


## libEscansión: A Recursive Precedence Approach to Metrical Scansion

Fernando Sanz-Lázaro <fernando\_dot\_sanz-lazaro\_at\_oeaw\_dot\_ac\_dot\_at>, Austrian Academy of Sciences  <https://orcid.org/0000-0002-8815-6741>

### Abstract

In recent years, there has been impressive progress in the automatic analysis of Spanish verse. However, despite the excellent results of previous attempts, there is still room for improvement, as this article shows. It presents libEscansión, a new automatic scansion tool that, although initially designed for analysing the verses of plays, outperforms all its predecessors in the scansion of fixed-meter poetry. The library uses two novel approaches to achieve this result: linguistically, it uses Castilian syllabification instead of the more diphthong-prone American syllabic division. It uses a dedicated module that can also be run independently. Methodologically, the library resolves metrical ambiguities using a perceptibility-based recursive prioritisation system. The combination of both allows libEscansión to perform a scansion 0.78% more accurately than the best of its peers against a gold standard corpus, reaching an accuracy of 97.01% (99.50% discarding non-erroneous disagreements).

### Introduction and motivation

Spanish Golden Age plays are composed in verse, and their analysis requires considering their metrical features. A noteworthy property of these works is their polymetry – that is, they combine sections of different metres without a solution of continuity –, although octosyllables are predominant. Nevertheless, there are further challenges beyond the polymetry as the ortho-typography may be unorthodox, intercalations of words, verses, or even whole stanzas in different languages are not uncommon, and, most importantly, the poetic creation serves the speech and not the other way around. Spoken language is, thus, the actual gold standard to which we have to aspire. Unfortunately, working with recordings turns out to be a difficult endeavour, as it raises several practical challenges but, in the first place, involves theoretical issues regarding sample representativity [Sánchez Jiménez 2017]. Therefore, we have opted for a simplified model, which phonology fulfils in the case of speech, even if assuming some general metrical principles.

1

Digital approaches to classical Spanish theatre have experienced a small golden age in recent years, largely thanks to the role played by stylometry in the attribution to Lope de Vega of the obscure play *La francesa Laura* as a result of research by Álvaro Cuéllar and Germán Vega García-Luengos [Cuéllar and Vega García-Luengos 2023a]. This groundbreaking discovery brought worldwide media attention to digital methodologies applied to philological studies. However, such an astonishing finding is not an isolated event but the culmination of the rapid development of Digital Humanities in the literary studies of the Spanish Golden Age theatre. Studies dedicated to this historical genre with digital methodologies have flourished over the last decade, making them a mainstream tool in the field. This has spilt over effects on other areas, timidly opening the door to their adoption, notwithstanding the commendable work of some notable scholars dedicated to other periods and genres [Calvo Tello 2021] of Hispanic literature within the digital humanistic field – some even renowned pioneers of Spanish digital philology [Fradejas Rueda 2010] – and a well-grounded theoretical scaffolding on Digital Humanities in general [Río Riande 2022]. However, in spite of these almost isolated, albeit fruitful, efforts, the convergence of research groups in these other areas has not, in general, reached a critical mass that would give it a momentum comparable to that enjoyed by the theatre of the Golden Age, which has already come to the point where even well-informed self-reflection takes place [Hernández Lorenzo 2020].

2

Recent developments in digital philology focused on Spanish Early Modern theatre should be understood in their context and, more specifically, considering a crucial factor that has prompted such an explosion: the relative availability of well-curated editions. The colossal annotated edition by Presotto [Vega 2015], which sets the quality standards that further attempts should pursue, has found worthy successors in project Prolope [Prolope 2024], DICAT [DICAT 2024] or ISTAE [Istae 2024]. Even though they have more modest pretensions than their predecessor, as the resources are shared among multiple plays, they provide a fixated text with enough enrichments to perform distant analyses of multiple types on it.

3

These projects have paved the way for further endeavours in the field of digital humanities by providing qualitative editions ready to be used as subjects for computer-assisted analysis. Stylometry, together with digital editions, is arguably the spearhead of this advance [Calvo Tello 2016][Cuéllar 2024][Cuéllar and Vega García-Luengos 2023b], which undoubtedly benefits from recent advances in palaeographic text recognition [Rosa, Cuéllar and Lehmann 2024][Cuéllar 2023b], metrical analysis of plays [Kroll and Sanz-Lázaro 2023][Kroll and Sanz-Lázaro 2022][Sanz-Lázaro 2024b], and dramatic analysis [Ehrlicher et al. 2020].

4

Even if in a more modest way, several new approaches have arisen in recent years to facilitate automatic metrical scansion of Spanish verses. The ways they tackle ambiguity resolution are diverse, but, in most cases, the outcomes are fairly positive as they tend to fall into the narrow margin of human inter-annotator agreement. In other words, they perform as well as their human counterparts. However, this also means that bias among human annotators may produce a more tolerant gold standard and increase the margin of the inter-annotator agreement. This seems to happen with the corpus of sonnets ADSO 100 [Navarro-Colorado et al. 2016] used by precedent studies as a gold standard [Marco and Gonzalo 2021]. Therefore, we have some grounds to believe that the 96% inter-annotator agreement Navarro-Colorado [Navarro-Colorado 2017] reported may be an underestimation. Consequently, we considered whether metrical analysis could benefit from another take on the issue, especially regarding the two primary sources of inconsistencies: syllabification and verse ambiguities.

5

The particular issues raised by a play must be adequately addressed to analyse Golden Age theatre as such pieces are intended for the stage and, consequently, an adequate interpretation leaves little place for orthographical idealisations. On the other hand, speech follows phonological rules to solve the interaction between neighbouring sounds. Indeed, one such situation is when vowels interact between word boundaries, a prominent Spanish verse adjustment device. Therefore, phonological rules apply. However, staged Golden Age plays substantially differ from ordinary speech, as they are intended not to be ordinarily spoken but declaimed. Consequently, the most extreme assimilation issues typical of rapid and careless speech in everyday conversation are virtually nonexistent in the genre. Conversely, plays provide additional information that poetry lacks, which allows informed guesses on the length of particular verses.

6

Considering those premises, we have designed a tool for theatre-analysis purposes, which produces state-of-the-art outcomes when tested against fixed metres. We aimed to produce a reusable library to scan Spanish verses that is able to cope with all the shortcomings stated above. That is, the result should be robust enough to handle orthographic and linguistic irregularities, consider phonological exceptions, and solve ambiguities in the best way possible. We have built on previous works, considering that Navarro-Colorado's [Navarro-Colorado 2017] PoS-tag/syllabification model is conceptually sound and the nearest to manual analysis. Posterior developments of the idea [Rosa et al. 2020b] have succeeded by focusing on those areas more prone to be the culprit of scansion errors, setting the right track to follow. The result is three Python libraries for syllabification [Sanz-Lázaro 2021b], phonological transcription [Sanz-Lázaro 2021a], and metrical scansion [Sanz-Lázaro 2023b].

7

## Related work

Studies on automatic metrical scansion of Spanish verse have already yielded significant payoffs. There were promising results already at the very beginning of the 2000s when Gervás [Gervás 2000] conducted a groundbreaking study on the matter. He primarily used a dictionary approach to syllabification, resorting to rules only for those words lacking a matching key in the dictionary. This process provides a list of syllables, the syllable count, prosodic stresses and rhyme. As with any pioneer method, this one presented various shortcomings that must be addressed. The first issue arises due to stress identification, as it only considers orthography. As a result, unstressed syllables such as, for instance, the determiners *la* and *mi* are marked as stressed. Therefore, the first verse of Quevedo's famous sonnet *Amor constante más allá de la muerte*, with which Gervás illustrates its paper, would be parsed as in Example 1 below.

8

*Cerrar podrá mis ojos la postrera*  
*ce rrár po drá mís ójos lá pos tré ra*  
- + - + + + - - + - -

Example 1.

Even though the classification of this particular verse in one of the four major hendecasyllabic types – which is Gervás' approach – would not be affected, a more fine-grained analysis of the verse shows the problem: this short heroic <sup>[1]</sup> verse would be classified as a long one. Furthermore, improper stresses may alter even the primary classification if located on one of the first four syllables. The second issue is related to the sequential way Gervás analyses synaloephas. The resolution is straightforward if the number of supernumerary syllables equals the potential synaloephas. However, if the numbers do not match, parsing the verse from left to right one potential synaloepha at a time may place a synaloepha between the wrong pair of vowels. Moreover, Gervás acknowledges that his programme does not offer a satisfactory method to solve ambiguities derived from tropes such as synaereses and hiatuses. In any case, the author reported an 88.73% accuracy against ADSO 100 [Navarro-Colorado et al. 2016], a corpus of 100 manually annotated sonnets that has become the gold standard for fixed-metre scansion in the last years [Marco et al. 2021].

9

Navarro-Colorado [Navarro-Colorado 2015] noted those shortcomings, which led him to develop a more sophisticated method. In order to improve stress identification, the words are first PoS-tagged – that is, they are annotated with information regarding their grammatical category – to decide whether they are stressed. Roughly described, classes such as, for instance, verbs or nouns carry prosodic stress in Spanish, while prepositions or conjunctions do not, , regardless of the word's intrinsic lexical stress [Quilis 2013]. The next step is syllabifying and placing prosodic stresses only in those words with the appropriate PoS-tag. If the syllable count is eleven, a hendecasyllabic verse is solved. Otherwise, the programme launches an adjusting module. If it produces multiple potential results, it calls a disambiguation module to consider all possible alternatives in the corpus. This approach solves Gervás' stress placement problem and substantially improves ambiguity resolution, which raises the accuracy to 94.44% [Marco et al. 2021] against the same test corpus.

10

Some years later, Agirrezabal, Alegría, and Hulden [Agirrezabal, Alegría and Hulden 2017] approached the issue by employing neural networks, with a reported accuracy of 90.84%. However, the best results have not been obtained until the current decade. More concretely, Rantanplan [Rosa et al. 2020b] reaches 96.22% against ADSO 100 [Marco et al. 2021] following Navarro-Colorado's classic approach, making this last project the most accurate up to the present date. Furthermore, it reduces the processing time from 40 min to 3s. Almost simultaneously, another team at the same university developed a new approach, Jumper [Marco and Gonzalo 2021], which neither syllabifies the words nor relies on PoS-tagging. These innovations improve the scansion speed even more with minimal accuracy loss, reaching an astounding 94.97% in 0.31s analysing Navarro-Colorado's corpus.

11

## Verse analysis

As previous methods do, we have divided metrical scansion into two stages: naive analysis and syllabic adjustment, but we approach the latter in a different fashion. Firstly, we perform the syllabification and PoS-tagging of the words. Then, we try to adjust the verse if it does not match the expected syllable count. Our library introduces two novelties – apart from the phonological transcription, which is for convenience rather than conceptual – namely, full exceptions-handling in the syllabification and its approach to ambiguity resolution (see Figure 1).

12

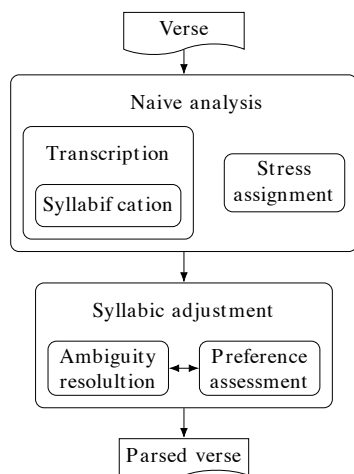


Figure 1.

## Naive analysis

The first step is a word-wise analysis of the verse by the libEscansión class `PlayLine`, whose practical usage is explained below. It consists of dividing the words into syllables and determining whether a particular word carries prosodic stress. The former task is carried out by the library `Fonemas` [Sanz-Lázaro 2021a], whilst the latter resorts to `Stanza` [Qi et al. 2020] PoS-tagger to determine the grammatical category and whether it is stressed. This class has two public attributes: a string containing the verse and a list of its words, each of whose elements is, in turn, another list with the corresponding word syllables, which are marked if they carry metric stress.

## Syllabification

Even though a phonological transcription is not strictly necessary to analyse a verse, it is convenient for diverse reasons. Firstly, using IPA notation with a single Unicode character for each phoneme expedites the straightforward handling of syllabic division, as we can treat each character equally without exceptions<sup>[2]</sup>. Spanish digraphs (*ch*, *rr*, *ll*, *qu*, *gu*) or multiple phonemes represented by a single grapheme (e.g., *x*) are no longer required. Second, the output allows an out-of-the-box rhyme analysis as the relation phoneme-character is biunivocal, which spares a special treatment for ambiguous graphemes (*b*, *v*, *c*, *z*, *g*). In this step, the text is not only transcribed but also preprocessed. Stranded letters are transliterated to their spelling (e.g., *B* would become *be*) as well as non-Spanish diacritics (e.g., Catalan *què* would become *ké*); other potential sources of errors are also handled here (e.g., typographic ligatures). Finally, it allows a straightforward treatment of vowel clusters, as syllabic and non-syllabic vocalic allophones do not share a common representation.

The syllabification proper is carried out by the independent library `Silabeador`, called by the transcriber. First, the programme looks for non-transliterated Latin words and tries to transliterate them according to the Spanish orthographic rules. Contrary to other foreign words, Latin words are not phonologically adapted and tend to respect not only stress positions contradicting Spanish rules but also the original syllabic division, which may also differ from Spanish due to the disparate sets of diphthongs of both languages. Then, the programme marks forced syllabic divisions where an exception should occur according to an external file. This file consists of two columns with *regular expression* substitutions; the first field contains a matching pattern, and the second corresponds to its substitution, including an underscore where a syllabic division must be forced. These phonological exceptions include ubiquitous words whose Castilian pronunciation differs from the orthographic hyphenation. To name some extremely common in any kind of speech, this rule matches the hiatuses of verbs in *-iar*, *-uar*, *-uir* [Quilis 2019], *-eir*, and *-oir* [Ríos 1999], adjectives in *-uoso* [Quilis 2019], some particular nouns [Navarro 1957], and other words [Canellada and Madsen 1987]. Adjectives in *-mente* are also exceptional since they carry secondary stress [Quilis 2013], which Spanish phonology otherwise lacks. They are marked too in the morpheme, and the syllabification module receives the lexeme alone; finally, both parts are rejoined after their separate treatment.

There are ambiguous words, such as *juez*, *suave* or *fiel* [Dominguez 2012], in which the diphthong is often broken in Spanish Golden Age verses. They might be typographically marked with a diaeresis (*í* or *ü*) by the editor of the source, in which case they are parsed as usual, but this may not be the case. Therefore, these ambiguous clusters are marked as a diaeresis, although their sinaeresis will be given priority later.

The resulting text is divided in its syllabic nuclei, arranging the diphthongs except for those explicitly separated in the previous step. Non-syllabic vowels – the non-nuclear sounds in a vocalic glide – are marked as such by substituting them for the character representing the corresponding allophone. The outcome is a list of syllabic nuclei and free consonants. Finally, those free consonants are assigned to the proper syllable on coda or onset according to Spanish phonological rules [Quilis 2019]. The output is a list of phonological transcribed syllables and the prosodic stress on the due ones.

## Stress assignment

Word stress does not correspond to metric stress exactly<sup>[3]</sup> as some word classes are considered unstressed regardless of their pronunciation in isolation [Quilis 2013]. The class `VerseMetre` of libEscansión determines which words carry metric stress. It takes a verse as an argument, cleans it, and divides its words into syllables using the library `Fonemas`. It also uses `Stanza` [Qi et al. 2020] to PoS-tag each word. After this, it decides whether a particular word carries stress according to its PoS-tag, but also considering the relation to neighbouring words when forming a compound as it may deprive from their prosodic stress to otherwise stressed words (e.g. *Santo Tomás* /san.to.to'mas/ but *el santo padre* /el'san.to'pa.dre/ or *cuarenta* /kwa'ren.ta/ but *cuarenta y dos* /kwa.ren.tai'dos/), in a similar fashion as it occurs, for example, with English numerals.

PoS-tagging may produce consistent errors or not be detailed enough. It can be bypassed in such situations by marking the words according to lists of stressed or unstressed occurrences. Undetected errors, however, are still possible. The greatest difficulty is detecting when an unstressed word should not carry prosodic stress as it involves several open word classes, whilst unstressed words belong to close ones. Therefore, a relatively high proportion of wrongly tagged words has no impact on the outcome as long as the tag represents an equivalent class in terms of stress.

## Syllabic adjustment

The syllabic adjustment resorts to libEscansión class `VerseMetre`, a child of `PlayLine`. This class takes a sorted list of possible numbers of syllables apart from the verse as an argument. For a play, a reasonable value for the list would be at least [8, 11, 7, 6] as most verses are octosyllabic, but hendecasyllables or these mixed with heptasyllables, hexasyllables for sung verses, and even more exotic combinations may appear. On the other hand, regular sonnets would encounter no problem with a list with one element [11].

20

The syllables are then parsed considering word boundaries by the private method `__adjust_metre()`, which takes a list of words and a sorted list of potential metres. Firstly, it tests whether the syllabic count plus the rhyme-stress correction (in other words, adding a syllable if the rhyming word is oxytonic or subtracting one if it is proparoxytonic) matches the first element of the list, which would be a trivial solution. Otherwise, it tries to adjust the metre.

21

The function has subroutines to search for potential synaloephas, hiatuses, and epentheses and apply them when necessary. If the verse length equals the number of orthographic syllables minus the number of potential synaloephas, it applies the synaloephas to the list of words. If this is not the case, it tries to do the same, but removing potential epentheses this time. In the opposite situation, if the orthographic syllables count is lower than the expected number, the function tries to apply hiatuses in suitable positions.

22

Finally, the function checks whether the corrected length matches the expected one more time. If this is the case, it returns the list of words flattened to a list of syllables. Otherwise, it calls itself recursively passing the list of expected lengths after removing its first element. Consequently, in the next recursion level, the method repeats the same process but regarding the second element of the original list as the expected length.

23

Syllabic splitting and joining are done with equivalent procedures. Placing synaloephas and synaereses or hiatuses and diaereses follows the same steps. First, a function locates all potential instances of the trope in question and produces a list of potential positions. Second, another function assesses and ranks those instances according to a system of preferences. Finally, another function applies the figure according to each position's rank until it matches the expected verse length. The three functions are invoked sequentially by a recursive function. With the sorted list of potential adjustments, the function applies the first one and invokes itself for the new list of words.

24

Let us illustrate it with the synaloephas and synaereses as the process involves additional steps. Finding potential syllabic mergers involves finding pairs of vowels in syllabic boundaries, providing they are compatible according to the perceptibility scale [Navarro 1957]. That means finding those combinations in which the perceptibility decreases in both directions of the nucleus.

25

| w | j | e | o | a |   | á | ó | é | í | ú |
|---|---|---|---|---|---|---|---|---|---|---|
| x | x | x | x | - | a | - | x | x | x | x |
| x | x | x | - | o |   | - | x | x | x |   |
| x | x | - |   | e |   |   | - | x | x |   |
| x | - |   |   | i |   |   |   | - | x |   |
| x | - |   |   | u |   |   |   |   | - |   |

Table 1. Perceptibility scale

That validates the transcription Example 2a and Example 2b, as its long vocalic is built around a nucleus *o* with higher perceptibility than the non-syllabic semiconsonant at its side in Example 2a or semivowel in Example 2b. Conversely, Example 2c and Example 2d contradict the Spanish phonology, as *o* is the vowel of highest perceptibility in this example and, therefore, cannot follow a less perceptible *u*. In the best case, three words can join in a five-vowel synaloepha, such as *envidia a Eustaquio* (*/j̞oæ̞e̞u/*).

26

*uno u otro*

Example 2a. /'u.no 'wo.tro/

Example 2b. /'u.nou 'o.tro/

Example 2c. /'u 'nouo.tro/

Example 2d. /'u 'ngwo.tro/

Example 2.

With the list of potential synaloephas, each element is evaluated and given points according to the prosodic circumstances colluding [Canellada and Madsen 1987] and their relative position according to the perceptibility scale and the Spanish preference for crescent diphthongs [Alarcos 1964]. Potential synaloephas between hemistichs have a penalisation – exceptional cases such as the conjunctions *y* and *o* need to be addressed individually. Additionally, a correction to potentially aspirated *h* is added. The latter responds to the poets' custom of making it aspirated (*/h/*) when in need of breaking a synaloepha, despite having become mute in the regular speech a long time ago and being kept in the orthography on pure etymological grounds [Lapesa 2008]. Once the inter-word boundaries are parsed, the same is done with the in-word syllables, not without adding an offset to prioritise synaloephas over synaereses. The list of positions is sorted according to their scores.

27

Finally, another function rearranges the list of syllables by joining words where required. This step involves not only moving phonemes and joining syllables but also marking non-syllabic vowels as such in the newly created clusters.

28

Diaereses are solved in a similar fashion. Firstly, the programme composes a list of potential divisions. Then, it assigns priorities and applies them recursively until meeting the expected count, or it tries with the next length value of the list if no suitable arrangement can be found.

29

## Mixed verse scansion

Using a sorted list of expected verse lengths confers great flexibility to libEscansión. For example, our approach to parsing plays consists of creating an instance of `VerseMetre`, providing an initial list with typical lengths such as [8, 11, 7, 6, 5, 4]. The verses usually are octosyllabic, so we put 8 in the first position. However, if a hendecasyllabic appears, the new instance of `VerseMetre` for the next verse will be given a list with 11 on top. We also know that poetic forms composed of hendecasyllabic and heptasyllabic verses are common. Therefore, we relegate 8 to the third position. Furthermore, as we work with parsed plays, we can use additional data to infer the length. For example, if the speaker is a musician or plainly `MUSIC`, it would probably be a popular form in hexasyllabic verses rather than a more prescriptive one; thus, we could parse that particular verse directly with 6 on top.

For example, the first *estrambote* verse of Quevedo's "Pintando la vida de un señor mal ocupado" has seven syllables instead of eleven (*Ca.ba.lle.ros chan.flo.nes*). We pass [11, 7, 8, 6] to the disambiguation module when analysing sonnets. After confirming that there are no possible divisions to obtain eleven syllables, the module calls itself removing the first element of the list of lengths and passing [7, 8, 6] to the next recursion level. As 7 is indeed possible, it returns this resolution.

In addition to attributes for syllables and rhythm, the class `VerseMetre` provides others for full and partial rhymes, syllabic nuclei, and ambiguity. The last one takes an integer value and can be used for a second pass to improve mixed-verse results.

## Evaluation

### Syllabification

We have tested the syllabification accuracy against the 106,362 words corpus EDFU [Rosa and Pérez 2020]. Using an *American* configuration as the corpus does, that is, without non-orthographic hiatuses and the group *tl* as an indivisible onset (*trans-a-tlán-tico* instead of *trans-at-lán-ti-co*), the syllabification module achieves 99.92% accuracy. This improves Agirrezabal's [Agirrezabal, Alegría and Hulden 2017] 98.06% and Navarro-Colorado's [Navarro-Colorado 2015] 98.35%. However, at first sight, it lies behind 99.99% of Rantanplan [Rosa et al. 2020a]. That is, disregarding errors like improper diphthongs in the corpus, such as *amambaiense a-mam-baien-se* instead of *a-mam-ba-ien-se*. All in all, 86 words differ from the corpus. From those, after a visual inspection, 37 differences are wrongly syllabified in the corpus, 41 have alternative yet acceptable differing syllabifications, and our programme wrongly syllabifies eight words Table 2.

|   | Word             |
|---|------------------|
| 1 | alz-hei-mer      |
| 2 | alz-héi-mer      |
| 3 | em-men-thal      |
| 4 | in-ter-vi-ú      |
| 5 | o-ui-ja          |
| 6 | pen-thou-se      |
| 7 | s-pas-pag-het-ti |
| 8 | s-pin-ning       |

Table 2. Syllabification errors

The words *penthouse* and *ouija* do not obey the Spanish rules and are syllabified by approximating the original phonology, which would require to treat them as an exception. Since this would not be productive, we have opted to accept this kind of error. When considering only these eight proper errors, the accuracy rises to 99.99%.

### Metrical scansion

To test the accuracy<sup>[4]</sup>, we downloaded the corpus ADSO 100 [Navarro-Colorado 2017] using the tool Averell [Díaz et al. 2021] to compare our results with previous studies, which is the gold standard they used. ADSO 100 is the only manually annotated corpus of its size with a reported inter-annotator agreement. It comprises 1,404 verses.<sup>[5]</sup> Our hardware was an Intel® Core™ i5-3470 CPU @ 3.20GHz, 16 GB of DDR4 RAM and a GPU NVIDIA GeForce GTX 1060 with 6GB. The PoS-tagging module ran on the GPU. For the syllabification, the options for Castilian prosody have been enabled, as well as epenthesis. The analysis takes 87 seconds ( $\pm 1.5$  s) in all cases (running Stanza on the GPU), disregarding the approximately 7 seconds Stanza takes to load.

In short, libEscansión scores 97.01% accuracy against the corpus when running with its standard settings (*Castilian* syllabification), an improvement over previous scansion engines (Table 3). The *American* syllabification renders slightly worse results, with 96.51%. However, as noted above, this corpus considers interjections such as *oh* and *ay* as unstressed, whilst our library does not. This behaviour can be reproduced with an ad hoc three-line hack in the code, in which case the accuracy would reach 98.65% for the standard setting and 98.15% for the *American* syllabification. We have considered inconsistencies in the corpus notation for the test as some proparoxytonic verses of the corpus are marked by omitting one of the last two unstressed syllables.

| Method              | Accuracy |
|---------------------|----------|
| Gervás              | 70.88    |
| Agirrezabal         | 90.84    |
| Navarro-Colorado    | 94.45    |
| Jumper              | 94.97    |
| Rantanplan          | 96.23    |
| libEscansión (ours) | 97.01    |

Table 3. Scores on Navarro-Colorado's 100 sonnets corpus as reported by Marco et al. [Marco et al. 2021]. Best scores in bold.

Apart from the disagreements due to interjections, there are 19 disparate verses left for other reasons and another with disagreements beyond the interjections

(Table 4, row 18). There are ten disagreements due to spelling deviations (rows 1–5) and annotator mistakes (rows 6–12). In the first case, the manual annotation reflects the pronunciation despite the lack of accents, such as in words like *Priamo* and the more subtle diacritics of *aún* and *dó*. Consequently, humans have an advantage over machines as the former can recognise the proper pronunciation despite the spelling. The second case ranges from non-marked stresses in content words, disregarding the primary stress of an adverb in *-mente* or an adjective, or the opposite, stressing an unstressed syllable like in an honorary title preceding a noun or the locution *ya que*.

38

Two disagreements (rows 18–19) deserve a special mention. The first is due to a PoS-tagging mistake, as *para* has been tagged as a preposition due to the lack of a vocative comma, which illustrates what we said above. Even though other potential errors of a similar kind have been intercepted, ambiguities resembling row 18 cannot be bypassed since the PoS-tagger has operated as expected, and it is unviable to predict any potential typo or edition mistake. The second mistake (row 19) is rather interesting. The syllabicator breaks the diphthong, but *VerseMetre*, after pondering all the options, joins the vowels again as it prioritises placing the caesura between words when keeping editorial diacreses is not explicitly forced.

|    | Verse  |
|----|--|
| 1  | falte a <b>Priamo</b> tierra y falte fuego.        |
| 2  | Si rico, ¿ <b>do</b> tus bienes vinculados?        |
| 3  | <b>aun</b> no pudo, de lástima, dar muerte,        |
| 4  | <b>aun</b> le quedaba brío y lozanía.              |
| 5  | del muerto Lilio es; que <b>aun</b> no perdona     |
| 6  | quien <b>promete</b> buen fin a mi viaje,          |
| 7  | trocó mi mal en <b>bien</b> , mi pena en gloria.   |
| 8  | este que, siendo <b>solamente</b> cero,            |
| 9  | del <b>señor</b> Amadís, como tú hiciste           |
| 10 | <b>ya que</b> no se conserva, se previene          |
| 11 | ¡Oh <b>santo</b> bodegón! ¡Oh picardía!            |
| 12 | Llámasme, <i>gran</i> Señor; nunca respondo.       |
| 13 | <i>confiada en</i> la vista vencedora,             |
| 14 | <b>Maestro</b> era de esgrima Campuzano,           |
| 15 | Yo <b>vi</b> unos ojos bellos, <i>que hirieron</i> |
| 16 | Quien <i>no, huya y no</i> escuche mi lamento,     |
| 17 | Lo caduco <i>esta uma</i> peregrina,               |
| 18 | <i>Para</i> cochero. El coche está en palacio.     |
| 19 | amenazando una total <i>rúina</i>                  |

Table 4. Differing annotations. In 13-17, corpus in bold, LibScansión in italics

If we leave the remaining inter-annotator disagreements, be they errors or different yet valid takes on the verse, libEscansión reaches a 99.50% accuracy.

39

## Requirements, installation and use

libEscansión was intended to be an internal part of a more complex mechanism. As the aim was to parse hundreds of thousands of verses from various sources automatically [Sanz-Lázaro 2023c], the verse was the last step after digging into *œuvre*, play, and act. Therefore, it was conceived as a Python library to be called by the programmes in charge of working at a higher level [Sanz-Lázaro 2023a]. Nonetheless, it can be used as a standalone tool even if lacking knowledge of Python or programming, as long as a Python 3.9 or later distribution is installed on the computer. It has been released under the Free Software GNU LGPL licence, making it available for embedding in other software without requiring an equivalent licence as long as the library's code remains unmodified; otherwise, Free Software licencing would be required.

40

As libEscansión relies on several libraries, these must also be installed (see Example 3).

41

```
pip install silabeador fonemas stanza==1.7.0 libEscansion
```

Example 3.

In order to carry the PoS-tagging properly, language models for Spanish should be available. In our case, we have used Stanza for the task, which provides functions to download them. In short, after importing the library within the Python environment should be executed (see Example 4).

42

```
import stanza
stanza.download(lang="es",
                package=None,
                processors={"ner": "ancora",
                           "tokenize": "ancora",
                           "pos": "ancora",
                           "lemma": "ancora",
                           "mwt": "ancora",
                           "constituency": "combined_charlm",
                           "depparse": "ancora",
                           "sentiment": "tass2020"})
```

Example 4.

To deploy libEscansión, the library must first be imported into the Python environment. Then, an instance of the class `VerseMetre` can be created, passing the verse to scan as a parameter. For convenience, the class `VerseMetre` can be imported independently. The resulting object possesses a number of attributes to represent verse features. Even though IPA notation has been preserved to a great extent – and thus, Unicode support is required for optimal visualisation – the main concern was automating tasks. To simplify text processing, stressed syllables have upper-case letters as their nuclear vowel instead of the IPA stress symbol. A typical workflow is represented in Figure 2.

43

```
In [1]: from libEscansion import VerseMetre

In [2]: text = "Érase un hombre a una nariz pegado"

In [3]: verse = VerseMetre(text)

In [4]: print(f"Syllable count:\t\t{verse.count}\n\
Syllabic nuclei:\t\t{verse.nuclei}\n\
Transcription:\t\t{verse.syllables}\n\
Consonance:\t\t\t{verse.rhyme}\n\
Assonance:\t\t\t{verse.asson}\n\
Rhythm:\t\t\t\t\t{verse.rhythm}")

Syllable count:      11
Syllabic nuclei:    EaU0AaaIeAo
Transcription:      ['E', 'ra', 'səUn'], ['h0m', 'brəAw', 'na'], ['na', 'rI0'], ['pe', 'gA', 'do']]
Consonance:         ado
Assonance:          ao
Rhythm:             +-+++--+--
```

Figure 2.

## Limitations

The library libEscansión has been designed with Spanish Golden Age theatrical verses in mind. Consequently, it has been modelled according to the usual metres in such pieces. The aforementioned manually annotated corpus of sonnets was the only one providing inter-annotator agreement, but it fulfilled the project's purposes. That means that no tests have been carried out with other types of verse, such as Alexandrines, regardless of their popularity and usage prior to and after the period, and further tests are required to measure libEscansión's accuracy beyond the Spanish Golden Age theatre versifying customs.

44

Furthermore, the tool was implemented as a means to build a data frame rather than a standalone tool. Therefore, speed has not been considered, and the processing time reflects it. Even though the hardware substantially differs from that used by Marco et al. [Marco et al. 2021] to carry out their tests, the processing time is still two orders of magnitude higher than the fastest method's score.

45

A weak spot is a dependency on a reliable PoS-tagger to mark prosodic stresses, which, in turn, depends on a language model. Consequently, erroneous tags have to be intercepted to bypass the PoS-tagging. This is not always possible, as some errors are unpredictable, and creating an exception for others would provoke side effects.

46

## Conclusions and further work

We have presented a new library for analysing Spanish verses that surpasses other automatic scansion tools for Spanish fixed-metre verses. The contribution of this work is twofold, as it involves different linguistic choices and a novel approach to ambiguity resolution. On the one hand, we have opted for Castilian syllabification rather than the more diphthong-prone American (and orthographical) varieties. According to the results, this may be a sensible choice as it outperforms the latter under identical conditions. On the other hand, we have presented a perceptibility-based recursive prioritisation system, which, combined with the Peninsular syllabification, raises the score of libEscansión well above its peers.

47

Additionally, working with Unicode-based phonological transcriptions removes layers of abstraction and the necessity of mapping correspondences to go back and forth from written to spoken text. This dramatically simplifies the handling of vocalic clusters since it explicitly discriminates between syllabic and non-syllabic

48

vowels.

The high degree of modularisation of the tools, with separate libraries for syllabification, phonological transcription and scansion, extends the range of uses beyond metrical analysis, as the lower layers can be used independently for different purposes. Syllabification and phonological transcription undoubtedly have uses beyond the narrow spectrum of historical theatre, in fields such as linguistics and language teaching.

49

With regard to poetry, which is the main focus of this work, there are many examples of the library being used in a real-life scenario. It has been developed primarily to provide a corpus for remote reading analyses, taking into account metric features. The corpus annotated with a previous version of libEscansión is available under a Creative Commons BY licence as an SQL database [Kroll and Sanz-Lázaro 2024], which is also hosted online and available for external queries through a graphical interface [Sound and Meaning 2022].

50

There is plenty of room for further improvement. However, comprehensive testing requires manually annotated corpora. Hopefully, in the near future, manually annotated corpora of polymetric pieces and other poetic forms besides the sonnet used here will be prepared. This would allow for further testing, which would undoubtedly reveal shortcomings and help improve libEscansión.

51

Although libEscansión was conceived as an auxiliary tool, it has transcended its original purpose. However, its interface limits it to a specialised niche with some technical knowledge. A graphical interface could help a broad public to access the capabilities of libEscansión without coding. As for users familiar with programming, thanks to the Free Software licence, libEscansión can be integrated into new developments – including third-party frontends – and its code can be adapted for new purposes.

52

## libEscansión Resources

The library libEscansión, documentation, and associated materials for installation and use may be found at <https://pypi.org/project/libEscansion/> and on GitHub at <https://github.com/fsanzl/libEscansion>.

53

## Funding

This work is part of the project *Sound and Meaning in Spanish Golden Age Literature*, funded by the Austrian Science Fund (FWF P 32563). <https://doi.org/10.55776/P32563>

54

This work was supported by the project *The Interpretation of Childbirth in Early Modern Spain*, funded by the Austrian Science Fund (FWF P 32263). <https://doi.org/10.55776/P32263>

55

## Notes

[1] We use Varela Merino, Moíno Sánchez and Jauralde Pou [Varela Merino, Moíno Sánchez and Jauralde Pou] systematics here for convenience. The library libEscansión aims to describe the features of the verse and considers the criteria used by Varela Merino, Moíno Sánchez, and Jauralde Pou as well as Navarro Tomás [Navarro Tomás 2015] in their respective taxonomies. It provides enough elements of judgment to classify a verse according to both systems but does not perform the classification itself.

[2] A SAMPA transcription, for instance, would share the disadvantages of the orthographic text in this regard.

[3] To use a simile with similar effects in verse - with all their differences - think of the full and reduced forms of function words in English.

[4] The script used is available in libEscansión development Git repository [Sanz-Lázaro 2024] in the subdirectory `utils`.

[5] The reason of this figure rather than 1,400 is Quevedo's inclusion of four additional verses, a *estrambote*, in "Pintando la vida de un señor mal ocupado".

## Works Cited

- Agirrezabal, Alegría and Hulden 2017** Agirrezabal, Manex, Iñaki Alegría and Mans Hulden. "A comparison of feature-based and neural scansion of poetry". In *RANLP 2017 – Recent Advances in Natural Language Processing*, 18–23. 2017. [https://doi.org/10.26615/978-954-452-049-6\\_003](https://doi.org/10.26615/978-954-452-049-6_003).
- Alarcos 1964** Alarcos Llorach, Emilio. *Fonología española*. Madrid: Gredos, 1964.
- Calvo Tello 2016** Calvo Tello, José. "Entendiendo Delta desde las humanidades". *Caracteres. Estudios culturales y críticos de la esfera digital* vol. 5.1 (2016): 140–176. <https://doi.org/10.5944/signa.vol30.2021.29299>
- Calvo Tello 2021** Calvo Tello, José. "Corpus de novelas de la Edad de Plata, en XML-TEI". *Signa: Revista de la Asociación Española de Semiótica* vol. 30 (2021): 83–107. <https://doi.org/10.5944/signa.vol30.2021.29299>
- Canellada and Madsen 1987** Canellada, María Josefa and John K. Madsen. *Pronunciación del español. Lengua hablada y literaria*. Barcelona: Castalia, 1987.
- Cuéllar 2023a** Cuéllar, Álvaro. "Cronología y estilometría: datación automática de comedias de Lope de Vega". *Anuario Lope de Vega. Texto, literatura, cultura*, vol. 29 (2023): 97-130. <https://doi.org/10.5565/rev/anuariolopedevega.483>.
- Cuéllar 2023b** Cuéllar, Álvaro. "La Inteligencia Artificial al rescate del Siglo de Oro. Transcripción y modernización automática de mil trescientos impresos y manuscritos teatrales". *Hipogrifo. Revista de literatura y cultura del Siglo de Oro*, vol. 11.1 (2023): 101–115. <https://doi.org/10.13035/H.2023.11.01.08>.
- Cuéllar 2024** Cuéllar, Álvaro. "Stylometry and Spanish Golden Age theatre: An evaluation of authorship attribution in a control group of one hundred undisputed plays". In *Digital stylistics in Romance Studies and beyond*, edited by Robert Hesselbach, José Calvo Tello, Ulrike Henny-Krahmer, Christof Schöch and Daniel Daniel. Heidelberg: Heidelberg University Press, 2024. [https://www.academia.edu/53262871/Stylometry\\_and\\_Spanish\\_Golden\\_Age\\_Theatre\\_An\\_Evaluation\\_of\\_Authorship\\_Attribution\\_in\\_a\\_Control\\_Group\\_of\\_Undisputed\\_Plays](https://www.academia.edu/53262871/Stylometry_and_Spanish_Golden_Age_Theatre_An_Evaluation_of_Authorship_Attribution_in_a_Control_Group_of_Undisputed_Plays). [preprint]. Accessed 09 February 2024.
- Cuéllar and Vega García-Luengos 2023a** Cuéllar, Álvaro and Germán Vega García-Luengos. "*La francesa Laura. El hallazgo de una nueva comedia del Lope de Vega*". *Anuario Lope de Vega*, vol. 29 (2023): 131–198. <https://doi.org/10.5565/rev/anuariolopedevega.492>.
- Cuéllar and Vega García-Luengos 2023b** Cuéllar, Álvaro and Germán Vega García-Luengos. "Un nuevo repertorio dramático para Andrés de Claramonte". *Hipogrifo. Revista de literatura y cultura del Siglo de Oro*, vol. 11.1 (2023): 117–172. <https://doi.org/10.13035/H.2023.11.01.09>.
- DICAT 2024** DICAT. *DICAT. Grupo de investigación teatral*. <https://dicat.uv.es/elgrupodicat>



- Dominguez 2012** Domínguez Caparrós, José. "Prosodia rítmica de vocales contiguas en interior de palabra". *Rhythmica. Revista Española de Métrica Comparada* vol. 10 (2012): 11–44. <https://doi.org/10.5944/rhythmica.13067>
- Díaz et al. 2021** Díaz Medina, Aitor, Álvaro Pérez Pozo and Javier de la Rosa. *Averell: A corpus management tool to transform poetic corpora into a JSON format compliant with the POSTDATA ontology, 2021. v1.2.2* <https://doi.org/10.5281/zenodo.5702404>
- Ehrlicher et al. 2020** Ehrlicher Hanno, Jörg Lehmann, Nils Reiter and Marcus Willand. "La poética dramática desde una perspectiva cuantitativa: la obra de Calderón de la Barca". *Revista de Humanidades Digitales* vol. 5 (2020): 1–25 <https://doi.org/10.5944/rhd.vol.5.2020.27716>
- Fradejas Rueda 2010** Fradejas Rueda, José Manuel. "La codificación XML/TEI de textos medievales". *Memorabilia* vol. 12 (2010): 219–247. <http://dx.doi.org/10.17613/M6DH35>
- Fradejas Rueda 2016** Fradejas Rueda, José Manuel. "El análisis estilométrico aplicado a la literatura española: las novales policíacas e históricas". *Caracteres. Estudios culturales y críticos de la esfera digital* vol. 5.2 (2016): 196–245. <http://dx.doi.org/10.17613/M6J66F>
- Gervás 2000** Gervás, Pablo. "A logic programming application for the analysis of Spanish verse". In *Computational Logic - CL 2000, First International Conference*, 1330–1344. Berlin: Springer, 2000. [https://doi.org/10.1007/3-540-44957-4\\_89](https://doi.org/10.1007/3-540-44957-4_89)
- Hernández Lorenzo 2020** Hernández Lorenzo, Laura. "Humanidades Digitales y Literatura española: 50 años de repaso histórico y panorámica de proyectos representativos". *Janus: estudios sobre el Siglo de Oro* vol. 9 (2020): 562–595. <https://www.janusdigital.es/articulo.htm?id=154>
- Istae 2024** ISTAE. *ISTAE. Impresos sueltos del teatro antiguo español*. <https://istae.uv.es/>
- Kroll and Sanz-Lázaro 2022** Kroll, Simon and Fernando Sanz-Lázaro. "Romances teatrales entre Mira de Amescua, Calderón y Lope: ritmo, asonancia y cuestiones de autoría". *Revista de Humanidades Digitales* vol. 7.1 (2022): 1–18. <https://doi.org/10.5944/rhd.vol.7.2022.31620>
- Kroll and Sanz-Lázaro 2023** Kroll, Simon and Fernando Sanz-Lázaro. "Ritmo, autoría y género: nuevas perspectivas sobre teatro lopesco desde las humanidades digitales". *Anuario Lope de Vega* vol. 29 (2023): 351–375. <https://doi.org/10.5565/rev/anuariolopedevega.491>
- Kroll and Sanz-Lázaro 2024** Kroll, Simon and Fernando Sanz-Lázaro. *Versos y estructuras teatrales áureos. A corpus of Spanish plays with dramatic and metric information*. 2024 <https://doi.org/10.25365/phaidra.471>
- Lapesa 2008** Lapesa, Rafael. *Historia de la lengua española*. Madrid: Gredos, 2008.
- Marco and Gonzalo 2021** Marco, Guillermo and Julio Gonzalo. "Escansión automática de poesía española sin silabación". *Procesamiento del Lenguaje Natural* vol. 66 (2021): 77–87. <https://doi.org/10.26342/2021-66-6>
- Marco et al. 2021** Marco, Guillermo, Javier de la Rosa, Julio Gonzalo, Salvador Ros, and Elena González-Blanco. "Automated Metric Analysis of Spanish Poetry: Two Complementary Approaches". *IEEE Access* vol. 9 (2021): 51734–51746. <https://doi.org/10.1109/ACCESS.2021.3069635>
- Navarro 1957** Navarro Tomás, Tomás. *Manual de pronunciación española*. Madrid: Consejo Superior de Investigaciones Científicas, 1957.
- Navarro Tomás 2015** Navarro Tomás, Tomás. "El octosílabo y sus modalidades". In *Los poetas en sus versos: desde Jorge Manrique a García Lorca*, 35–66. Madrid: Centro para la edición de los clásicos españoles, 2014.
- Navarro-Colorado 2015** Navarro-Colorado, Borja. "A computational linguistic approach to Spanish Golden Age Sonnets: Metrical and semantic aspects". In *Proceedings of NAACL-HLT Fourth Workshop on Computational Linguistics for Literature*, 105–13. Denver, CO: Association for Computational Linguistics, 2015. <https://doi.org/10.3115/v1/W15-0712>
- Navarro-Colorado 2017** Navarro-Colorado, Borja. "A metrical scansion system for fixed-metre Spanish poetry". *Digital Scholarship in the Humanities* vol. 33.1 (2017): 112–127. <https://doi.org/10.1093/llc/fqx009>
- Navarro-Colorado et al. 2016** Navarro-Colorado, Borja, María Ribes, and Noelia Sánchez. "Metrical annotation of a large corpus of Spanish sonnets: representation, scansion and evaluation". *10th edition of the Language Resources and Evaluation Conference 2016 Portorož, Slovenia*, 2016. [http://www.dlsi.ua.es/~borja/navarro2016\\_MetricalPatternsBank.pdf](http://www.dlsi.ua.es/~borja/navarro2016_MetricalPatternsBank.pdf)
- Prolope 2024** Prolope. *Prolope. Grupo de investigación sobre Lope de Vega de la Universidad Autónoma de Barcelona*. 2024. <https://prolope.uab.cat/>
- Qi et al. 2020** Qi, Peng, Yuhao Zhang, Yuhui Zhang, Jason Bolton, and Christopher D. Manning. "Stanza: A Python Natural Language Processing Toolkit for Many Human Languages". In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. 2020. <https://nlp.stanford.edu/pubs/qi2020stanza.pdf>
- Quilis 2013** Quilis, Antonio. *Métrica españolas*. Barcelona: Ariel, 2013.
- Quilis 2019** Quilis, Antonio. *Tratado de fonología y fonética españolas*. Madrid: Gredos, 2019.
- Rosa and Pérez 2020** Rosa, Javier de la and Álvaro Pérez. *linhd-postdata/edfu: First version of EDFU syllabification corpus*. 2020. <https://zenodo.org/record/3898684>
- Rosa et al. 2020a** Rosa, Javier de la, Álvaro Pérez, Laura Hernández, Salvador Ros, Elena González-Blanco. *PoetryLab: An open source toolkit for the analysis of Spanish poetry corpora*. 2020. <https://postdata.linhd.uned.es/resultats/poetrylab/poetrylab-app/>
- Rosa et al. 2020b** Rosa, Javier de la, Álvaro Pérez, Laura Hernández, Salvador Ros, and Elena González-Blanco. "Rantanplan, fast and accurate syllabification and scansion of Spanish poetry". *Procesamiento del Lenguaje Natural* vol. 65 (2020): 83–90. <https://doi.org/10.26342/2020-65-10>
- Rosa, Cuéllar and Lehmann 2024** Rosa, Javier de la, Álvaro Cuéllar and Jörg Lehmann. "El proyecto Modernífa: modernización ortográfica del teatro del Siglo de Oro con modelos de lenguaje". *Anuario Lope de Vega* vol 30 (2024): 410–425. <https://doi.org/10.5565/rev/anuariolopedevega.530>
- Río Riande 2022** Río Riande, Jimena del. "Humanidades Digitales o las Humanidades en la intersección de lo digital, lo público, lo mínimo y lo abierto". *Publicaciones de la Asociación Argentina de Humanidades Digitales*. vol. 33 (2022). <https://doi.org/10.24215/27187470e038>
- Ríos 1999** Ríos Mestre, Antonio. "La transcripción fonética automática del diccionario electrónico de formas simples flexivas del español: estudio fonológico en el léxico". *Estudios de Lingüística del Español* vol. 4 (1999): 1–2. <http://elies.rediris.es/elies4/>
- Sanz-Lázaro 2021a** Sanz-Lázaro, Fernando. *Fonemas*. 2021a <https://doi.org/10.25365/phaidra.466>
- Sanz-Lázaro 2021b** Sanz-Lázaro, Fernando. *Silabeador*. 2021b <https://doi.org/10.25365/phaidra.467>
- Sanz-Lázaro 2023a** Sanz-Lázaro, Fernando. "Del fonema al verso: una caja de herramientas digitales de escansión teatral". *Revista de Humanidades Digitales* vol. 8 (2023): 74–89. <https://doi.org/10.5944/rhd.vol.8.2023.37830>
- Sanz-Lázaro 2023b** Sanz-Lázaro, Fernando. *libEscansión*. 2023 <https://doi.org/10.25365/phaidra.465>

**Sanz-Lázaro 2023c** Sanz-Lázaro, Fernando. "Planteamientos digitales del drama aurisecular: automatización de la escansión métrica de obras teatrales". In *La escritura en el tiempo: investigaciones en torno a la literatura hispánica*, edited by Inés González Cabeza, Érika Redruello Vidal and Raquel de la Varga Llamazares, pp 109–126, León: Universidad de León, 2023.

**Sanz-Lázaro 2024** Sanz-Lázaro, Fernando. *libEscansión*. 2024 <https://github.com/fsanzl/libEscansion>

**Sanz-Lázaro 2024b** Sanz-Lázaro, Fernando. *Ritmo y estructura de la comedia áurea: posibilidades y límites del análisis digital automático del teatro español clásico* [PhD thesis]. University of Vienna, 2024. <https://doi.org/10.25365/thesis.75102>

**Sound and Meaning 2022** Sound and Meaning in Spanish Golden Age Literature. *Search corpus*. 2022 [https://soundandmeaning.univie.ac.at/?page\\_id=175](https://soundandmeaning.univie.ac.at/?page_id=175)

**Sánchez Jiménez 2017** Sánchez Jiménez, Antonio. "Acentos contiguos en los romances de la *Arcadia* (1598), de Lope de Vega". *Atalanta. Revista de las letras barrocas* vol. 5.1 (2017): 5–61. <https://doi.org/10.14643/51A>

**Varela Merino, Moíno Sánchez and Jauralde Pou** Varela Merino, Elena, Pablo Moíno Sánchez and Pablo Jauralde Pou. *Manual de métrica española*. Madrid: Castalia, 2005.

**Vega 2015** Lope de Vega, Félix *La dama boba*, edited by Marco Presotto. Collab. Sònia Boadas, Eugenio Maggi and Aurèlia Pessarrodona. Barcelona/Bologne: Prolope/Alma Mater Studiorum - Università di Bologna, 2015. <https://doi.org/10.6092/UNIBO/LADAMABOBA>



This work is licensed under a Creative Commons Attribution-NoDerivatives 4.0 International License.