

# The Evolution of Spanish Word Order in *El Cantar de Mio Çid*

## *La evolución del orden de palabras en El Cantar de Mio Çid*

XOSE ANTONIO PADILLA GARCÍA

Dpto. de Filologia espanyola, Lingüística general i Teoria de la literatura  
Àrea de Lengua Española  
Universitat d'Alacant  
xose.padilla@ua.es  
<https://orcid.org/0000-0003-4333-3174>

RECIBIDO: 21 DE ENERO DE 2024  
ACEPTADO: 15 DE JULIO DE 2024

**Abstract:** Our work has two primary objectives: a) to examine how the transition from SOV to SVO is manifested in *El Cantar de Mio Çid* (CMC); and b) to explore the reasons that can explain the different word orders found in the text. To test our hypothesis, we structured our study into three parts. Firstly, we conducted a typological description of the CMC by analysing its 3800 verses. Secondly, we applied the chi-square technique to assess whether there is a statistically significant association between the orders of the text and several descriptor variables (verb type, hemistich, weight [number of syllables], context [givenness], and poetic [rhyme]). In the last part of the study, we will discuss the results and present the conclusions. The obtained data indicate that SVO order can be defined as the basic order in CMC. Nevertheless, our data also suggest that other orders (SOV, OVS) can coexist and, in some cases, fluctuate in CMC. This variability can be attributed to typological, pragmatic and informational factors.

**Keywords:** Basic Pattern. Mio Çid. Syntactic Evolution. Typology. Word Order.

**Resumen:** Nuestro trabajo tiene dos objetivos principales: a) examinar cómo se manifiesta la transición de SOV a SVO en *El Cantar de Mio Çid* (CMC); y b) explorar las razones que pueden explicar los diferentes órdenes de palabras que se encuentran en el texto. Para comprobar nuestra hipótesis, estructuramos nuestro estudio en tres partes. En primer lugar, realizamos una descripción tipológica del CMC mediante el análisis de sus 3800 versos. En segundo lugar, aplicamos la técnica chi-cuadrado para evaluar si existe una asociación estadísticamente significativa entre los órdenes del texto y una serie de variables descriptoras (tipo de verbo, hemistiquio, peso [número de sílabas], contexto [dada] y poética [rima]). En la última parte del estudio, discutiremos los resultados y presentaremos las conclusiones. Los datos obtenidos indican que el orden SVO puede definirse como el orden básico en el CMC. Sin embargo, nuestros datos también sugieren que otros órdenes (SOV, OVS) pueden coexistir y, en algunos casos, fluctuar en el CMC. Esta variabilidad puede atribuirse a factores tipológicos, pragmáticos e informativos.

**Palabras clave:** Patrón básico. Mio Çid. Evolución sintáctica. Tipología. Orden de palabras.

### CÓMO CITAR ESTE ARTÍCULO

Padilla García, Xose Antonio. 2025. «The Evolution of Spanish Word Order in *El Cantar de Mio Çid*». *Rilce* 41.2: 723-43. DOI: <https://doi.org/10.15581/008.41.2.723-43>



Traditionally, one of the most significant linguistic shifts in the transition from Latin to Romance languages involves the transformation of the basic word order from SOV (Subject, Object, Verb) to SVO (Subject, Verb, Object).<sup>1</sup> The exact moment and reasons for this change are difficult to determine. However, it seems evident that this shift could not have occurred abruptly; instead, it unfolded through several phases or stages, influenced by various factors (Pinkster 1991; Gimeno 2016).

Within the context of the Spanish language, the *Cantar de Mio Çid* (hereafter referred to as CMC)<sup>2</sup> is a text of particular interest for the research of the word order evolution. The relevance of CMC is anchored in two key attributes: a) it stands as the oldest Castilian text, tracing its origins to the 11th (Menéndez Pidal 1950) or 12th centuries (Smith 1972), and b) it holds the distinction of being an (oral) epic poem, categorized as a minor genre in contrast to religious or juridical texts. This last characteristic potentially afforded the CMC a certain degree of flexibility, by enabling a meaningful depiction of the substantial syntactic transformations occurring during that period.

Scholars' perspectives on CMC's basic word order vary widely. Some researchers argue for a VSO basic order, similar to Arabic (López-García 2000; Bossong 2006; Pinto 2015), others assert an SOV pattern, reminiscent of Basque and/or classical Latin (Montgomery 1977; Alonso 1962; Menéndez Pidal 1950). Alternatively, some authors suggest an SVO basic order, aligning with modern Spanish<sup>3</sup> (Suárez 2017).

More recent works have not specifically studied word order in the CMC, but they have examined word order in general medieval Spanish (12th-15th centuries). These studies deal with two main inquiries: a) revisiting the basic order of the language (Meyer-Hermann 1988); and b) exploring the potential categorization of medieval Spanish as a v2-language<sup>4</sup> (López Izquierdo/Casti-

- 
1. Comparative linguistics indicates that, in the instances of syntactic change, the predominant direction has been from SOV to SVO (Bichakjian 1987; 1988; Gell-Mann/Ruhlen 2011).
  2. In our study, we rely on the edition by Smith (1972).
  3. Like all Romance languages, Spanish is typically classified as a Subject-Verb-Object (SVO) language. Besides the canonical word order of SVO, it exhibits all the typological characteristics typical of SVO languages, such as the use of prepositions (instead of postpositions) or postnominal genitives (instead of prenominal) (Greenberg 1960; 2005; López-García 2000).
  4. In modern German, for instance, which follows SVO as its main order, the second position in the schema must consistently and mandatorily be filled by the v: Ich (s) *kaufe* (v) die Zeitung (o) jeden Tag [I (s) *buy* (v) the newspaper (o) every day]. If the initial position of the sentence is occupied by a constituent other than the subject (topicalization), the v is retained, nonetheless, in the second position (v2): Die Zeitung (o) *kaufe* (v2) ich (s) jeden Tag or Jeden Tag *kaufe* (v2) ich (s) die Zeitung (o) (Elvira 2015a; 2015b).

llo Lluch 2015). This last question refers to whether medieval Spanish was a language in which the verb must appear obligatorily in the second position of the sentence (verb-second constraint).<sup>5</sup> Opinions on this matter vary among researchers. Authors such as Adams (1976; 1977a; 1977b), Warner (1980), or more recently, Bauer (2009) or Wolfe (2015; 2016) support this proposal; researchers such as Batllori (2015), Sitaridou (2011; 2015), or Elvira (2015a; 2015b) are against.

In order to explore to what extent the CMC fits the different typological descriptions and hypothesis, we have designed a study divided into three parts. In the first part, we will define the concepts and methodology and we will approach the typological description of the CMC. For this, we will analyse the 3800 verses comprising the poem, which belong to the Cantares I, II and III. We will also establish which order is the most frequent in the text, providing data on the higher or lower percentages of each of the six possible orders.<sup>6</sup> Secondly, we will analyse which variables could explain the higher or lower frequency of these orders in relation to the topic position (TOP). We will apply the chi-square test to determine if there is a statistically significant association between the orders found in the text and a series of explanatory variables that we will define later on. In the last part of the study, we will discuss the results obtained and present the conclusions derived from the analyses conducted.

In sum, our work has two main objectives: a) to examine how the transition from SOV to SVO is manifested in the CMC; and b) to explore the (typological, pragmatic, informational) reasons that can explain the different orders that appear in the text.

#### WHAT IS THE BASIC WORD ORDER?

Establishing the basic word order in a given language is always a process of simplification. The syntax of languages consists of various constructions, which can influence the decision to define a main order. Similarly, the form of the verb (transitive, intransitive, personal, impersonal), the type of sentence (declarative, interrogative, exclamatory), and its complexity (simple, complex, main, subordinate) could favour one tendency over another. Despite this,

5. That is, [x]v[x] with x = s or o alternatively.

6. Six possible combinations: SVO, SOV, VSO, VOS, OSV, OVS. Mathematically, this is expressed as  $3 \times 2 \times 1 = 6$ .

typological studies (Greenberg 1960; 2005; Comrie 1981; Dryer 2011; 2013; etc.) tend to use the *simple transitive declarative sentence* as a pattern, and this choice appears to be justified by semantic and perceptual reasons.

Transitive declarative sentences align with prototypical semantic descriptions. The semantics of the transitive sentence reflect how actors are organized in the world (López-García 2000; Kemmerer 2012). According to this description, in the most common scenario, a volitional *agent*, usually human, acts upon a *patient*, usually inanimate, changing its state or position:

- (1) *Blanca* (agent) [s] moved (transitive verb) [v] *the chair* (patient) [o] ~  
[S + V + O]

Most languages in the world encode this scenario in this manner (Creissels 2016). Semantic descriptions seem to suggest, therefore, the prominence of this type of verb over others, justifying in some way the connection between argument structure and syntax.

This prototypical pattern also appears to be useful to examine human cognitive processing (Hahn et al. 2020). Neurolinguistics has highlighted the relevance of the basic pattern (with three main components) to analyse how the brain processes information (Gabrić 2019). Some studies, for example, have shown that placing information about the agent (s) before the patient (o) seems to facilitate, from a procedural point of view, predictions about future action (v) (Kemmerer 2012). This way of processing information indirectly justifies the fact that the SVO and SOV orders (subject prominence and verb-object constituent juxtaposition) are the most frequent orders in the languages of the world (Dryer 2013).

Therefore, the basic pattern is both a simple and complex unit at the same time. These two characteristics probably explain its success as a unit for typological classification (Greenberg 2005; Dryer 2011; 2013). Using the basic word order will always mean focusing on a limited part of syntax and may not be easy to apply to all languages. Nevertheless, despite the controversies, the basic pattern can be considered a useful tool for examining the evolution and typological classification of many languages.<sup>7</sup>

7. Examples have been identified for each of the potential word orders in various languages. There are languages with SVO basic order (Mary has a book, English), SOV order (Mariak liburu bat du, Basque), VOS order (Manana boky i María, Malagasy), VSO order (Istarat Mariam kitab, Arabic), or OVS order (Kanawa yano toto [The canoe took the man], Hixkaryana); and languages, such as Apurinã (Brazilian Amazon), that seem to have an OSV basic order (Amaruru n-unama-ka-namu-ta [I am making the child sleep]) (Dryer 2013).

## HYPOTHESES ABOUT BASIC ORDER IN CMC

Applying the basic order pattern to the particular analysis of CMC has resulted in two possible orders, different from (modern Spanish) SVO. Several hypotheses can account for it.

Some studies (Alonso 1962; Menéndez Pidal 1950; Lapesa 1981; etc.) suggest that the SVO and SOV orders in CMC coexist in a particularly striking way. Verb postposition (SOV) is, for these researchers, more common in CMC than in later texts.

According to these studies, two potential hypotheses could explain this pattern. The first is a *literary condition*. The (non-regular) rhyme in the second hemistich forced the author of CMC to make certain changes in word order, and these changes account for the high number of SOV occurrences (Smith 1972; Martínez-Gil 1989). The second hypothesis is related to *linguistic contact*. For some researchers, the SOV order is a Latin archaism, latent or intentional (Alonso 1962; Menéndez Pidal 1950; etc.). For others, the SOV order hypothetically derives from interference from the Basque substrate or adstrate (Montgomery 1977), because the basic order in Basque is SOV, similar to Classical Latin.

As an alternative to the SOV order, some studies propose VSO for medieval Spanish (Hinojo 1988; Martínez-Gil 1989; Blake 1991). A hypothetical connection at the time between medieval Spanish and languages such as Arabic,<sup>8</sup> which also has a VSO word order, could support this pattern in CMC (López-García 2000; Bossong 2006; Pinto 2015).

In sum, opinions on this matter vary among researchers (Padilla 2018). The disparity between them can likely be ascribed to two potential explanations: a) some classic studies exclusively focus on syntactic factors (not informative or pragmatic factors), and b) some studies combine various types of sentences (transitive, intransitive, constructions with clitics, etc.)<sup>9</sup> to establish the typological order.

Our study, as previously mentioned, will adopt the typological canon (declarative transitive simple sentence with three constituents) as a starting point and will add two other structures: copulative (or copula) and participle sentences (Suñer 2015). However, the data from each group will be presented se-

8. This phenomenon could be attributed to the influence of Mozarabic (Christians residing in Muslim regions) scribes, who were proficient in both Arabic and Romance languages.

9. A review of the different types of constructions appears in Padilla (2018).

parately (table 1). In addition to syntactic factors, we will examine the informative and pragmatic factors that could account for the various orders observed in the text.

## METHODOLOGY

The CMC is an oral poetic composition, a *cantar*. Cantars were originally transmitted orally and later transcribed into manuscripts.<sup>10</sup> Its oral origin implies that the text is adorned with numerous expressions tied to a hypothetical dialogue between the minstrel and his audience.<sup>11</sup> In fact, the oral nature of the text is important to provide justification for certain word order usages, as we will see later.

The *Cantar*'s 3800 verses are formed by two hemistiches with internal caesura (H1 + H2); and the rhyme, when it occurs, affects the second hemistich (H2).

From a methodological point of view, our first (descriptive) analysis has examined all the sentences that appeared in the text (*Cantares* I, II and III) provided that they met two criteria:

- a) They are composed of three components (S, V, O),
- b) They belong to one of these three groups:
  - b.1 The sentence is transitive
 

[H1]/ todos [S] dizian [V] una razon [O]<sup>12</sup>  
 everyone said            an    explanation  
 'Everyone provided an explanation'.  

[SVO; H2, verse 13]
  - b.2 The sentence has a copulative verb
 

Alegre [O]    es [V] Myo Çid [S]/ [H2]  
 Happy            is        Myo Çid  
 'Mío Cid is happy'.  

[OVS; H1, v. 2470]

10. Depending on what is deemed more significant in characterizing the text, scholars opt to designate it as either a poem or a *cantar* (Smith 1972).

11. Expressions like the following are not uncommon in CMC:  
 Fablo Martin Antolinez/ odredes            la que    a dicho (v. 70)  
 Spoke Martin Antolinez/ you will hear    what    he said.  
 'Martín Antolínez spoke and now you will hear what he said'.

12. We have used *The Leipzig Glossing Rules* to translate the examples (<https://www.eva.mpg.de/lingua/pdf/Glossing-Rules.pdf>).

- b.3 The sentence has a past participle  
 [H1]/ amos [s] tornados [O] son [V]  
 both returned have  
 ‘Both have returned’.  
 [SOV; H2 v. 2561]

Therefore, this selection excludes intransitive sentences (2) and sentences that omit the (s)VO (3):<sup>13</sup>

- (2) De parte de orient vino [V] vn coronado [s]/ [H2]  
 From the eastern side came a crowned man  
 ‘There came a king from the East’.  
 (v. 1288)
- (3) Tiene [V] dos arcas lennas de oro esmerado [O] / [H2]  
 Has two coffers full of gold fine  
 ‘He has two coffers full of fine gold’.  
 (v. 113)

In the description of the selected corpus, and to check its possible influence, we have also differentiated whether the examined sentence, and its order, occurred in the first hemistich (H1) of the verse; in the second (H2); or it was distributed between the two (H1 + H2).

An example of each case (separated by slashes: /) is shown in (4):

- (4) [H1]:  
 Myo<sup>14</sup> Çid [s] ganno [V] a Alcoçer [O]/ [H2]  
 My Çid won to Alcocer  
 ‘Mío Cid conquered (the city of) Alcocer.’  
 (SVO, H1, v. 610)
- [H2]:  
 [H1]/ Dios [s] sabe [V] el aiuntar [O].  
 God kwows the joining in  
 ‘Only God knows with whom we will end up!’  
 (SVO, H2, v. 373)

13. This option –namely, an omitted subject– is the most frequent in modern spoken Spanish (Padilla 2001; 2005).

14. *Myo* (my, of mine) is a medieval expression which implies both affection and respect. In modern Spanish, the expression would be *mi Cid* (My Cid) or *mi querido Cid* (My dear Cid).

[H1 + H2]:

Grande duelo [O] avien [v]/ las yentes cristianas [S]  
 a lot of sadness had the people Christian  
 ‘The Christian community felt profound sorrow’.

(OVS, H1 + H2, v. 29)

### *Variables in inferential statistics*

In the second phase of our analysis, our objective was to investigate the potential associations between the word orders found in the text and several explanatory variables. Thus, beyond the previously described variables, such as *verb type* and location within a specific *hemistich* (H1, H2, or H1 + H2), this section examines three new additional variables: *weight*, *context*, and *poetic*.

Firstly, the variable *weight* measures the number of syllables in a constituent, exploring Behaghel’s Law of Increasing Terms.<sup>15</sup> According to this law, given two phrases, when possible, the shorter precedes the longer.

In our analysis, a constituent that is made up of a maximum of two syllable is called (*light-*)*weight*. An example is provided in (5):

(5) [one syllable = light-weight]

Dad las a qui quisieredes uos, ca yo [S] pagado [O] so [V]  
 Give them to whomsoever you will, for I glad am  
 ‘Give them to whomever you wish, for I am pleased with it’.

(SOV, H2, v. 2089)

vs.

[seven syllables = non light-weight]

*Los ynfantes*<sup>16</sup> *de Carrion* [S] lieuan [V] grandes conpannas [O]  
 The ynfants of Carrion carry great companions  
 ‘The infants of Carrion have many noble friends’.

(SVO, H1, v. 1981)

15. Behaghel’s Law of Growing Terms (or Constituents) is one of the principles proposed by the German linguist Otto Behaghel. This law postulates that in Indo-European languages, the elements of a sentence are arranged in a specific way, that is, shorter or less important elements precede longer or more important elements (Behaghel 1909; Arnold et al. 2000).

16. *Ynfante* is an ancient nobility title. Today it applies to the sons and daughters of the kings of Spain who are not the direct heirs to the throne.



Secondly, the variable *context* assesses the potential connection of a constituent with the previous discourse (givenness).<sup>17</sup> In (6), for instance, this connection is established through the summarizer-provided deictic element *esto* (it/this).

- (6) Esto [O] me an buelto [V] myos enemigos malos [S]  
 This me have given my enemies evil  
 ‘I have received this from my evil enemies’.  
 (OVS, H1, v. 9)

Lastly, the variable *poetic* examines whether the constituent, in this case located in H2, rhymes with another word from the previous verse. An example is provided in (7) (rhyme = *-idos*).

- (7) Con L vezes mill de armas/ todos [S] fueron [V] *conplidos* [O]  
 With L times thousand of weapons/ all were glad  
 ‘With fifty thousand of weapons, everyone was joyful’.  
 Entraron sobre mar/ en las barcas son *metidos*.  
 They entered on sea/ in the boats are put.  
 ‘They embarked on the sea, placed in the boats’.  
 (SVO, H2, v. 1026)

This second phase of the analysis will assess the connection between the mentioned variables and the first position into the hemistich (H1, H2, or H1 + H2). That is, this analysis will measure the hypothetical statistically significant association between the variables and *topic* (TOP)<sup>18</sup> position.

The identification of a statistically significant association between the observed data and the TOP position would additionally suggest a consistent and regular placement of a specific constituent in the word order pattern.<sup>19</sup> In

17. The concepts of theme/rheme (givenness) work on a syntagmatic (or horizontal) axis; the topic/comment, on the contrary, operates on an imaginary paradigmatic (or vertical) axis linked to the first position of the sentence (Padilla 2005; 2006).

18. Adopting Chafe's (1976) topic/comment classic framework, we designate the TOP position as the initial point for conveying subsequent information. From a *cartographic perspective*, this position aligns, to some extent, with the concept of the *left periphery*, as discussed by Rizzi (2004) and Rizzi/Bocci (2017).

19. Several authors have pointed out the hypothetical connection between the topic and the grammaticalization of the subject in the first position of the basic sentence scheme (Venneman 1974; Givón 1992; Elvira 2015a; 2015b). According to this proposal, the SVO structure would actually respond to the fact that the subject is the sentence constituent that most frequently occupies this position.

such cases, it would imply that the examined constituent is progressing towards a hypothetical grammaticalized (or syntactized)<sup>20</sup> association with this informative place, ceasing to have a pragmatic-informative value. Conversely, the presence of another constituent in this location would indicate a deviation from the basic word order, which means (pragmatic-informative) marked cases (or topicalization).

Based on the above, the six examined orders result in the following variables: S\_TOP (<SVO, SOV), O\_TOP (<OVS, OSV), and V\_TOP (<VOS, VSO).

All the variables in our study are categorical. We used the SPSS statistical program (version 28) for analysis.

## RESULTS

After examining the 381 occurrences that meet the established criteria, out of the 3800 verses that compose the CMC, we have obtained the results shown in the following table:

	TRANSITIVE					COPULATIVE					PARTICIPLE		TOTAL SUM.	
	H1	H2	H1 + H2	TOTAL SUM.	%	H1	H2	H1 + H2	TOTAL SUM.	%		%		%
SVO	24	31	31	86	46	30	16	4	50	37	19	34	155	40
OVS	16	14	16	46	24	43	17	7	67	50	23	41	136	35
VOS	3	–	12	15	8	–	2	1	3	2	1	2	19	5
VSO	4	1	6	11	6	–	–	1	1	1	–	–	12	3
SOV	3	7	15	25	13	1	4	6	11	8	12	21	48	12
OSV	3	1	2	6	3	3	–	1	4	3	1	2	11	3
	53	54	82	189		77	39	20	136		56		381	

Table 1. List of occurrences from CMC.

Table 1 shows that the most frequent basic order in transitive sentences is SVO (46 %). This is true for the analysed examples (381 occurrences) regardless of whether the example was in H1, H2, or it was distributed between the two hemistiches (H1 + H2). When considering the overall data, the percentage of SVO

20. We understand *syntactization* in the sense explained by Lehmann (1992, 425): «Reduction of syntagmatic variability includes the fixation of word order. This is why grammaticalization goes hand in hand with the loss of word order freedom». For syntactization, see also Seiler (2015).

is 40 %. The remaining occurrences are distributed among the other five possible orders (see *supra*).

The statistical analysis of CMC also presents two particularly interesting findings.

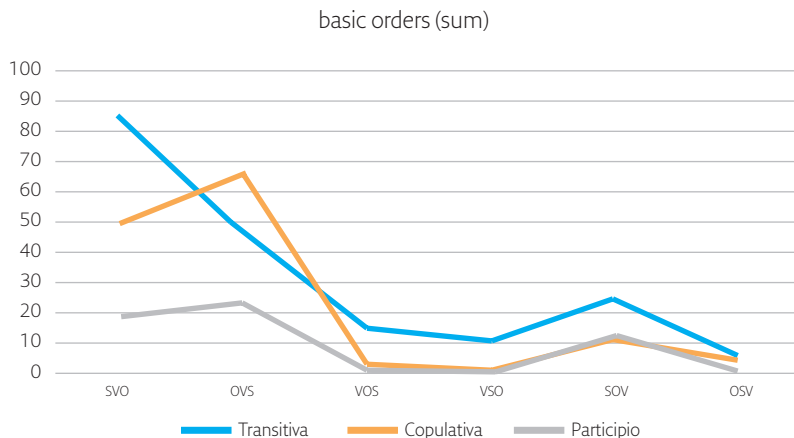


Figure 1. Word order pattern data.

The first finding, as indicated by the highest peak in figure 1, is that OVS is the most frequent order in copulative sentences (50 %) and participle sentences (41 %). Secondly, word orders with a final verb (SOV), unlike in modern Spanish,<sup>21</sup> have a relatively significant percentage (12 % total sum; 13 % transitive; 21 % participle).

### *Inferential Analysis*

The data reported in the descriptive statistical analysis were subjected, in a second phase, to Pearson's chi-square test (figure 2). The goal was to verify if there was a possible association between the examined variables and the initial position of the hemistich (TOP).

$$\chi^2 = \sum \frac{(\text{observed}_{ij} - \text{model}_{ij})^2}{\text{model}_{ij}}$$

Figure 2. Pearson's chi-square test.

21. In contemporary colloquial Spanish, the SOV order is nearly non-existent/absent, accounting for only 1 % (Padilla 2001). Examples of SOV order include sentences such as: Yo (s) un cubata (o) pagaría (v) por verte bailar (*I [s] would pay [v] for a mixed drink [o] to see you dance*).

In order to do this, as previously mentioned, we grouped together those orders with the same constituent in the TOP position. In other words, we merged the general data into three different groups: S\_TOP (<SVO, SOV), O\_TOP (<OVS, OSV), and V\_TOP (<VOS, VSO).

We analysed the possible association of each variable (transitive [yes/no], copulative [yes/no], participle [yes/no], [low-]weight [yes/no], context [yes/no], poetic [yes/no], H1 [yes/no], H2 [yes/no], and H1 + H2 [yes/no]) with one of the possible constituent positions.

## Results

The results derived from the analysis can be summarized as follows. The first important finding is that the variables (low-)weight ( $\chi^2(2) = .720$ ,  $p = .698$ ) and context ( $\chi^2(2) = 2.787$ ,  $p = .248$ ) lacked statistical significance in any scenario.

Based on this result, it can be concluded that neither the preceding context (T/R, givenness) nor the weight (number of syllables) of the word or group of words offer any insight into which constituent occupies the TOP position. In essence, any element, regardless of its syllable count or its association with the preceding context, has the potential to occupy the TOP position.

The variable poetic [rhyme] is, however, significant  $\chi^2(2) = 137.363$ ,  $p < .001$ . In fact, this variable differentiates the V\_TOP option from the other two variables (S\_TOP and O\_TOP). This data reveals that V\_TOP, which represents a small option (8 %), is significantly associated with very specific marked conditions.

The second, and perhaps more important, finding is the association of different orders with a specific hemistich. The pattern O\_TOP is predominantly and significantly associated with H1,  $\chi^2(2) = 27.063$ ,  $p < .001$ ; the placement S\_TOP is associated, on the other hand, with the variables H2,  $\chi^2(2) = 8.365$ ,  $p < .005$ ; and H1 + H2,  $\chi^2(2) = 26.474$ ,  $p < .001$ . There is no definition, in this case, for V\_TOP.

Regarding the type of verb (transitive, copulative, participle), the results are also relevant. The relationship of the transitive variable  $\chi^2(2) = 27.165$ , for S\_TOP,  $p < .001$ ; and the copulative variable  $\chi^2(2) = 15.880$ ,  $p < .001$ , for the O\_TOP order, are statistically significant. Conversely, the association of the participle sentence with any of the three groups of orders is not significant,  $\chi^2(2) = 4.112$ ,  $p = .128$ .

In sum, the data derived from the descriptive analysis find confirmation in the first two cases: with a transitive verb and with a copulative verb.

## DISCUSSION

The data resulting from the statistical analysis provide relevant information about 1) the basic constituent order in CMC, and 2) the pragmatic use of the available possibilities.

The first finding is that the descriptive statistical analysis confirms the SVO order as the most frequent basic order in CMC. This is the case whether we look at sentences with transitive verbs (46 %), which are typologically prototypical, or if we examine the entire set of examples (40 %) in the corpus (table 1). Descriptive percentages are lower for other word order options. Among these alternatives, the VSO order, suggested as basic in certain studies (López-García 2000; Bossong 2006), is infrequent, as indicated in table 1 (see also figure 1).

The data also do not seem to support the definition of CMC as a v2-language (verb-second constraint). In CMC, v2 position prevails in terms of percentages. Nevertheless, v3 position in (8) or even v1 position in (9) are equally possible:

- (8) ca yo [s] pagado [O] so [v3]  
for I paid am  
'Great, I am pleased to hear that'.  
(SOV, H2, v. 2089)
- (9) Meçio [v1] Myo Çid [s] los ombros [O]/ [H2],  
Rocked Myo Çid the shoulders  
'Mío Cid rocked his shoulders'.  
(VSO, H1, v. 13)

The variable of S\_TOP, and consequently, the association of s with the initial position, is statistically significant in transitive sentences; and it happens similarly, when s is placed in two of the hemistich positions: H2 and H1 + H2. That is, the hypothesis proposed by various authors (Venneman 1974; Givón 1992; Elvira 2015a; 2015b) asserting that the subject tends to grammaticalize the TOP (first) position remains valid.

The text, however, also presents significant associations between the TOP position and the V and O constituents, although the relationship between this association and a lower frequency of cases is also revealing.

The V\_TOP order, which is the least frequent (8 %), is significantly associated with the poetic variable. These cases should therefore be explained primarily as a result of a marked context. However, the quantitative insignificance of these examples rejects the variable of poetic mentioned in section 3 for the text as a whole.

Finally, the O\_TOP order (<OVS, OSV) is, as we have indicated, predominant in copulative sentences (50 %) and is significantly associated with the placement of the example in H1.

*What does the text reveal about the transition from SOV to SVO?*

The data from the descriptive and inferential analysis seem to indicate that the transition from SOV to SVO in CMC had been almost definitively completed. The SVO order is the most frequent order in transitive sentences (46 % in transitive sentences; 40 % in the overall data), which generally define the typological description of the language. Furthermore, there is a statistically significant association between S and the TOP position. From the results of the analysis, it can be inferred that the SVO order is the basic pattern, and therefore, the (well-formed) basic scheme from which order changes are explained or generated.

Any modification of the SVO order must hypothetically therefore be attributed to a special communicative, pragmatic or informative meaning.

The OVS order and the association of O in copulative sentences with the TOP position should be considered, in this sense, cases of marked orders (or topicalization):

- (10) *Grandes* [TOP] son [V] los poderes [S]  
Big are the powers  
'The powers are big'.  
(OVS, H1, v. 669)

*Firme* [TOP] son [V] los moros [S]  
Determined are the Muslims  
'The Muslims are determined'.  
(OVS, H1, v. 755)

The statistically significant association of O\_TOP with H1 is equally relevant. H1 is the furthest hemistich from the poetic need for rhyme and therefore could be considered the more free context for varying the order of constituents.

The oral-dialogic nature of the text could additionally support the high frequency of the OVS order. The (poet-)minstrel constructed his text orally in front of an audience. Consequently, he had the need to highlight elements by changing their position in the scheme so that the listeners' attention could be directed towards the most relevant information of the narrative.<sup>22</sup> The fact that the audience hypothetically already knew this information, as a result of the likely regular repetition of the show, also reinforces the need to place the most important informational elements in a relevant (marked) position.

Finally, the SOV word order represents a smaller part (12 %, total sum). Most of the examples that appear in the text are an inversion of participles (21 %):

(11) Çiento e quinze caualleros/ todos [S] *iuntados* [O] son [V]  
all joined are  
'All come together'.  
(SOV) (v. 261)

However, the SOV order in participle clauses is not systematic. The preposed participle fluctuates and coexists throughout the text with the postposed participle, similar to modern Spanish:

(12) Mynaya<sup>23</sup> Albar Fanez      desto es *pagado*:/  
 Mynaya Albar Fanez      of this is paid  
 'My brother, Albar Fañez, is pleased with this'.  
 (...)

22. Some authors define these uses as *narrative inversion* (Fontana 1993).

23. The expression *my naya* is a mixture of Spanish and Basque: *my* (in modern Spanish *mí*) is equivalent to *my* in English, and (*a*)*naya* (in modern Basque *anaia*) means *brother*. The translation could be: *my brother* or *my brother in arms*.

por ir      con      el omnes      son *contados*  
 for going with      the men      are few  
 ‘The men who go with him are very few’.  
 (v. 826)

There are no significant statistical data to support the association between the participle clause and a specific type of order.

Despite the above, the importance of examples with SOV order cannot be relativized, as they still sound strange to modern Spanish speakers (well-formed condition). Examples as in (13) could therefore indicate the presence of traces of an incomplete process (SOV>SVO shift). This unfinished process is most likely, only present in written language:

(13) todos (s) *iuntados* (o) son (v)  
 all              joined              are  
 ‘All come together’.

The SOV order of Basque, in a hypothetical bilingual environment (Basque-Spanish) (López-García 2000), may have functioned as a facilitating substrate/adstrate force.

## CONCLUSIONS

Our study aimed to achieve two primary objectives: a) to examine how the transition from SOV to SVO is manifested in the CMC; and b) to explore the (typological, pragmatic, informational) reasons that can explain the different orders that appear in the text.

To test our hypothesis, we structured our study into three parts. Firstly, we conducted a typological description of the CMC by analysing its 3800 verses. Secondly, we applied the chi-square technique to assess whether there is a statistically significant association between the orders of the text and several descriptor variables (verb type, hemistich, weight [number of syllables], context [givenness], and poetic [rhyme]). In the last part of the study, we will discuss the results and present the conclusions.

The key findings from the analyses offer valuable insights into the typological description of CMC and the word order variations within the text. Significantly, the SVO order emerges as the basic order, constituting 46 % of transitive sentences and 40 % of the overall examples. However, the OVS order prevails in copulative sentences (50 %) and participle sentences (41 %).



Upon evaluating these initial descriptive results, the chi-square test indicates a non-significant result for the variables of context (givenness) and weight but it establishes statistical significance for the variable of poetic, defining the marking of the V\_TOP scheme. The chi-square test consistently identifies a statistically significant association between the TOP position in the first hemistich (H1) and the S\_TOP scheme in transitive sentences, as well as between the O\_TOP scheme and copulative sentences. No significant relevance is found for participle sentences. Moreover, the results reveal a significant association between the variable hemistiches and the examined forms: H2 and H1 + H2 for S\_TOP, and H1 for O\_TOP.

The interpretation of the data supports the prevalence of a basic SVO order and a relationship between the TOP position and subject, aligning with the observations of certain authors (Venneman 1974; Givón 1992; Elvira 2015a; 2015b) who assert the tendency of the subject to grammaticalize the TOP position.

Due to the non-prototypicality of copulative sentences, the O\_TOP scheme may be viewed as a case of marked order (topicalization). The characterization of the text as a *cantar*, coupled with its oral nature, could explain – among other reasons – the frequent utilization of the OVS order to emphasize the most relevant narrative information.

Despite the fact that the SOV order represents only 12 % of the analysed examples, its significance cannot be overlooked. The results suggest that the SOV order (ca yo [S] *pagado* [O] so [V] (SOV), todos [S] *iuntados* [O] son [V] (SOV) (v. 261), etc.) would not sound very unfamiliar in the Middle Ages conversely to modern Spanish. A latent Basque influence (SOV), or a Latin archaism (SOV), as some authors have noted, could have played a reinforcing role.

Finally, regarding the transition from the SOV order to SVO, which is mostly a completed process, our data suggest that other orders could coexist and fluctuate in CMC, particularly the anteposition and postposition of the participle. Additionally, the significant occurrence of OVS order, particularly in copulative sentences, can be ascribed to marked cases and the (originally) oral characteristics of the text.

## REFERENCES

- Adams, James N. 1976. *The text and language of a Vulgar Latin Chronicle (Anonymous Valesianus II)*. London: University of London.

- Adams, James N. 1977a. «A typological approach to Latin word order». *Indogermanische Forschungen* 81: 70-99.
- Adams, James N. 1977b. *The Vulgar Latin of the Letters of Claudius Terentianus*. Manchester: Manchester UP.
- Alonso, Martín. 1962. *Evolución sintáctica del español*. Madrid: Aguilar.
- Arnold, Jennifer, Anthony Losongco, Thomas Wasow y Ryan Ginstrom. 2000. «Heaviness vs. Newness: The Effects of Structural Complexity and Discourse Status on Constituent Ordering». *Language* 76(1): 28-55. <https://dx.doi.org/10.1353/lan.2000.0045>.
- Batllo, Montserrat. 2015. «Análisis comparativo de las construcciones de dislocación a la derecha y su incidencia en el orden de palabras del español, catalán y portugués medievales». En López Izquierdo/Castillo Lluch, 83-110.
- Bauer, Brigitte. 2009. «Word order». En *New Perspectives on Historical Latin Syntax, 1: Syntax of the Sentence*, eds. Pirluigi Baldi y Philip Cuzzolin, 241-316. Berlin: Mouton de Gruyter.
- Behaghel, Otto. 1909. «Die Stellung des Objects im Germanischen». *Indogermanische Forschungen* 24: 1-100.
- Bichakjian, Bernard H. 1987. «The Evolution of Word Order: A Paedomorphic Explanation». En *Papers from the 7<sup>th</sup> International Conference on Historical Linguistics*, eds. Anna Giacalone-Ramar, Onofrio Carruba and Giuliano Bernini. Amsterdam: John Benjamins.
- Bichakjian, Bernard. H. 1988. *Evolution in Language*. Ann Arbor: Koroma.
- Blake, Robert. 1991. «Syntactic aspects of latinate texts». En *Latin and the Romance Languages in the Early Middle Ages*, ed. Roger Wright, 219-32. London: Routledge.
- Bossong, Georg. 2006. «La sintaxis de las Glosas Emilianenses en una perspectiva tipológica». En *Actas del VI Congreso internacional de historia de la lengua española*, eds. J. Jesús Bustos Tovar y José Luis Girón Alconchel, 529-43. Madrid: Arco Libros.
- Chafe, Wallace. 1976. «Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View». En *Subject and topic*, ed. Charles N. Li, 25-55. New York: Academy Press.
- Comrie, Bernard. 1981. *Universales del lenguaje y tipología lingüística*. Madrid: Gredos.
- Creissels, Denis. 2016. *Transitivity, valency and voice*. European Summer School in Linguistic Typology, Porquerolles. <http://www.deniscreissels.fr/public/Creissels-ESSLT.pdf>. [15 January 2023].

- Dryer, Matthew S. 2011. «The evidence for word order correlations». *Linguistic Typology* 15: 335-80. <https://doi.org/10.1515/lity.2011.024>.
- Dryer, Matthew S. 2013. «Order of Subject, Object and Verb». En *The World Atlas of Language Structures Online*, eds. Matthew S. Dryer y Martin Haspelmath. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/81>. [15 January 2023].
- Elvira, Javier. 2015a. *Lingüística histórica y cambio gramatical*. Madrid: Síntesis.
- Elvira, Javier. 2015b. «¿V2 en español antiguo?». En *El orden de palabras en la historia del español y otras lenguas romances*, eds. Marta López y Mónica Castillo, 27-47. Madrid: Visor.
- Fontana, Josep M. 1993. *Phrase structure and the syntax of clitics in the history of Spanish*. Ph.D. Dissertation, University of Pennsylvania.
- Gabrić, Petar. 2019. «The evolution of syntax: What do we know?». *Linguistisches Kolloquium 2019/20*. Johannes Gutenberg-Universität Mainz. [https://www.researchgate.net/publication/343629337\\_The\\_evolution\\_of\\_syntax\\_What\\_do\\_we\\_know](https://www.researchgate.net/publication/343629337_The_evolution_of_syntax_What_do_we_know) (accessed 15 January 2023).
- Gell-Mann, Murray, and Merritt Ruhlen. 2011. «The origin and evolution of word order». *Proceedings of the NASUSA* 108(42): 17290-95.
- Gimeno, Francisco. 2016. «Variación y cambio del latín al romance: el calco sintáctico del orden semítico de palabras». *Aemilianense* 4: 291-365.
- Givón, Talmy. 1992. «On interpreting text-distributional correlations: some methodological issues». En *Pragmatics of Word Order Flexibility*, ed. Doris L. Payne, 305-25. Amsterdam: J. Benjamins. <http://dx.doi.org/10.1075/tsl.22.13giv>.
- Greenberg, Joseph H. 1960. «A quantitative approach to the morphological typology of language». *International Journal of Typology of Language* 26: 178-94. <http://dx.doi.org/10.1086/464575>.
- Greenberg, Joseph H. (1966). 2005. *Language Universals*. Berlin/New York: De Gruyter Mouton. <http://dx.doi.org/10.1515/9783110899771>.
- Hahn, Michael, Dan Jurafsky and Richard Futrell. 2020. «Universals of word order reflect optimization of grammars for efficient communication». *Proceedings of National Academy of Sciences* 117(5): 2347-53. <https://doi.org/10.1073/pnas.1910923117>.
- Hinojo, Gregorio. 1988. «Del orden de palabras en castellano medieval». En *Actas del I Congreso Internacional de la Lengua Española*, eds. Manuel Ariza et al., 437-47. Madrid: Arco Libros.
- Kemmerer, David. 2012. «The cross-linguistic prevalence of SOV and SVO word order reflects the sequential and hierarchical representation of

- action in Broca's area». *Language and Linguistics Compass* 6(1): 50-66.  
<https://doi.org/10.1002/lnc3.322>.
- Lapesa, Rafael. 1981. *Historia de la lengua española*. Madrid: Gredos.
- Lehmann, Christian. 1992. «Word order change by grammaticalization». En *Internal and External Factors in Syntactic Change*, eds. Marinel Gerritsen y Dieter Stein, 395-416. Berlin/Boston: De Gruyter Mouton.  
<https://doi.org/10.1515/9783110886047.395>.
- López-García, Ángel. 2000. *Cómo surgió el español. Introducción a la sintaxis histórica del español antiguo*. Madrid: Gredos.
- López Izquierdo, Marta, y Monica Castillo Lluch, eds. 2015. *El orden de palabras en la historia del español y otras lenguas iberorromances*. Madrid: Visor.
- Martínez-Gil, Fernando. 1989. «Las inversiones del orden de palabras en el romancero». *Hispania* 72(4): 1-27. <http://dx.doi.org/10.2307/343568>.
- Menéndez Pidal, Ramón. 1950. *Orígenes del español*. Madrid: Espasa-Calpe.
- Meyer-Hermann, Reinhard. 1988. «¿Se debe la posposición del sujeto en el español a una influencia árabe?». *Revista de Filología Española* 68(1/2): 67-96.
- Montgomery, Thomas. 1977. «Basque models for some syntactic traits of the *Poema de Mio Cid*». *Bulletin Of Hispanic Studies* 54: 95-99.  
<http://dx.doi.org/10.3828/bhs.54.2.95>.
- Padilla, Xose A. 2001. «El orden de palabras en el español coloquial». Tesis doctoral, Universitat de València.  
<https://www.cervantesvirtual.com/obra/el-orden-de-palabras-en-el-espanol-coloquial--0/>.
- Padilla, Xose A. 2005. *Pragmática del orden de palabras*. Alicante: SPUA.
- Padilla, Xose A. 2006. *La organización informativa del discurso y el orden de palabras*. Madrid: Liceus.
- Padilla, Xose A. 2018. «El orden básico de constituyentes en el español primitivo». En *Estudios lingüísticos en homenaje a Emilio Ridruejo*, eds. Antonio Briz, María José Martínez Alcalde, Nieves Mendizábal de la Cruz, Mara Fuertes Gutiérrez, José Luis Blas Arroyo y Margarita Porcar Miralles, 1041-56. València: Universitat de València.
- Pinkster, Harm. 1991. «Evidence for SVO in Latin». En *Latin and the Romance Languages in the Early Middle Age*, ed. Roger Wright, 69-82. London: Routledge.
- Pinto, Carlos F. 2015. «Los contactos de lengua y el orden vs en la historia del español». En López Izquierdo/Castillo Lluch, 49-82.

- Rizzi, Luigi. 2004. «Locality and Left Periphery». En *Structures and Beyond: The Cartography of Syntactic Structures*, ed. Antonio Belletti, 223-51. New York: Oxford UP.
- Rizzi, Luigi, and Guliano Bocci. 2017. «Left Periphery of the Clause: Primarily Illustrated for Italian». En *The Wiley Blackwell Companion to Syntax*, eds. Martin Everaert y Henk C. Riemsdijk, 1-30. London: Blackwell Publishing Ltd.
- Seiler, Guido. 2015. «Syntactization, Analogy and the Distinction between Proximate and Evolutionary Causations». En *Variation in Language: System- and Usage-based Approaches*, eds. Aria Adli, Marco García García y Göz Kaufmann, 239-63. Berlin/Boston: De Gruyter.  
<https://doi.org/10.1515/9783110346855-010>.
- Sitaridou, Ioanna. 2011. «Word order and information structure in Old Spanish». *Catalan Journal of Linguistics* 10: 159-84.  
<https://doi.org/10.5565/rev/catjl.36>.
- Sitaridou, Ioanna. 2015. «La anteposición de participio en español antiguo debida a la estructura de la información del discurso». En López Izquierdo/Castillo Lluch, 111-41.
- Smith, Colin. 1972. «Introducción». *Poema de Mio Çid*. Madrid: Cátedra.
- Suárez, Mercedes. 2017. *Sobre el orden de palabras en castellano medieval*. Lugo: Axac.
- Suñer, Avelina. 2015. «Estructura informativa y orden sujeto-participio en las cláusulas absolutas de anterioridad». En López Izquierdo/Castillo Lluch, 319-75.
- Vennemann, Theo. 1974. «Topics, subjects, and word order: from SXV to SVX via TVX». En *Historical linguistics*, eds. John M. Anderson y Charles Jones, 339-76. Amsterdam: North-Holland.
- Warner, Richard. 1980. «Word order in Old Latin: Copulative clauses». *Orbis* 29: 251-63.
- Wolfe, Sam. 2015. «The nature of Old Spanish verb second reconsidered». *Lingua* 164: 132-55. <https://doi.org/10.1016/j.lingua.2015.06.007>.
- Wolfe, Sam. 2016. «Verb-Initial Orders in Old Romance: A Comparative Account». *Revue Roumaine de Linguistique* 60: 147-72.