Peeking Inside the Rhythmic Possibilities of the Portuguese *Decassílabo*

Adiel Mittmann

Universidade Federal de Santa Catarina, Brazil
adiel@mittmann.net.br
🄾 0000-0003-2184-2955

Gabriel Esteves

Universidade Federal de Santa Catarina, Brazil
gabrielesteues@gmail.com
🄾 0000-0003-4719-6672

Alckmar Luiz dos Santos

Universidade Federal de Santa Catarina, Brazil
alckmar@gmail.com
🄾 0000-0002-7896-0103

Abstract

The Portuguese *decassílabo* has a rich internal structure, which is usually analyzed in terms of the rhythmic patterns revealed by verse scansion. In this article, we aim to explore how the stress of one syllable influences another. In order to achieve this goal, we use raw rhythmic patterns, that is, ones in which stress clashes have not been resolved. To exploit these patterns, we apply three methods: tile plots, indices and graphs. Based on our investigation of a corpus of 24 poets, we find that poets who employ similar rhythmic patterns may not share the same preferences concerning stress clashes and may produce different dependence relations. Because many of the associations among syllables reflect the use of one of the two basic types of *decassílabo*, namely the heroic and Sapphic variants, we also propose verse classifications that are useful for understanding and comparing the works of different poets. The main conclusion of this study is that although the *decassílabo* has certain general features, much room remains for poets to create their own interpretations of the form.

1 Introduction

The *decassílabo* has a long history in the poetry of Portugal and Brazil. Throughout its development, much has been written about its features, but there have been few efforts to analyze the large corpora of works that use the form. Such an analysis may, however, promote a better understanding of the internal structure of the *decassílabo*. While it is common to talk about the rhythmic patterns that poets use, little quantitative data has been produced to show how the stress of one syllable influences the stress of another or how stress clashes are handled by different poets.

This article introduces methods to visualize and analyze the relationships between the different syllables in a verse. These methods use the output produced by Aoidos (Mittmann, Wangenheim, et al. 2016), our automatic scansion tool, in a special way: instead of considering actual rhythmic patterns, they harness information about each syllable's raw stress. As such, stress clashes are left unresolved.

Our methods can be divided into three groups: 1) those that deal with dependence between syllables, 2) those that classify verses into basic categories related to traditional variants (heroic or Sapphic) and 3) those that explore stress clashes. These methods produce percentages and indices, and the dependence relationships among the syllables are also depicted graphically using tile plots and graphs.

Although only the Portuguese *decassílabo* is considered here, there is no reason to believe that these methods would not work equally well with other types of verse and with at least some other languages such as Spanish and Italian.

Throughout this article, we use the following conventions: the smallest number in a table column is underlined; the largest one is shown in bold. Numbers that appear within parentheses were calculated from a small set of verses (fewer than 100) and should therefore be interpreted more carefully. Verses listed in a single block are not sequential in the poems that they were taken from; instead they are merely grouped together for convenience's sake.

This article is structured as follows: The Section 2 describes the corpus used in our analyses. The Section 3 then gives general information about the *decassílabo* and raw rhythmic patterns. The Section 4 describes our methods and the analyses that they produced. Our concluding remarks are given in the Section 5.

2   Corpus

This article analyzes a corpus of 24 works, which were written at various times between the 16th and 21st centuries, as shown in Table 1. Each work was assigned a three-letter code derived from its author's name. In this article. these codes are used to refer to the relevant work or its author.

While most of the works analyzed are single books, some (i.e. MAT, NTA, RAB, ANJ) represent the complete works of their respective authors. Some of the works contain *decassílabos* exclusively (this is the case for long epic poems such as CAM and TEI) while others include many other types of verse. Concerning the latter, this study only considers lines of 10 syllables, and the Table 1 only take these into account. One work, PIN, is a translation of Dante's *Divina Commedia* while all the other works were originally written in Portuguese.

The number of *decassílabos* in each work varies greatly, ranging from 1,192 in EMM to 62,517 in GLM. Since works are analyzed individually, there is no risk of a single work dominating the calculations. For our purposes, the presence in each unit of analysis (i.e. the poet's work) of more than 1,000 lines is more important than the total size of the overall corpus (more than 180,000 lines).

Code	Birth	Poet	Work	Verses
CAM	1524	L. de Camões	*Os Lusíadas*	8,816
MEN	1600	Sá de Meneses	*Malaca Conquistada*	10,637
MAT	1636	G. de Matos	Complete works	4,187
DUR	1722	S. R. Durão	*Caramuru*	6,672
COS	1729	C. M. da Costa	*Obras Poéticas*	5,736
NTA	1740	N. T. de Almeida	Complete works	2,074
GAM	1741	B. da Gama	*O Uraguai*	1,377
GON	1744	T. A. Gonzaga	*Cartas Chilenas*	4,172
JAM	1761	J. A. de Macedo	*O Oriente*	8,694
BOC	1765	M. M. B. du Bocage	*Poesias Eróticas*	3,335
MAG	1811	G. Magalhães	*Suspiros Poéticos*	3,787
PIN	1822	X. Pinheiro	*Divina Comédia*	14,227
DIA	1823	G. Dias	*Os Timbiras*	2,004
RAB	1826	L. Rabelo	Complete works	2,250
FAG	1841	F. Varela	*Anchieta*	8,481
SIL	1854	D. Silveira	*Lises e Martírios*	2,070
EMM	1866	E. de Meneses	*Mortalhas*	1,192
TIG	1822	B. Tigre	*Bromilíadas*	3,406
ANJ	1884	A. dos Anjos	Complete works	5,843
NUN	1897	C. A. Nunes	*Os Brasileidas*	8,502
GLM	1951	G. Mattoso	*Sonnettudo*	**62,517**
COL	<1968	C. A. de Oliveira Leite	*Caxias*	3,866
MNC	<1994	M. N. Costa	*Antônio Conselheiro*	2,265
TEI	<2016	J. C. S. Teixeira	*Famagusta*	9,273
				185,383

Table 1: Poets and works in the corpus used in this article, as ordered by the year of the poet's birth

Most of the poets are well known to mainstream literary critics, however COL, MNC and TEI are exceptions. Their years of birth are unknown to us, and the years shown in Table 1 are the years of publication of their respective works. Those three works are lesser-known epic poems; COL and MNC deal respectively with the lives of Duke of Caxias and Antônio Conselheiro, while TEI is a fantastic retelling of the fall of Famagusta to the Ottomans. While the versification of COL, and especially TEI, is more traditional, MNC uses the *decassílabo* in a distinctive way, as we will see.

Not all the source texts are of equal quality. The XML file read by our scansion tool may contain errors that were introduced in any of several previous stages. The print edition may, for example, have included misprints that could only be "fixed" by comparing different editions; the OCR algorithm may have failed to recognize words properly; human editors may have inadvertently mangled verses when updating the text's orthography. As a consequence, the longer a text has been in our corpus, the greater its reliability; while some texts (such as CAM or GAM) have been improved and corrected over many years, others (such as JAM or COL)are recent additions that have not yet undergone this refinement.

The text of GLM is a very recent addition to the corpus and was particularly challenging to analyze. The author made the text freely available to us as a carefully prepared digital edition, which should have simplified the task of feeding it into scansion tools. However, the text used a non-standard spelling system based on a pre-1943 orthography; substantial effort was therefore required to "update" many words in the work.

3   Background

The Portuguese *decassílabo* consists of 10 syllables which are counted from the first syllable of a verse to its last stressed syllable. Any additional syllables are not counted and typically are not rhythmically relevant. Strictly speaking, the word "decassílabo" may refer to any verse that is 10 syllables long; however, this term usually refers to a more specific type of verse, which should technically be called *decassílabo italiano* based on its alleged Italian origins.[1] This is the kind of *decassílabo* that this article is concerned with.

The *decassílabo* is characterized by a mandatory stress on either the sixth or the fourth syllable. Verses that follow the former type are called *heroic*; those of the latter variety are called *Sapphic*. The sixth-syllable variant is the more important one; no poets known to us object to the stress placement on the sixth syllable, but some do reject the choice of the fourth syllable. A verse whose main stress is on the sixth syllable may also contain a stress on any of its other syllables. However, a verse whose main stress is on the fourth syllable has one further restriction: its seventh syllable cannot be stressed. This means that either the eighth syllable must be stressed (so that the seventh is weakly stressed) or the seventh syllable must have no stress at all. If, however, a 10-syllable verse contains stresses on the fourth and the seventh syllables, then it is usually not considered a *decassílabo italiano* but a verse of Provençal or Iberian origin depending on one's critical framework.[2] Verses that do not follow any of these patterns may also be found in works that are thought to use only *decassílabo*, but typically, however, there are few such variations and they can often be plausibly explained as scansion tool errors or mistakes in the original text.

Dividing verses into heroic and Sapphic variants can be challenging when both the fourth and the sixth syllables are stressed. Consider the following verse by CAM:

> Abrindo a porta ao vasto mar patente,

Readers of this line may place stronger stress on either the fourth ("**por**-ta") or the sixth syllable ("**vas**-to"). There is no simple answer to the question of whether this verse is heroic or Sapphic. The scansion of a given line also depends

---

[1] Spina (2003), for instance, argues that "when we say that Sá de Miranda introduced the *decassílabo italiano* into the Portuguese poetry, this does not mean that this meter is genuinely Italian; it had already been widely used among troubadours as well as in France from the mid-10th century. What was taken from the Italian *decassílabo* was its caesura."

[2] Bandeira (1997), for instance, calls this variation "decassílabo de gaita galega" while Said Ali (1999) describes it as "hendecassílabo ibérico" and Proença (1955) understands it as a "verso provençal" and so on.

on the preceding and following lines, the internal structure of syllabic groups and the semantic weight of words. There are also cases where one classification may be more likely, but it remains possible that the author preferred the other option:

> Vai de Calisto ao seu contrário Polo,

When reading this verse, also written by CAM, it may be tempting to strongly stress the fourth syllable ("Ca-**lis**-to") and so make it a Sapphic variant; this is especially compelling because of the symmetry between the first two prosodic units, each of which has three unstressed syllables followed by a stressed one. However, the argument could also be made that the author did not use Sapphic verses. Supporters of this view might highlight the rather weak stress on "seu" to read this as a heroic verse. This reading is not inconceivable since verses such as this one can also be found in CAM:

> Lembrando-se do seu passado pranto,

This line can only be plausibly understood as a heroic verse whose strongly stressed sixth syllable has its origins in the weak lexical stress on "seu".

The classification of verses with possible stresses on their fourth and sixth syllables as either heroic or Sapphic requires a judgment call that this article does not attempt to make. As will be seen, when classifying verses, we focus instead on those that are unambiguously Sapphic or heroic.

In usual scansion, the rhythmic patterns that are quantified and analyzed attempt to mimic the sequence of stressed and unstressed syllables that a human reader would produce when reading verses. This study, however, analyzes a different type of pattern, one that does not resolve any kind of stress clash. Consider, for instance, the following two lines by NUN:

> excruciantes vindo ela a contorcer-se,
> do grande rio, vindo essas guerreiras

Both these lines contain a stressed sixth syllable, but in the first case, the stress on "**vin**-do" is weakened since the sixth syllable ("**e**-la") must be stressed by the very nature of the line. In the second case, the stress falls on "**vin**-do" in the sixth position and therefore weakens the stressed syllable in the next word ("**es**-sas").[3] When the patterns of these verses are given as 3-6-10 and 2-4-6-10, these lexical stress clashes are hidden from view. Instead the clashes are always effectively resolved by weakening one of the two successive syllables.[4] In fact, traditional versification techniques prescribe that poets should avoid lexical clashes whenever possible. However, this study considers all lexical stress possibilities, regardless of whether they involve two or more adjacent stressed syllables. As such, the patterns for these verses are understood as 3-6-7-10 and 2-4-6-7-10, respectively, even though this does not reflect the way these lines are read aloud. These patterns are referred to as "raw" and opposed to the patterns that arise when stress clashes are resolved.

---

[3] It should also be noted that in both cases, the final syllable of "vindo" is joined to the next syllable by a synaloepha so that there are no intervening syllables that would prevent a lexical stress clash.

[4] Said Ali (1999) notes that "the rhythmic movement does not support the collision of two strong syllables pronounced with the very same intensity."

Given the nature of raw patterns, this study often states that a syllable "can" or "may" be stressed. We also refer to rhythmic possibilities since the actual stress patterns are not known before the clashes are resolved.

It is important to state openly that "raw rhythmic patterns" are artificial in the sense that they do not correspond to how a verse is actually read. Indeed, they may contain stress clashes that could be considered unpronounceable. The justification for using raw patterns anyway is that they are a useful signal that can be measured from verses. That this signal is more or less detached from the phonetic reality is not especially relevant so long as the signal is correctly measured and provides useful information.

4   Methods

This section presents the methods used to analyze raw rhythmic patterns in this study. These methods were applied to all works in the corpus, and we describe the results below.

4.1   Dependence

Three kinds of tile plots were used in this study: positive (left-hand column of Figure 1), negative (right-hand column of Figure 1) and contrasting (Figure 2).

In the *positive* tile plot, each numbered row filters the verses available to include only those that may be stressed on the syllable corresponding to the row number. The total number of verses (in thousands) that have successfully passed through the filter is shown to the right of each row. Row number 1, for example, only contains information about verses whose first syllable may be stressed. The number in each cell is the percentage of verses that may be stressed on the syllable indicated by the corresponding column so that, for instance, the fourth cell in the first row tells us the percentage of verses that may be stressed on the first and the fourth syllables;this figure is relative to the total number of verses that may be stressed on the first syllable. The color of the tiles ranges from dark blue (for lower values) to light blue (for higher values).

The plot is not symmetrical because percentages are calculated relative to the total number of verses considered in each row. In the first positive tile plot of Figure 1, for example, the cell in the first row of the second column and the cell in the second row of the first column have the values of 19% and 12% respectively. Nevertheless they both reflect the same total number of verses— 537. The 10th row in a positive tile plot counts all the verses since all those included in the plot are stressed on the 10th syllable. As such, this row gives a summary of the rhythmic possibilities for all these verses.

A quick glance at a positive tile plot can reveal interesting patterns. In the case of both CAM and GLM, the appearance of the even-numbered columns in a brighter shade of blue indicates that these poets exhibit iambic tendencies. It is clear, however, that GLM favors this type of verse more frequently. Turning to the last row, we can create a simple *iambicity index* by adding up the percentages

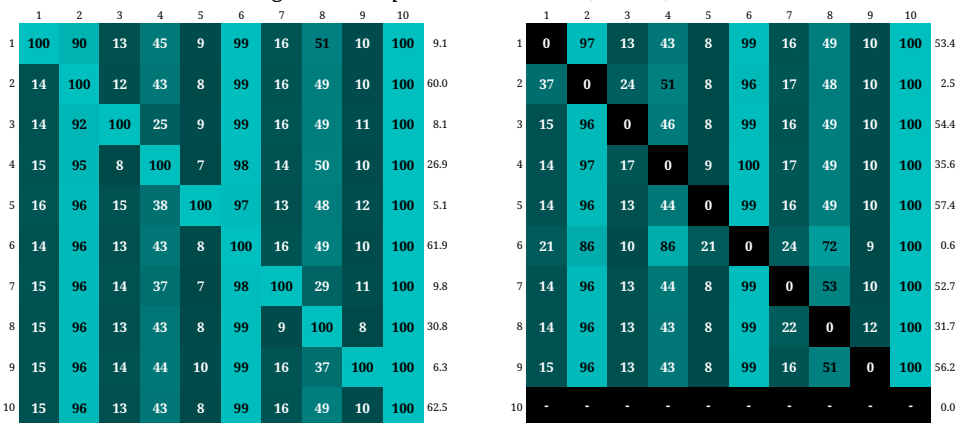|	1	2	3	4	5	6	7	8	9	10	
1	100	19	40	53	6	98	5	42	2	100	2.8
2	12	100	13	55	6	97	5	45	2	100	4.4
3	37	19	100	16	6	99	5	43	3	100	3.0
4	33	53	11	100	3	96	5	43	2	100	4.5
5	32	50	35	28	100	96	5	45	3	100	0.6
6	32	49	35	50	6	100	5	42	2	100	8.6
7	27	51	30	51	6	95	100	26	2	100	0.5
8	31	52	34	51	6	95	3	100	1	100	3.8
9	31	49	41	51	8	100	3	15	100	100	0.2
10	32	49	34	51	6	98	5	43	2	100	8.8

	1	2	3	4	5	6	7	8	9	10	
1	0	64	31	51	6	98	6	44	2	100	6.0
2	51	0	55	48	6	98	5	41	2	100	4.5
3	29	65	0	70	6	97	6	44	2	100	5.8
4	31	46	59	0	9	100	5	44	2	100	4.3
5	32	49	34	53	0	98	5	43	2	100	8.3
6	32	54	13	95	11	0	12	91	0	100	0.2
7	32	49	34	52	6	98	0	44	2	100	8.4
8	33	48	35	52	6	100	7	0	4	100	5.0
9	32	49	34	52	6	98	5	44	0	100	8.6
10	-	-	-	-	-	-	-	-	-	-	0.0

(a) L. de Camões' *Os Lusíadas* (CAM), with 8,816 verses.

	1	2	3	4	5	6	7	8	9	10	
1	100	20	51	48	9	87	17	53	9	100	1.6
2	21	100	20	49	7	86	17	53	8	100	1.6
3	49	19	100	16	9	97	18	43	9	100	1.7
4	47	48	16	100	6	74	14	60	8	100	1.7
5	50	41	56	36	100	92	18	50	12	100	0.3
6	43	42	50	37	8	100	17	44	8	100	3.3
7	47	45	49	39	8	94	100	26	10	100	0.6
8	45	45	39	52	7	78	8	100	6	100	1.9
9	48	43	50	48	12	89	20	39	100	100	0.3
10	44	43	45	44	8	88	16	51	8	100	3.8

	1	2	3	4	5	6	7	8	9	10	
1	0	60	41	42	7	90	15	49	7	100	2.1
2	61	0	64	41	8	90	15	49	8	100	2.2
3	39	63	0	68	6	81	15	57	7	100	2.1
4	41	39	69	0	9	100	17	44	7	100	2.1
5	43	43	44	45	0	88	16	51	7	100	3.5
6	50	51	11	98	5	0	8	98	7	100	0.4
7	43	43	44	45	7	87	0	55	7	100	3.2
8	42	41	52	36	8	100	24	0	10	100	1.9
9	43	43	45	44	7	88	16	52	0	100	3.5
10	-	-	-	-	-	-	-	-	-	-	0.0

(b) G. Magalhães' *Suspiros Poéticos* (MAG), with 3,787 verses.

	1	2	3	4	5	6	7	8	9	10	
1	100	90	13	45	9	99	16	51	10	100	9.1
2	14	100	12	43	8	99	16	49	10	100	60.0
3	14	92	100	25	9	99	16	49	11	100	8.1
4	15	95	8	100	7	98	14	50	10	100	26.9
5	16	96	15	38	100	97	13	48	12	100	5.1
6	14	96	13	43	8	100	16	49	10	100	61.9
7	15	96	14	37	7	98	100	29	11	100	9.8
8	15	96	13	43	8	99	9	100	8	100	30.8
9	15	96	14	44	10	99	16	37	100	100	6.3
10	15	96	13	43	8	99	16	49	10	100	62.5

	1	2	3	4	5	6	7	8	9	10	
1	0	97	13	43	8	99	16	49	10	100	53.4
2	37	0	24	51	8	96	17	48	10	100	2.5
3	15	96	0	46	8	99	16	49	10	100	54.4
4	14	97	17	0	9	100	17	49	10	100	35.6
5	14	96	13	44	0	99	16	49	10	100	57.4
6	21	86	10	86	21	0	24	72	9	100	0.6
7	14	96	13	44	8	99	0	53	10	100	52.7
8	14	96	13	43	8	99	22	0	12	100	31.7
9	15	96	13	43	8	99	16	51	0	100	56.2
10	-	-	-	-	-	-	-	-	-	-	0.0

(c) G. Mattoso's *Sonnettudo* (GLM), with 62,517 verses.

Figure 1: Tile plots of rhythmic possibilities. In the plots on the left, each row indicates the percentage of rhythmic possibilities when the corresponding syllable **may** be stressed; the plots on the right show this percentage when the syllable **may not** be stressed.
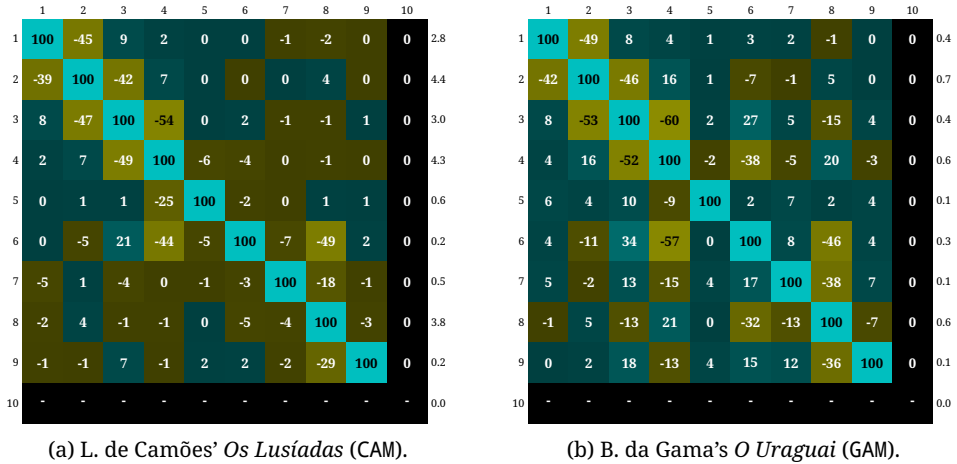
	1	2	3	4	5	6	7	8	9	10	
1	100	-45	9	2	0	0	-1	-2	0	0	2.8
2	-39	100	-42	7	0	0	0	4	0	0	4.4
3	8	-47	100	-54	0	2	-1	-1	1	0	3.0
4	2	7	-49	100	-6	-4	0	-1	0	0	4.3
5	0	1	1	-25	100	-2	0	1	1	0	0.6
6	0	-5	21	-44	-5	100	-7	-49	2	0	0.2
7	-5	1	-4	0	-1	-3	100	-18	-1	0	0.5
8	-2	4	-1	-1	0	-5	-4	100	-3	0	3.8
9	-1	-1	7	-1	2	2	-2	-29	100	0	0.2
10	-	-	-	-	-	-	-	-	-	-	0.0

(a) L. de Camões' *Os Lusíadas* (CAM).

	1	2	3	4	5	6	7	8	9	10	
1	100	-49	8	4	1	3	2	-1	0	0	0.4
2	-42	100	-46	16	1	-7	-1	5	0	0	0.7
3	8	-53	100	-60	2	27	5	-15	4	0	0.4
4	4	16	-52	100	-2	-38	-5	20	-3	0	0.6
5	6	4	10	-9	100	2	7	2	4	0	0.1
6	4	-11	34	-57	0	100	8	-46	4	0	0.3
7	5	-2	13	-15	4	17	100	-38	7	0	0.1
8	-1	5	-13	21	0	-32	-13	100	-7	0	0.6
9	0	2	18	-13	4	15	12	-36	100	0	0.1
10	-	-	-	-	-	-	-	-	-	-	0.0

(b) B. da Gama's *O Uraguai* (GAM).

Figure 2: Contrasting tile plots of positive and negative rhythmic possibilities (as seen in Figure 1)

Code	Index	Code	Index	Code	Index	Code	Index
MAG	<u>106.4</u>	FAG	131.3	DUR	144.2	COL	159.9
MNC	114.0	TIG	135.3	PIN	144.7	CAM	162.0
EMM	114.9	BOC	135.5	NTA	150.2	JAM	168.9
MAT	120.2	NUN	137.1	MEN	154.4	DIA	185.2
ANJ	129.4	COS	141.4	SIL	156.4	GON	191.0
RAB	129.8	TEI	142.7	GAM	159.0	GLM	**226.1**

Table 2: Iambicity index for all poets, in ascending order

of syllables in even positions (excluding the 10th syllable) and subtracting those in odd ones; the results are shown in Table 2. The index reflects what can be seen in the plots. It is striking also that GLM stresses the second syllable in the vast majority of cases. This suggests a very peculiar usage of the *decassílabo*.

The *negative* tile plot is similar in most respects to the positive one. The main difference is that each numbered row selects only the verses that may *not* be stressed on the corresponding syllable. When a percentage cannot be calculated due to the lack of any verses in the row, a dash is shown instead of a number. When the percentage is exactly zero, the cell is colored black. The color scheme distinguishes values that are exactly zero from those that have only been rounded to zero. The latter are shown in dark blue instead of black.

The *contrasting* tile plot is calculated by subtracting the negative plot from the corresponding positive plot. Negative values are common in this type of plot and are shown in yellow: the lighter the tile, the more negative the number.

In a contrasting tile plot, each cell contains dependence information. Positive values indicate that if the syllable in the row number position may be stressed,

then the syllable corresponding to the column number also tends to be stressed. Conversely, negative values mean that the other syllable tends not to be stressed.

To some degree, the information in the tile plots makes explicit the relationship between heroic and Sapphic verses. For instance, the negative tile plots for all poets show that when the sixth syllable is not stressed, the fourth and the eighth syllables light up; in other words, when a verse cannot be heroic, it will most likely be Sapphic. But the reverse relationship does not always hold; the negative tile plots for Figure 2 show that when the fourth syllable cannot be stressed, only the poet MAG puts a noticeably increased stress on the sixth syllable. This is due to the larger proportion of Sapphic verses in his work. The reasoning here is as follows: when there are few Sapphic verses, the lack of a stress on the fourth syllable does not mean much for the sixth one. On the other hand, the lack of a stress on the sixth syllable influences the fourth syllable since the only realistic alternative to a heroic verse is a Sapphic one, however few of them there might be.

Among poets who use Sapphic verse more often, a similar effect can be seen regarding the eighth syllable: when it cannot be stressed, a stress on the sixth syllable is more common. Such poets also display other subtler dependence relationships. For instance, Figure 2 shows that the third syllable has a positive influence on the sixth one. In the case of CAM, who uses few Sapphic verses, this influence is minimal. It is, however, very significant for GAM, who employs Sapphic verses more often.

The fifth and the seventh syllables would never be stressed in a realistic reading of a *decassílabo*. This means that when they may be stressed, they will likely be involved in a stress clash that allows them to become weakened in favor of the fourth, sixth or eighth syllables. This can be seen in the positive tile plots for Figure 1: in the rows for the fifth and seventh syllables, the sixth syllable may almost always be stressed as well. The poet GLM shows a marked preference for a stressed second syllable; this is visible in his positive tile plot where a possible stress on the first and the third syllables is very frequently associated with a stress clash with the second syllable.

By adding together the magnitude of the values shown in a contrasting plot, we can measure the general influence of a potential stress in one syllabic position on the possible stress in another position. We call this measure the *dependency index*. When calculating it, we include only cells that have been calculated based on the information from at least 100 verses in both the positive and negative tile plots—the intention is to exclude cells whose values might be too noisy. Table 3 shows the dependency index calculated for all works in the corpus.

The dependence relationship between syllables, as shown in a contrasting tile, can also be visualized as a graph in which nodes represent syllables and edges show their dependence on one another. Graphs for six works are shown in Figure 3. The thickness of the edges reflects the degree of dependence. To reduce the amount of clutter, only cells with an absolute value greater than or equal to 5.0 are represented as edges. Edges have an arrowhead that touches the syllable that is being influenced. The shape of the arrowhead indicates whether the influence is positive (a regular arrowhead) or negative (an inverted

Code	Index	Code	Index	Code	Index	Code	Index
GLM	<u>3.87</u>	MAT	8.70	MEN	9.60	EMM	11.05
TEI	6.33	COS	9.05	BOC	9.77	NTA	11.07
CAM	7.36	MNC	9.06	MAG	10.22	DUR	11.69
COL	8.11	PIN	9.27	ANJ	10.29	JAM	11.82
GON	8.27	RAB	9.51	FAG	10.45	SIL	11.95
TIG	8.50	NUN	9.60	DIA	10.81	GAM	**13.62**

Table 3: Dependency index for all poets, in ascending order

arrowhead). In Figure 3c, for example, one edge joins the sixth and the third syllables; the edge's arrowhead is regular so the relationship is positive. The edge points from the sixth to the third syllable, which means that when the sixth syllable is stressed, the third one tends to be stressed too. This edge is directly based on the sixth row of the third column of the contrasting plot in Figure 2a.

These graphs also highlight certain properties of the corpus. The tendency to avoid stress clashes is clear: there are often (though not always) negative edges between adjacent syllables. We also observe simpler networks among the poets who rely heavily on the heroic *decassílabo*, i.e. (GLM, TEI and CAM). Notably in the work of these poets the sixth syllable is not influenced by the other syllables, presumably because it is almost always stressed anyway.

4.2 Verse Types

Readers of heroic and Sapphic verses can usually identify the two forms. Not many verses truly allow for both heroic and Sapphic readings and when both interpretations are permitted, readers may interpret these verses either way. Since this article only considers rhythmic possibilities, verses are classified into the traditional types according to certain criteria.

A verse is said to be *purely heroic* if it can be stressed on the sixth syllable but cannot be stressed on the fourth syllable. A purely heroic verse, thus, does not allow for a Sapphic reading and must be read as a heroic verse.

A verse is said to be *purely Sapphic* if it can be stressed on the fourth syllable but cannot be stressed on the sixth syllable. A purely Sapphic verse, thus, does not allow for a heroic reading. A further constraint is that a purely Sapphic verse must either not be stressed on the seventh syllable or, if such a stress is allowed, then the eighth syllable must be able to carry a stress too. In that case, when the verse is read aloud, the stress clash is resolved in favor of the eighth syllable.

An *ambiguous* verse is one that can be either heroic or Sapphic because it can be stressed on the fourth and the sixth syllables. Such verses are only ambiguous in the strict sense that they theoretically allow for either a heroic or a Sapphic reading; in practice, however, readers treat most if not all of these verses as either heroic or Sapphic.
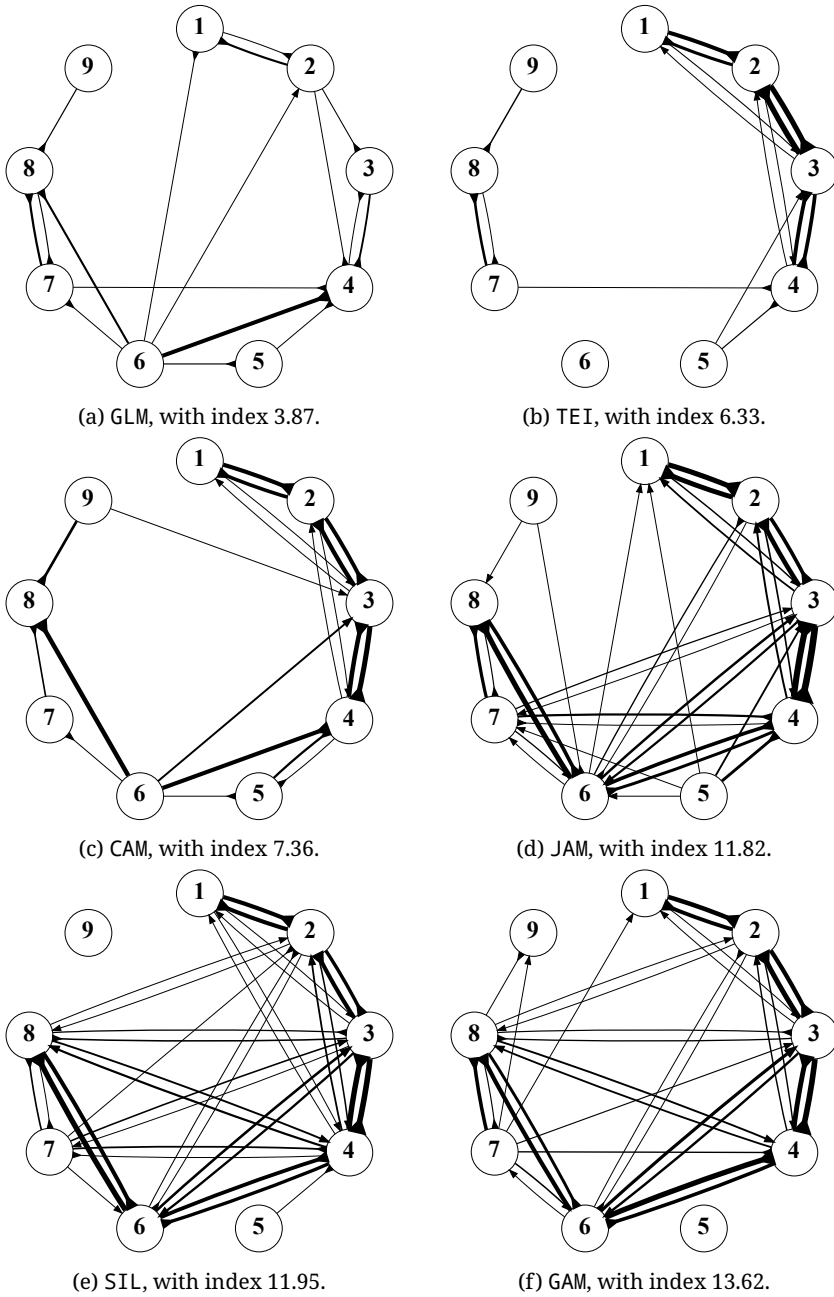
(a) GLM, with index 3.87.

(b) TEI, with index 6.33.

(c) CAM, with index 7.36.

(d) JAM, with index 11.82.

(e) SIL, with index 11.95.

(f) GAM, with index 13.62.

Figure 3: Influence of a possible stress on one syllable on other stress placements

Code	Heroic	Sapphic	Ambiguous	Provençal	Other	Eighth
NUN	58.53	11.86	<u>29.33</u>	(0.11)	(0.18)	94.0
FAG	57.65	12.50	29.81	(0.00)	(0.05)	99.3
GON	64.74	5.18	30.03	(0.00)	(0.05)	**100.0**
MNC	51.83	8.57	32.54	**(3.58)**	**(3.49)**	<u>30.4</u>
GAM	46.04	20.92	32.75	(0.00)	(0.29)	96.5
MAG	55.45	11.33	32.98	(0.05)	(0.18)	99.8
TEI	**65.16**	(0.73)	34.00	(0.00)	(0.11)	(69.1)
RAB	54.09	10.22	35.20	(0.18)	(0.31)	96.5
ANJ	43.44	20.54	35.58	(0.21)	(0.24)	88.0
DUR	48.49	14.22	37.22	(0.03)	(0.04)	98.7
BOC	48.46	12.92	37.69	(0.48)	(0.45)	98.8
EMM	43.12	16.61	39.26	(0.25)	(0.76)	87.4
MAT	52.16	6.31	40.29	(0.43)	(0.81)	92.4
MEN	48.30	11.04	40.42	(0.05)	(0.19)	97.3
DIA	37.72	20.91	40.77	(0.15)	(0.45)	99.0
NTA	44.07	13.36	41.76	(0.48)	(0.34)	98.6
GLM	56.78	0.79	42.24	(0.05)	(0.14)	82.9
COS	49.04	7.83	43.08	(0.00)	(0.05)	**100.0**
SIL	31.06	**24.83**	43.77	(0.19)	(0.14)	99.2
TIG	51.26	3.49	44.98	(0.03)	(0.23)	91.6
COL	51.63	(2.35)	45.01	(0.85)	(0.16)	(60.4)
JAM	<u>30.91</u>	22.46	45.97	(0.37)	(0.29)	99.4
PIN	40.25	11.53	47.72	(0.17)	(0.33)	97.8
CAM	48.38	2.16	**49.25**	(0.09)	(0.12)	94.7

Table 4: Frequency of verse types (%) for all poets, sorted based on their use of ambiguous verses. The *Eighth* column indicates the frequency (%) of a potentially stressed eighth syllable in purely Sapphic verses.

A *Provençal* verse is one that can be stressed on the fourth and seventh syllables, but cannot be stressed on either the sixth or the eighth syllables. In other words, a basic 4-7-10 rhythmic pattern is the only possible reading of such a verse.

Table 4 shows the percentage of each type of verse for all works in the corpus. Works are sorted based on the presence of so-called ambiguous verses.

An old question in the realm of Portuguese versification is whether the eighth syllable must be stressed in Sapphic verses. The last column of Table 4 shows the percentage of purely Sapphic verses whose eighth syllable may be stressed. For most works, the rate is more than 90% and among those with at least 100 purely Sapphic verses, only one work had a result below 80% (MNC, with 30.4%). It should be noted that in purely Sapphic verses, as defined in this study, there can be no stress on the fifth, sixth, seventh and ninth syllables. A consequence is that if the eighth syllable is not stressed, there will be a total of five unstressed syllables in a row, which is highly uncommon in poetry. We may therefore expect the frequency of a potentially stressed eighth syllable to

be high. As such, a rate of less than 95% probably indicates that in the poet's style, the eighth syllable tends to bear a rather artificial stress since it is not generally derived from lexical stress.

The last column of Table 4 shows that for two poets, (GON and COS), up to a rounding error of one decimal place, 100% of their purely Sapphic verses allow for a stressed eighth syllable. This syllable will most likely be stressed when read since the last stressed syllable was the distant fourth one. Another five poets (FAG, MAG, DIA, SIL and JAM) have a rate of more than 99%.

On the other hand, there are poets such as ANJ, whose purely Sapphic verses may have a stressed eighth syllable only 88.0% of the time. We can therefore find verses such as these in his works:

> Da Fantasia nos itinerários
> Eu, de saudades me despedaçando,
> Outras cabeças aparecerão
> Hoje, porém, que se desmoronou

We would fully expect many, if not most, human readers to stress the eighth syllable in such verses despite the absence of a primary lexical stress there. The sequence without any stress from the fourth to the 10th syllables simply continues for too long for speakers not to add intermediate stresses. In the case of A. dos Anjos', this artificial stress on the eighth syllable in his purely Sapphic verses goes hand in hand with his highly artificial synaloephas.

4.3 Stress Clashes

A *stress clash* takes place when two adjacent syllables could potentially be stressed. When the verse is read aloud, such clashes are resolved by the reader since the two adjacent syllables cannot both be stressed. Consider the second line of Camões' *Os Lusíadas*, for example:

> Que da ocidental praia Lusitana,

Here, in principle, either the last syllable of "o-ci-den-**tal**" or the first one of "**prai**-a" could be stressed by the reader, but a stress on both syllables is not possible. Experienced readers will identify in advance that in this type of verse, a fairly regular heroic *decassílabo*, the fifth syllable is not stressed and so resolve the clash by stressing the sixth syllable, that is, the first syllable of "**prai**-a".

Although these latent stress clashes are not evident in an actual reading, their analysis can reveal stylistic tendencies since not all poets make equal use of them. A *stress clash index* can be calculated as follows. Stress clashes are represented in positive tile plots by the cells adjacent to the main diagonal; however, those percentages are relative to the number of verses in each row and, if we were to average them all, the resulting index could be skewed by the high percentages in rows with few verses. The index is instead calculated based on the total number of clashes, which can be obtained by adding together all the stress clash tiles after multiplying them by the total number of verses in the relevant row. Furthermore, because stress clashes are symmetrical (a stress clash between the first and second syllables also counts as a clash between the

Code	Index	3+4	4+5	5+6	6+7	Code	Index	3+4	4+5	5+6	6+7
MNC	<u>0.23</u>	3.61	11.34	6.22	10.22	ANJ	0.43	6.83	13.50	4.96	13.91
NUN	0.30	4.76	7.64	4.70	9.08	SIL	0.47	8.17	6.23	6.84	14.93
FAG	0.32	3.77	3.40	<u>3.46</u>	15.05	NTA	0.48	6.50	13.72	8.75	14.33
COL	0.34	(4.40)	(2.20)	3.96	14.83	DIA	0.51	3.82	8.11	7.28	**21.83**
JAM	0.35	5.94	1.18	6.03	17.34	RAB	0.56	9.57	3.91	8.71	15.45
CAM	0.35	10.53	11.05	9.26	<u>5.30</u>	BOC	0.57	9.51	3.48	5.82	19.55
GAM	0.35	<u>3.47</u>	4.51	5.68	11.51	MAT	0.58	6.44	10.23	10.94	18.27
TEI	0.37	(4.41)	**(26.47)**	6.37	9.90	COS	0.59	8.02	10.69	10.13	18.88
GON	0.39	5.56	<u>0.93</u>	5.59	5.89	PIN	0.61	**10.55**	8.90	12.50	18.53
DUR	0.41	8.85	1.16	7.02	13.54	EMM	0.63	10.10	7.58	**14.20**	18.29
MEN	0.41	9.20	4.60	9.38	13.45	MAG	0.64	9.79	4.66	8.52	17.29
TIG	0.42	10.08	11.76	8.30	11.68	GLM	**0.73**	7.66	15.12	8.79	17.20

Table 5: Stress clash index and stress clash frequency between syllable pairs. Poets are sorted by the relevant index. The minimum value in each column is underlined; the maximum value is in bold.

second and first), only numbers above the main diagonal are included. This total number is then divided by the total number of verses in the tile. In other words, the stress clash index is the expected number of stress clashes per verse.

Table 5 shows the stress clash index for all works in the corpus, with values ranging from 0.23 (MNC) to 0.73 (GLM) clashes per verse. In general, we may expect a lower index to correspond with stricter adherence to the traditional clash avoidance rule. The high rate of stress clashes in the verses of GLM may relate to the single-mindedness with which he wrote verses with a 2-6-10 pattern. As the reader already knows where the stresses will be, the poet can allow more clashes into his poetry and perhaps use them to his own stylistic advantage. On the other hand, the low index score of MNC may be explained by the infrequency of synaloephas in his work.[5] Synaloephas usually join unstressed syllables and so bring stressed syllables closer together. A low rate of synaloephas therefore translates to fewer stress clashes.

Of particular interest are stress clashes that involve a verse's main ictus. Since there is less room for interpretation in this case, the poet can produce interesting effects by forcing the reader to stress a syllable that would not normally be favored. This is because when there is a stress clash between, say, the first and the second syllables, the reader may resolve it either way and, if the poet specifically intends for one syllable to be stressed, the reader may not notice. In contrast, when the clash occurs between the sixth and the seventh syllables, the reader knows that given tradition and the innate stanzaic tendency towards regularity, it is the former syllable that must bear the stress, whatever word happens to be there. Table 5 therefore also shows the frequency (%) of verses with a clash between the fourth and sixth syllables and their adjacent syllables. In the former case, only purely Sapphic verses are counted; in the latter, we consider only purely heroic ones.

---

[5] In experiments yet to be published, we show that synaloephas are eight times more common in the works of CAM than in those of MNC.

The following examples illustrate the effect achieved through a clash between the sixth and seventh syllables of purely heroic verses. CAM, who uses such clashes in only 5.30% of cases, provides this model:

> A sazão e o lugar, fazem cruezas
> Assi foi do Saber, alto e profundo,

The effect is remarkable because words that one would expect to be stressed ("**fa**-zem", "**al**-to") are instead robbed of their rhythmic strength. The following two lines come from DIA, who used stress clashes in 21.83% of cases:

> Tornou-lhe Jurucei: "Paz aos Gamelas,
> Ministros de Tupã, núncios da glória?"

5   Conclusion

The methods proposed by this article are a form of distant reading (Moretti 2013). Some of the works in our corpus are not well known (MNC, COL and TEI) and would normally receive little to no attention based on close reading approaches. However, because Aoidos (Mittmann, Wangenheim, et al. 2016) is used to scan all verses automatically, we can include such works in our analyses. We believe it is important not to exclude works solely because of their lack of popularity; we would like to include as many works as possible in our future corpora and so let the verses speak for themselves.

The methods introduced in this article are capable of distinguishing poets who might otherwise appear rhythmically very similar. This is possible because in examining raw stress patterns, we look beyond the actual resolution of a line's rhythm: instead we pay attention to stress clashes before they are resolved. The poets CAM and MAT, whose work is rhythmically similar (Mittmann, Pergher, et al. 2019) are, thus, shown to have different tolerance of stress clashes. On the other hand, the texts of MEN and DUR, though rhythmically different, show a very similar tolerance of these clashes. The analysis of such clashes is therefore another tool for telling poets apart. Our previous work has shown how poet B. Tigre was able to skillfully mimic rhythmic patterns in his parody of L. de Camões' *Os Lusíadas*. The present article shows that Tigre allows stress clashes between the sixth and the seventh syllables of purely heroic verses with twice the frequency of Camões.

The main conclusion of this article is that although we may speak about the general form of the *decassílabo* or its development over time, individual poets can and do use the *decassílabo* differently and so effectively create their own types of verse. To take one example, although the structure of heroic verse is generally quite flexible, GLM imposes the rule that the second syllable must be stressed in addition to the sixth. In the case of Sapphic verses, for some poets (for example, GON, FAG and COS), a stress on the eighth syllable is a strict requirement. Others clearly do not see the need for such a stress (or at least not one derivable from lexical stress) on the eighth syllable.

Acknowledgments

References

Bandeira, Manuel (1997). *Seleta de prosa*. Rio de Janeiro: Nova Fronteira.

Mittmann, Adiel, Paulo Henrique Pergher, and Alckmar Luiz dos Santos (2019). "What Rhythmic Signature Says About Poetic Corpora". In: Quantitative Approaches to Versification (2019).

Mittmann, Adiel, Aldo von Wangenheim, and Alckmar Luiz dos Santos (2016). "Aoidos: A System for the Automatic Scansion of Poetry Written in Portuguese". In: 17th International Conference on Intelligent Text Processing and Computational Linguistics (2016).

Moretti, Franco (2013). *Distant reading*. London: Verso.

Proença, Manuel Cavalcanti (1955). *Ritmo e poesia*. Rio de Janeiro: Simões.

Said Ali, Manuel (1999). *Versificação Portuguesa*. São Paulo: Editora da Universidade de São Paulo.

Spina, Segismundo (2003). *Manual de versificação românica medieval*. São Paulo: Ateliê Editorial.