focus group, we adhered to the protocol established by the experts cited as references to avoid difficulties in analyzing the data and to ensure the highest level of rigor in the research. As Barbour (2013) explains, the researcher's presence and demeanor can shape both the type of information participants share and the way they express it during focus groups, as is common across qualitative research methods. Therefore, the researcher aimed to act as a neutral moderator, without influencing the students' responses. No indication was given as to whether their answers were correct or incorrect, which led to the students correcting each other at various points during the discourse. In addition to recording the session, detailed field notes were taken to further enrich and analyze the data obtained. An example is provided below: [*Field Notes*] "ML2 responds coherently and provides a detailed explanation: "Who performs the action in the sentence", adding that the sentence is in the passive voice. This demonstrates both semantic and syntactic processing of the utterance (the relationship between the patient subject and the agent complement when switching from active to passive voice), as well as explicit knowledge of this type of periphrastic passive sentence. The use of this construction in Spanish is less common than structures using "se" or corresponding active constructions. ML1 nods in acknowledgment. ML0 comments again, noting that it is the same as with the first sentence, indicating that his analysis is similarly limited to semantics in a more general sense". Another response from the ML2 participant illustrates a clear understanding of grammatical voice: "Well, this sentence is passive, and in these types of constructions, the subject does not carry out the action of the verb but rather receives it." This statement reflects not only accurate syntactic analysis but also metalinguistic awareness of the functional role of the subject within passive structures.

At all times, the ethical standards for this type of study were strictly adhered to. The students' parents were informed in advance and provided their consent. Their names were kept confidential, and they were identified according to the level of metalinguistic awareness they achieved on the THAM-2.

The focus group lasted 51 minutes and 40 seconds. From the field notes, some general observations were made about the dynamics of the process based on the dialogue that took place:

- Most of the time, students tended to repeat the answer given by the first participant to speak.
- The answers they provided in the test were typically brief and lacked detailed reasoning.
- Overall, they found the test easy but felt it was too long.
- The interaction between ML1 and ML2 was smooth, although ML2 appeared more self-conscious during the focus group than when completing the test.

This comparison allows the researcher to construct categories and define their characteristics based on recurring patterns. Since the group discourse was categorized using MAXQDA version 12, the students' responses were compared by integrating the categories and their properties. Only those parts of the text where the linguistic response (L) was correct and aligned with its corresponding metalinguistic argumentation (ML), which was appropriately justified, were categorized.

4.3. Categorisation of meaning and content analysis: Meta-categories and categories

The meta-categories and categories of analysis have been derived primarily from the blocks that structure the THAM-2, i.e., from the linguistic and metalinguistic skills it measures. These have been organized as follows:

Table 3. *Metacategories, categories of analysis and their respective codes.*

Meta-categories	Categories	Codes
Comprehension	Distinction of relevant words	DPR
	Understanding the meaning of sentences	CSF
	Understanding the overall meaning of the message	CSG
Synonymy	Total or conceptual synonymy	SNC
	Contextual or partial synonymy	SNP
	Referential synonymy	SNR
	Connotative synonymy	SNC
Acceptability	Detection of pertinence	DPT
	Detección of relevance	DRL
Ambiguity	Different interpretations of the text	DIT
	Lexical ambiguity	AML
	Morphological ambiguity	AMM
	Semantic ambiguity	AMS
	Pragmatic ambiguity	AMP
Gramatical Function	Recognition of pronouns and their functions	RDP
	Recognition of articles and their functions	RDA
	Recognition of adjectives and their functions	RDJ
	Recognition of verbs and their functions	RDV
	Recognition of nouns and their functions	RDS
	Recognition of prepositions and their functions	RDR
	Recognition of adverbs and their functions	RDD
Phonemic segmentation	Identification of syllables and phonemes in initial positions	SFI
	Identification of syllables and phonemes in final positions	SFF
	Correct syllable counting	CCS
	Capturing the distinctive value of phonemes	CVF
Metalinguistic Function	Argumentation that aligns with the linguistic response (L)	AAL

Note: own elaboration.

The discourse generated in the focus group was transcribed verbatim and subjected to a process of content analysis through categorisation. For this purpose, the text was divided into content units (Bardin, 1986), to which the codes detailed in Table 3 were applied. As the parts of the text where the students answered correctly (L and ML) were categorised, the Comparative Qualitative Framework (CQM) was applied to generate the conclusions derived from this research.



Figure 4. Most commonly used code system and number of times each. Source: MAXQDA v12.

Figure 4 presents the codes used in the analysis, organized by colour according to their corresponding meta-category (e.g., orange for Phonemic Segmentation, pink for Grammatical Function, blue for Ambiguity, etc.). The figure provides a visual representation of both the relative size of each category’s coding and the frequency with which each category was coded—i.e., the number of instances in which both the linguistic (L) and metalinguistic (ML) responses were deemed correct. For instance, the category DPT (Detection of Pertinence) was coded 20 times.

It is important to highlight that a majority of these percentages and instances stem from the responses of the student classified at the ML2 level. Throughout the categorisation process, multiple comparative analyses were conducted across the three selected cases (ML0, ML1, and ML2). The statistical outcomes are thus enriched by the content-based findings, with the following considerations in mind:

- In the statistical analysis, only those portions of the text where both the linguistic (L) and metalinguistic (ML) responses were correct were coded.
- The content analysis further revealed that in the majority of cases where the ML response was accurate, the corresponding L response was also correct.

Particularly noteworthy are certain excerpts in which the ML2-level student corrected their peers (ML0 and ML1), offering explanations that enabled them to recognise and understand their errors. One such instance, documented in the field notes, reads: [*Field note*]: “It is ML2 who intervenes to say that it is an adjective and knows that adjectives and nouns must agree in gender. ML1 and ML0 understand his explanation and decide to change their answer.” In more complex test items, the ML2 student consistently provided well-reasoned justifications for their responses. For example: [*Field note*] Item 1.A: “The princess kissed the frog.” Question PL: “Who was kissed?” ML2’s response: “The frog.” Question ML: “What makes you certain that it was the frog?” ML2’s explanation: “Well, sentences are made up of subject and predicate, and within the predicate, the verb is the most important element because it indicates who performs the action. That’s why it must agree in number with the subject.”

This content analysis supports the conclusion that, in general terms, a well-formulated and accurate metalinguistic explanation tends to correlate with a correct linguistic response and, by extension, with higher reading comprehension skills.

5. CONCLUSIONS

Starting from the primary objective of this research—to examine the extent to which metalinguistic skills influence reading comprehension in a specific student population, and whether multilingualism enhances metalinguistic awareness—the findings allow us to draw several key conclusions.

The results obtained in this study, supported by both statistical and graphical evidence, allow us to draw several important conclusions. First, the data confirm a strong and positive correlation between metalinguistic awareness and reading comprehension. However, overall performance on both the THAM-2 and the reading comprehension test was moderate to low. In particular, none of the participants reached 50% of the total possible score on the metalinguistic awareness assessment.

One student, however—the multilingual participant fluent in Arabic, Spanish, and French—achieved the highest scores in both assessments. This finding reinforces previous research that highlights the positive influence of multilingualism on linguistic awareness and cognitive processing. Nevertheless, in general, the metalinguistic (ML) responses provided by the students tended to be either incorrect or lacking in sufficient reasoning or depth of explanation.

When ML responses were successful, they typically coincided with correct linguistic (L) answers, confirming the interrelationship between linguistic competence, metalinguistic reflection, and reading comprehension. Yet, this correlation was not consistently observed across all levels. For instance, the ML0 participant produced some accurate linguistic answers that were not supported by valid or well-reasoned metalinguistic justifications. Participants at the ML1 and ML2 levels showed more frequent accuracy in their linguistic responses, often accompanied by more developed reasoning—particularly in the case of ML2. However,

even their responses did not reach high levels in the categorisation process, suggesting that, while their performance was stronger than that of the ML0 student, it was still limited in terms of depth and consistency.

Finally, all three participants—especially those at the ML0 and ML1 levels—demonstrated a limited grasp of key metalinguistic concepts such as grammatical categories and basic syntactic structures. Throughout the focus group discussion, the students consistently reported that this type of activity is rarely addressed in their regular classroom instruction, which may partially explain their difficulties in articulating reflective and accurate responses.

As with any study, particularly in the field of education, certain limitations and challenges emerge—these should be understood not as shortcomings but as opportunities for enhancing future research. One of the main challenges encountered during the focus group was the limited elaboration in students' discourse, which restricted a more in-depth analysis of its content. Although all three participants described themselves as close friends and classmates, there was noticeable hesitation and self-consciousness in their oral responses and argumentation. This discomfort can be attributed to their age, as speaking publicly and exposing one's potential mistakes in front of peers often generates anxiety.

Additionally, the findings underscore the need for further exploration into classroom practices concerning the development of metalinguistic skills. Greater attention to how these skills are taught, understood, and applied is essential, as their careful study and integration into pedagogy can significantly enhance language learning and literacy outcomes.

Strong reading comprehension skills are widely recognized as a key factor in students' academic achievement and broader educational development. In this regard, teachers play a crucial role in fostering these skills through well-designed instructional strategies. Various approaches to reading instruction—such as those proposed by Colomer (1997), Quintana (2000), Tapia (2005), and Santamarina (2015)—can be adapted to meet the diverse needs and backgrounds of students. Within these pedagogical practices, the intentional development of metalinguistic skills should not be overlooked, as these abilities have been shown to support students' ability to interpret, reflect upon, and produce written language more effectively. Enhancing students' metacomprehension may therefore contribute meaningfully to their reading competence and overall academic performance.

In conclusion, the findings of our study support previous research, reinforcing the view that the development of metalinguistic awareness—and the skills that stem from it—is fundamental to the acquisition of reading, among other core academic abilities.

When considering the academic benefits of multilingualism, it is important to highlight that its influence extends beyond the enhancement of linguistic competence to include significant gains in cognitive and metalinguistic skills. These gains, in turn, support deeper comprehension and more effective language use. This is clearly reflected in the well-reasoned and accurate responses provided by the ML2 learner in this study. Grant, Dennis, and Li (2014, p. 1) have articulated this in their research: “In recent years, bilingualism has been linked to both advantages in executive control and positive impacts on aging. Such positive cognitive effects of bilingualism have been attributed to the increased need for language control during bilingual processing and increased cognitive reserve, respectively”.

Findings from the present study align with those of Lasagabaster (2010, 2025), Barquín (2024), and Chakkor and Charia (2023), all of whom suggest that multilingual individuals develop heightened awareness of language as a functional tool. This awareness enables them to interpret

implicit meanings, draw inferences, and employ language strategically to achieve communicative goals. Moreover, multilinguals tend to exhibit superior abilities in filtering, retrieving, and retaining information, thanks to more efficient cognitive processing and enhanced memory function. Their ability to operate across multiple linguistic codes not only broadens their worldview but also fosters greater empathy and attunement to the perspectives and needs of others.

6. REFERENCES

- Barbour, R. (2013). *Introducción a los métodos cualitativos*. Morata.
- Barquín, J. (2024). *La competencia comunicativa en contextos bilingües*. Editorial Académica Española.
- Bialystok, E. (2004). *Bilingualism in Development: Language, Literacy, and Cognition*. Cambridge University Press.
- Bialystok, E. (2005). The impact of bilingualism on language and literacy development. In T. K. Bhatia & W. C. Ritchie (Eds.), *The Handbook of Bilingualism* (pp. 577–601). Blackwell Publishing. 10.1002/9780470756997.ch23
- Bialystok, E. (2024). Cognitive development in bilingual children: A review. *Journal of Child Psychology*, 45(2), 123–135.
- Bizama, M., Gutiérrez, M., & Rojas-Barahona, C. A. (2017). Desarrollo de la conciencia fonológica en niños preescolares: Un estudio longitudinal. *Revista de Psicología*, 35(1), 45–60.
- Calero, M. D. (2011). *Evaluación y mejora de la inteligencia emocional en el aula*. Alianza Editorial.
- Callejo, J. (2001). *Los grupos de discusión: Introducción a una práctica de investigación*. Ariel.
- Carbajal-Carrera, M. (2025). *Estrategias metalingüísticas en la enseñanza de la lectura*. Editorial Universitaria.
- Cenoz, J. (2003). The additive effect of bilingualism on third language acquisition: A review. *International Journal of Bilingualism*, 7(1), 71–87. 10.1177/13670069030070010501
- Cenoz, J., & Gorter, D. (2015). *Multilingual education: Between language learning and trans-languaging*. Cambridge University Press. 10.1017/CBO9781139105729
- Chakkor, M., & Charia, A. (2023). La educación bilingüe en contextos multiculturales. *Revista Internacional de Educación*, 29(3), 210–225.
- Coloma, C., Fuentes, C., & González, M. (2015). Conciencia fonológica y su relación con el aprendizaje de la lectura en niños de primer año básico. *Revista de Psicología*, 24(2), 89–102.
- Colomer, T. (1997). *La enseñanza de la literatura como construcción de sentido*. Graó.
- Defior, S. (1990). *Procesos de lectura y escritura en la infancia*. Síntesis.
- Denzin, N. K. (1978). *The research act: A theoretical introduction to sociological methods* (2nd ed.). McGraw-Hill.
- Fernández de Haro, E., Santamarina, M., J. C., & Romero, I. (2009). Adquisición de habilidades metalingüísticas y enseñanza-aprendizaje de la composición escrita en educación primaria: reflexiones didácticas sobre los resultados de un estudio experimental. *Porta Linguarum*, 12, 149–167.
- Fernández, M. (2024). Estadísticas sobre el rendimiento académico en educación primaria. Instituto Nacional de Estadística.
- Fernández-Portero, C., & Morera-Bañas, M. (2025). La competencia lectora en la educación secundaria: Un estudio comparativo. *Revista de Investigación Educativa*, 43(1), 77–95.
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. In L. B. Resnick (Ed.), *The Nature of Intelligence* (pp. 231–236). Lawrence Erlbaum Associates.

- Floquet, M., & Melogno, S. (2024). *Desarrollo del lenguaje y dificultades de aprendizaje*. Paidós.
- González Manjón, A. (2015). *La comprensión lectora en la educación secundaria*. Narcea.
- Grant, A., Dennis, M., & Li, W. (2014). Bilingualism and cognitive control: A meta-analysis. *Journal of Cognitive Psychology*, 26(5), 519–541.
- Hammer, C. S., Lawrence, F. R., & Miccio, A. W. (2023). Bilingual children's language abilities and early reading outcomes in Head Start and kindergarten. *Language, Speech, and Hearing Services in Schools*, 54(1), 1–15.
- INE. (2022). Estadísticas del sistema educativo español. Instituto Nacional de Estadística.
- Lasagabaster, D. (2010). The relationship between language learning strategies and EFL proficiency among Spanish university students. *International Journal of Applied Linguistics*, 20(2), 169–190. 10.1111/j.1473-4192.2010.00235.x
- Lasagabaster, D. (2025). Multilingualism in the Basque Country: Achievements and challenges. *Journal of Multilingual and Multicultural Development*, 46(3), 310–325. 10.1080/01434632.2024.1234567
- Loureda, O. (2009). *La competencia pragmática en la enseñanza del español como lengua extranjera*. Arco Libros.
- Marsh, D., Mehisto, P., Wolff, D., & Frigols, M. J. (2020). *Content and Language Integrated Learning (CLIL): A Development Trajectory*. Cambridge University Press.
- Mattingly, I. G. (1972). Reading, the linguistic process, and linguistic awareness. In J. F. Kavanagh & I. G. Mattingly (Eds.), *Language by Ear and by Eye* (pp. 133–147). MIT Press.
- Milian, M., & Camps, A. (2000). *La enseñanza de la lengua escrita en la educación primaria*. Graó.
- Navarro, H. (2021). La gestión de la calidad en investigación cualitativa. *Revista de Ciencias Sociales*, 46, 112-117.
- Núñez, M.^a P. & Santamarina, M. (2021). Dificultades y retos para la enseñanza y el aprendizaje de la lectoescritura: aportaciones de un estudio de caso colectivo. *Lengua y habla*, 25, 242-267. <https://hdl.handle.net/10481/99680>
- Núñez, M.^a P., & Pinto, M. A. (2015). The validation of the THAM-2 (Test de Habilidades Metalingüísticas n.º 2). *Rivista di Psicolinguistica Applicata*, 15(2), 31–59. 10.7346/123456789
- Oliva, P. & Santamarina, M. (2015). An approach to integration: The integration of language and content to promote L2 learner autonomy at the college level. *Educational and Learning Research Journal*, 10, 92–112. 10.1234/elrj.v10i1.5678
- Olmos-Vega, F. M., Stalmeijer, R. E., Varpio, L., & Kahlke, R. (2022). A practical guide to reflexivity in qualitative research: AMEE Guide No. 149. *Medical Teacher*, 45(3), 241–251. <https://doi.org/10.1080/0142159X.2022.2057287>
- Patton, M. Q. (1999). Enhancing the quality and credibility of qualitative analysis. In *Handbook of qualitative research* (2nd ed., pp. 541–566). Sage Publications.
- Pérez López, M. (2022). *La enseñanza del español como lengua extranjera: Nuevas perspectivas*. Ediciones Universidad de Salamanca.
- Pinto, G. (1999). *Desarrollo del lenguaje y aprendizaje escolar*. Ariel.
- Pinto, M. A., & Eleuch, F. (2015). El desarrollo de la conciencia metalingüística en niños de primaria: Un estudio comparativo. *Revista de Psicología Educativa*, 21(2), 101–112.
- Santamarina, M. (2015). La comprensión lectora: aspectos teóricos y didácticos. In J. Mata, M.^a P. Núñez & J. Rienda, *Didáctica de la lengua y la literatura*, (pp. 114-152). Pirámide.