

Combining rhythmic and syntactic analysis: an experiment on Dante’s *Comedy* with the new tool *TRIARS*

Ugo Conti

Università di Firenze, Florence, Italy

ugo.conti@unifi.it

Abstract

This paper is meant to introduce a specifically designed tool for the purpose of combining rhythmic and syntactic data to analyse literary works written in third rhyme and to look for identical rhythmical-syntactical structures in different tercets, finally studying their use in the many restrictions imposed by meter and its rules. *TRIARS* is still in its preliminary stage, and it has only been tested on Dante’s *Comedy*.

L’obiettivo del saggio è di illustrare un nuovo strumento informatico appositamente creato per permettere uno studio della terza rima dal duplice punto di vista ritmico e sintattico. Il fine è quello di ricercare e analizzare identiche strutture ritmico-sintattiche in diverse terzine, osservandone l’utilizzo nelle varie restrizioni imposte dal metro. Lo strumento, denominato *TRIARS*, è ancora allo stadio di prototipo e, attualmente, è stato testato solo sulla *Commedia*.

1. Introduction

Dante’s *Comedy* has been driving an ever-growing interest over the past decades, especially in recent years considering the upcoming seven hundredth anniversary in 2021. Concerning this paper, the last two decades have been crucial to the research on Dante’s works: several important and computer-based projects aimed at analysing all the different aspects of the poet’s works, among which a special mention needs to be made about those tools, such as *DanteSearch* and the *Archivio Metrico Italiano (AMI)*, essential to the analysis of Dante’s style from more than one perspective.

Having mentioned *DanteSearch* and *AMI*, the two main points on which the paper will focus, syntax and rhythm, are already defined. Both these aspects have been studied for years without the help of computers and an impressive amount of data and new ideas concerning Dante’s style have already been achieved and available; nonetheless, it is indeed possible to make a

further advancement in this direction with the help of digital humanities. Besides *DanteSearch* and *AMI*, many studies were already heading in this direction, such as [18], with a deep and accurate analysis of Dante's syntax and rhythm and a special attention to the *Comedy*, but no efforts have been made to define a tool that could explore both these aspects together. However, since syntax and rhythm are inextricably intertwined and therefore always influencing each other, an analysis focused on both these aspects is crucial to a fuller understanding of Dante's style.

This paper is meant to introduce *TRIARS* (*Terza Rima Informatizzata per l'Analisi Ritmica e Sintattica*), a specifically designed tool for the purpose of combining rhythmic and syntactic data to analyse literary works written in third rhyme and to look for identical structures in different tercets, finally studying their use in the many restrictions imposed by meter and its rules. The first results of this research, which began with the analysis of Dante's *Comedy*, have already been published on several occasions and are part of a Ph.D. project that led to the creation of *TRIARS* and to an introductory analysis of the whole *Comedy*, in order to confirm the truth of its theoretical basis and the mere feasibility of such an analysis [9]. As a matter of fact, although it generated these first and promising results ([8]; [10]), the tool itself is still preliminary and, as the paper will show, it needs to be improved and redefined. In these regards, the introduction and a detailed discussion of its features and main issues is a necessary step towards a definite version of *TRIARS*.

The paper is structured as follows: in Section 2, the main studies and scientific basis from the original project are reviewed; Section 3 reports the most important steps in the development of the early version of *TRIARS*; in Section 4, all the features and functionalities of *TRIARS* are described in detail; Section 5 outlines the major flaws and weaknesses of the current version of *TRIARS*, whilst all the essential ideas of improvement are examined; additionally, Section 7 focuses on the final purpose of the research, showing a brief example of compared analysis on the *Comedy* and other different poems in third rhyme, such as the *Caccia di Diana* by Giovanni Boccaccio and the *Trionfi* by Francesco Petrarca; finally, general conclusions are reported in Section 8.

2. Related works

The main inspiration for *TRIARS*, like many other stylistic studies and projects about Dante's *Comedy*, is the paper originally written by Gianfranco Contini in 1965, *Un'interpretazione di Dante* [11]. Contini had the unquestionable merit of leading the way on interpreting a new perspective on Dante's studies, which is a valid starting point even for today's research; as a matter of fact, he described the way in which Dante's memory works, stating that it is not "puramente verbale, per eccitazioni provenienti da oggetti affini, ma si organizza in figure ritmiche" and thus introducing the concept of "figura ritmica" in this research field.¹

1 For an English translation see [12]: "[Dante's memory] is not purely verbal, functioning through stimuli from similar objects; rather, it is organized by rhythmic figures".

Within just a few years the first concrete results of this influence on Dante’s studies emerged: during the early Seventies, other works about Dante’s style were published: *Ritmo e modelli ritmici* by Pier Marco Bertinetto [5] and two papers by Pietro Beltrami ([2]; [3]). Both these studies were crucial to gather together what have been defined as “echi di Dante entro Dante” by Contini and resulting in the first analysis of Dante’s “sintagmi ritmici” in the *Comedy*; particular mention needs to be made about Bertinetto’s work, which was undoubtedly innovative at the time, because it involved the use of a 360/44 IBM computer to assign the rhythmic ictuses in a semi-automatic way to Dante’s hendecasyllables.

Another work that was heavily influenced by Gianfranco Contini’s ideas is *L’autonomia del significante* by Gian Luigi Beccaria [2], where the relevance of any correlation between signified and signifier is heavily contested. This work offers a focal point for all stylistic studies up to present day, though it has been downsized for its inflexibility. Nonetheless, Beccaria’s pages suggest many paths towards the analysis of Dante’s “memoria ritmica”, with the extra value of keeping the third rhyme’s “grande ritmo” as a focal point together with the “piccolo ritmo” of the hendecasyllable. As these pages will show, this is one of the main concerns of this paper, since the *Comedy* is only a testing ground for a general theory on the analysis of the third rhyme, which for this reason is the key to the entire research.

Other important contributions have more recently been brought to Dante’s stylistic studies, ones that hark to the above-mentioned works. Generally speaking, since it is mostly dedicated to Boiardo’s poetry [15], the early works of Marco Praloran have been heavily influenced by this perspective, and by the dawn of the new century, together with Arnaldo Soldani and the Gruppo Padovano di Stilistica, they have led to the foundation of the *Archivio Metrico Italiano*,² whose first result was a book about Petrarca’s verse [16]. The *AMI* project started a wide analysis of early Italian poetry, from Giacomo da Lentini to Torquato Tasso’s verses, which resulted in a database where all these poets’ versification has been catalogued from different perspectives: metrical form, type of verses adopted, irregularities in the versification and, most importantly for this paper, rhythm. As Section 7 will show, one of the most relevant aspects that links this research to the *Archivio Metrico Italiano* rhythmic analysis is that a set of detailed and exact rules has been outlined to assure a homogeneous accentuation of Italian verses from all the archived poems and, more importantly, to prepare guidelines for the community of researchers that can provide equally homogeneous rhythmic analysis to the original database [17].

Moreover, together with the data from the *Archivio Metrico Italiano*, the other fundamental tool for *TRIARS* is the database *DanteSearch*, a project supervised by Mirko Tavoni, which consists in the lemmatization and grammatical analysis of all Dante’s works, and also in the syntactic annotation of *Convivio*, *Rime* and *Comedy*.³ The database allows the users to query the mentioned works on a morphological and syntactic level, but the most relevant aspect for this

2 <http://www.maldura.unipd.it/ami/php/index.php>. The database is currently under a redesigning process and it has been temporarily deactivated.

3 <http://www.perunaenciclopediaantescadigitale.eu:8080/dantesearch/>.

research is, of course, the syntactic annotation of Dante's *Comedy*, ideated and encoded in XML-TEI⁴ by Sara Gigli for her Ph.D. dissertation [13].

As the next Section will show more in detail, both these projects have been crucial to the realization of the first and preliminary experiments with *TRIARS*. In the next pages we are going to review its main features and its major flaws, with the purpose of highlighting its potential in relation to the analysis of the Italian third rhyme. Before that, a few notes on the development of the tool are necessary.

3. Developing *TRIARS*

As mentioned above, *TRIARS* stands for *Terza Rima Informatizzata per l'Analisi Ritmica e Sintattica*, so it is clear that the main point of the whole project is a more general analysis of the third rhyme as a structure of the Italian poetry: the aim of this research is not the single hendecasyllables, but the “grande ritmo” of its meter. Hence, Dante's *Comedy* is not the only work that is important to the project, even though the whole theoretical basis have been laid on Dante's third rhyme. The reason of this choice is obvious: having had the opportunity to repurpose data from already existing projects that were focused on both the main aspects considered by the research at issue, Dante's poem was a forced, but indeed very welcome, choice as a testing ground for a study on the third rhyme. In fact, this decision is also well integrated with the common practice brought on since the first studies after the publication of Contini's *Un'interpretazione di Dante*, in which the *Comedy* was indicated as a privileged testing ground for experimenting new theories about Italian's poetry stylistic aspects.

The development of *TRIARS* has been entirely entrusted to the Laboratorio di Informatica Umanistica of the Literature and Philosophy Department of the University of Florence,⁵ with the collaboration of dr. Giovanni Salucci and his *progettinrete*,⁶ who have already been active on several occasions in the digital humanities.

From a technical point of view, the web interface and the database have been created by using the web application *ASP.NET Core MVC*, a framework for the realization of web apps and API with a MVC (*Model-View-Controller*) conceptual scheme. The web interface, moreover, has been written in HTML5 and CSS3, while *TRIARS*' search engine uses the database MS SQL-Server; finally, both the search engine and the web interface have been realized with the platform WCM (*Web Content Manager*) powered by the already mentioned *progettinrete* and specifically designed for academic scientific projects involving the Digital Humanities.⁷

4 The TEI-XML encodings of *DanteSearch*'s are compliant with the P5 version of the TEI *Guidelines*, published in 2007 and continuously updated: <http://www.tei-c.org/Guidelines/P5/>.

5 The official page of the Laboratorio di Informatica Umanistica is the following: <https://www.liu.unifi.it/index.html>.

6 The home page of *progettinrete* may be reached at: <http://www.progettinrete.it>.

7 The official page of the WCM can be found here: <https://www.wcm.it/>.

Concerning the data from the *Archivio Metrico Italiano*, to build the database for *TRIARS* no adjustments were necessary: as already mentioned, one of the strongest points of the *AMI* is that its guidelines are clear and detailed, as to ensure that all the rhythmic analysis are homogeneous, so it has been decided not to intervene on the original accentuation as to be in line with all other related projects. On the other hand, *DanteSearch*’s data needed a direct intervention before it was possible to include it in the new database. Of course, *DanteSearch* data is as well thought as the *Archivio*’s, and, as a matter of fact, no editing work has been done concerning the theoretical aspects of the syntactic annotation; however, the highly detailed and in-depth analysis of Dante’s syntax was the first impediment to the realization of the actual database.

The main problem consisted in the nature of *DanteSearch* itself, which on one hand allows the user to perform in-depth syntactic queries, but on the other it sometimes causes the parameters to hide the very same results that are expected to be found by *TRIARS*, a tool that does not profit of all the distinctions established by Sara Gigli’s analysis. To provide an example, in *Inf.*, XXIX 91-3 and *Inf.*, XXX 79-81, the mere attribute of ‘rhetoric’ – assigned to the clause *ma che mi val* in *Inf.*, XXX 81, which has been marked up as ‘*coordinata avversativa interrogativa di tipo X retorica*’, but not present in the clause *ma tu chi sè* in *Inf.*, XXIX 93, which has been marked up as ‘*coordinata avversativa interrogativa di tipo X*’ –, hides the exact properties that *TRIARS* is supposed to be searching:

Inf., XXIX 91-3

“Latin siam noi, che tu vedi sí guasti	2 4 7 10
qui ambedue”, rispuose l’un piangendo;	1 4 6 8 10
“ <i>ma tu chi sè</i> che di noi dimandasti?”	2 4 7 10

Inf., XXX 79-81

Dentro c’è l’una già, se l’arrabbiate	1 4 6 10
ombre che vanno intorno dicono vero;	1 4 6 8 10
<i>ma che mi val</i> c’ho le membra legate?	4 7 10

The example here is quite relevant, since a minimal attribute as ‘rhetoric’ is capable of hiding the identical setting of the interrogative clause in the two tercets, which happen to be in a very similar rhythmic scheme for both the last hendecasyllables.

Having acknowledged this incompatibility between *DanteSearch* and *TRIARS*, the very first step has been to redefine the degree of depth of the syntactic analysis, which on *TRIARS* is set to the mere distinction of the macro-type of the clauses. To a better understanding of the kind

of intervention on the original XML codification, we submit the first three verses of the *Comedy* exactly as they are readable in the original mark-up downloadable from *DanteSearch*:

```
<s><cl type="esclam" function="princ">Ahi quanto<cl type="rel impl
lim" function="subord I" id="IfI1" next="IfI2"> a dir <cl type="int
x" function="subord II" id="IfI3" next="IfI4">qual era</cl></cl>
&grave; cosa dura<lb/> <cl type="rel impl lim" function="subord
I" id="IfI2"><cl type="int x" function="subord II" id="IfI4">esta
selva selvaggia e aspra e forte<lb/> <cl type="rel app antec"
function="subord III">che nel pensier rinova la
paura!</cl></cl></cl></lb/>
```

As to be more specific, at the level of this preparatory adjustments, the intervention has been localized in the definition of the *type* attribute, where all the detailed syntactic specifications have been simplified; to provide a more detailed example, attributes, like ‘rel impl lim’, which stands for ‘relativa implicita limitativa’, or ‘int x’, which stands for ‘interrogativa di tipo X’, have respectively been reduced to ‘rel’ and ‘int’. It follows that the relevant part of the XML code has now become:

```
<s><cl type="esclam" function="princ">Ahi quanto<cl type="rel"
function="subord I" id="IfI1" next="IfI2"> a dir <cl type="int"
function="subord II" id="IfI3" next="IfI4">qual era</cl></cl>
&grave; cosa dura<lb/> <cl type="rel" function="subord I"
id="IfI2"><cl type="int" function="subord II" id="IfI4">esta selva
selvaggia e aspra e forte<lb/> <cl type="rel" function="subord
III">che nel pensier rinova la paura!</cl></cl></cl></lb/>
```

This work of simplification has been carried out for the whole *Comedy*, thus reducing the original number of 227 different attributes in *DanteSearch*’s annotation to just 29 [9] and leaving to the user the task of discerning any further and significant syntactic distinctions.

Putting aside the few typos in the original code, no further changes have been performed on *DanteSearch* XML codification. This way, the theoretical basis behind the analysis have been accepted indiscriminately, even where the definition of the syntactic type was interpretable in different ways.

4. Features and functionalities

As stated above, the current version of *TRLARS* is not final: what is currently in use is a preliminary version of the tool, whose purpose is to prove whether the theoretical basis is solid or not, and to test the possibility to perform effective queries under syntactic and rhythmic restrictions at once. In spite of its temporariness, *TRLARS* already allows the users to explore its

features on the *Comedy*’s text, by reaching it at its homepage;⁸ to provide a more accurate idea of its functionalities, a description of its features is needed.

This is the main interface of *TRIARS*, from which it is possible to perform queries by filling out any of the available options:

Ricerca

Testo

Macrotipo sintattico

Livello di subordinazione

Ritmo

Opzioni di ricerca ritmica

Figure 1: *TRIARS*’ Home Page.

The interface allows many research options, both related to syntactic and rhythmic aspects of the literary works in the database. Of course, any research can be launched by fulfilling a single field; but, for every result, the system will still output all the relevant data for the other fields; alternatively, all the options can be selected to intertwine the rhythmic and syntactic aspect in a single research, thus making it ever more extensive. Even though both possibilities still stand on their own and can be useful in different situations, the purpose of *TRIARS* is to combine two different aspects of poetry’s language, so most of the queries would benefit from the fulfilment of more than one field. As a matter of fact, both syntactic and rhythmic research regarding Dante’s *Comedy* is already possible with *DanteSearch* and the *Archivio Metrico Italiano*, even though they are not exclusively focused on the third rhyme as a structure; on the other hand, *TRIARS* most peculiar side is to allow double research and to avoid the difficult and complex task to combine manually the results from two different databases.

Here’s a detailed description of all the different options, with examples of how to fill them:

1. The field ‘Testo’ allows the user to narrow the results down to a specific word or group of words that can be read in any single hendecasyllable of the database; so, it is

8 <http://triars.labdilef.it/>.

possible to type “vita” to find the verse of *Inf.* I 1 (“Nel mezzo del cammin di nostra *vita*”) and all the others where the word is mentioned, such as *Par.* XXVIII 1 (“Poscia che ’ncontro a la *vita* presente”), but it is also possible to type “cammin di nostra vita” and round the research down to the first verse of the poem only.

2. ‘Macrotipo sintattico’ limits the results to the only selected syntactic macro-type, such as “dichiarativa” or “finale”; so, for example, filling this field with the option “relativa”, *TRIARS* will find all the hendecasyllables with a relative clause in it, regardless of their functions in the sentence.
3. ‘Livello di subordinazione’, on the other hand, gives the user the possibility to further specify for the syntactic function of a main or subordinate clause, defining its level of coordination or subordination in the sentence; this last field can be also selected without specifying the ‘Macrotipo sintattico’, so it can be both used to search all the coordinated clauses in the poems or all the subordinates of a specific grade, but it can also be combined with the previous option to obtain, for example, all the 3rd grade relative clauses in the database.
4. The next option, ‘Ritmo’, requires the input of a series of digits, each separated by a blank space, that correspond to the ten positions of the hendecasyllable. As seen in the first picture, there are three different alternatives to be selected in the field ‘Opzioni di ricerca ritmica’:
 - 4.1. ‘Ricerca la presenza esatta degli ictus indicati’, it is the basic option, and it is automatically selected by the tool if no choice is selected by the user. It narrows the research down to all the hendecasyllables with the very same rhythmic sequence inputted by the user, so, when a sequence like 2 4 7 10 is researched, the tool returns all the tercets where there is at least one hendecasyllable with the rhythm marked up exactly as 2 4 7 10.
 - 4.2. ‘Ricerca la presenza almeno degli ictus indicati’, it allows more flexible queries, where at least the ictuses specified by the user must be present, but there can be others too in a single verse. This means that inputting a sequence like 4 6 8 10, the tool will find all the tercets where at least one hendecasyllable has ictuses on the positions 4 6 8 10, but also any other cases where other positions are accented, like 1 4 6 8 10, 2 4 6 8 10 or, if present in the database, even 1 2 3 4 5 6 7 8 9 10.
 - 4.3. ‘Ricerca la presenza esatta della serie di ictus indicata’, it is the most unusual option, and it allows to narrow all the results down to a specific sequence of ictuses in a portion of the hendecasyllable. It means that, when the digit 4 6 7 are inputted, the tool will find all the instances marked up as 4 6 7 10, but also any possible variations which would not interfere with the sequence 4 6 7, such as 1 4 6 7 10, 1 2 4 6 7 19, 1 3 4 6 7 8 10 and similar, though it would not find verses with a scheme like 4 7

10, 4 5 7 10 or 4 5 6 7 10. This option is not the most common choice, but it can be of use when it is needed to focus the user’s attention onto a specific portion, such as the first, the middle or the last part, of the hendecasyllable.

As obvious as it may sound from the description of the three options, *TRIARS*, according to the guidelines from the *Archivio Metrico Italiano*, does not include a limited number of hendecasyllable schemes, but it allows to input any rhythmic schemes, since no ictus’ sequence is considered to be impossible or wrong, in the metrical meaning of the word.

Here’s a short example of a research involving all the available options: if the sequence *ed elli* is inputted in the field ‘Testo’, selecting the categories of *dichiarativa* and the function of *principale* and specifying the ictuses at least on the 2, 4 and 10 positions, *TRIARS* is able to find 48 results of hendecasyllables that match all the requests. The tool then proceeds to open a new page where the total number of results is shown, followed by the single hendecasyllable characterized by all the specification, organized by their appearance order in the original works they have been taken from:

Risultati: 48

Inf, III, v. 13 Ed elli a me, come persona accorta:	2 4 5 8 10	dich princ	Vedi la terzina
Inf, III, v. 34 Ed elli a me: "Questo misero modo	2 4 5 7 10	dich princ, dich princ	Vedi la terzina
Inf, III, v. 76 Ed elli a me: "Le cose ti fier conte	2 4 6 10	dich princ, dich princ	Vedi la terzina
Inf, IV, v. 19 Ed elli a me: "L'angoscia de le genti	2 4 6 10	dich princ, dich princ	Vedi la terzina
Inf, V, v. 76 Ed elli a me: "Vedrai quando saranno	2 4 6 7 10	dich princ	Vedi la terzina

Figure 2: an example of how results are shown in *TRIARS*

At this point, the user can select the button ‘Vedi la terzina’ next to any of the results, and the tercets in which it is located will be entirely shown, complete with the rhythm of the other two related verses:

Risultati: 48

Inf, III, v. 13		
Ed elli a me, come persona accorta:	2 4 5 8 10	dich princ
Vedi la terzina		
13 Ed elli a me, come persona accorta:	2 4 5 8 10	dich princ
14 "Qui si convien lasciare ogni sospetto;	1 4 6 7 10	
15 ogni viltà convien che qui sia morta.	1 4 6 8 10	

Figure 3: the single results can be expanded to show the entire tercet.

This is the current level of development of *TRIARS*, which is enough to perform preliminary researches, ones that may even be as articulated as the example above, though only related to Dante's *Comedy*. However, as it may have been evident even from a short example like the one above, there are still many issues with the tool and a lot of work still has to be done to make it work properly. In the next two Sections the major issues of *TRIARS* will be described and discussed, with a paragraph focused on the necessary improvements that have to be taken to fix them.

5. Major flaws of the current version

In this Section the most important flaws of the current version of *TRIARS* will be considered from a technical point of view. Of course, observations about the technical issues of this tool do not end the debate about the critical points of this research, points that can be of a theoretical nature as well; nonetheless, the main focus of this paper is technical and those theoretical aspects – such as the insurmountable problem of a univocal definition of the syntactic type of a clause or the rhythmic profile of a verse – have already been touched upon other occasions ([8]; [9]; [10]). So, having set aside all these theoretical weaknesses, the technical issues that *TRIARS* incurred in during its development phase have to be examined, which can be rounded down to three major types of issues, all three connected to the very same original cause.

The first and easiest to detect flaw has to do with the accuracy with which *TRIARS* highlights the relevant clauses given as a result of any query. This issue is evident even when it comes to a clause contained in a small part of a single hendecasyllable, as in the case of *Inf*, I 4, where the interrogative clause *qual era* is not highlighted as to produce an easy distinction from the other phrasal elements in the verse or, when the user opens it, in the tercet:

Inf, I, v. 4	Ahi quanto a dir qual era è cosa dura	2 4 6 8 10	int subord II	Vedi la terzina
4	Ahi quanto a dir qual era è cosa dura	2 4 6 8 10	int subord II	
5	esta selva selvaggia e aspra e forte	1 3 6 8 10		
6	che nel pensier rinova la paura!	4 6 10		

Figure 4: an example of *TRIARS*' inaccuracy in showing the clauses in the verses.

This means that the only way to distinguish this kind of clauses is for the readers to carefully read them themselves, that is to autonomously analyse the sentence to understand its structure and the function of the interested clause in it. This apparently easy task is bound to become more and more difficult in all the cases where a single clause occupies more than one hendecasyllable or more than one tercet. On a closer inspection, this second kind of issue is more critical than it appears: every single verse in which the clause is settled is counted as a single result (at least it is without any exception when no rhythmic scheme is specified in the research), thus making the results count unreliable. An example of this issue is the case of the relative clause *che con lena affannata [...] si volge all'acqua perigliosa* in *Inf.*, I 22-4:

Inf, I, v. 22	E come quei che con lena affannata	2 4 7 10	rel subord I	Vedi la terzina
22	E come quei che con lena affannata	2 4 7 10	rel subord I	
23	uscito fuor del pelago a la riva	2 4 6 10		
24	si volge a l'acqua perigliosa e guata,	2 4 8 10		
Inf, I, v. 23	uscito fuor del pelago a la riva	2 4 6 10	rel subord I	Vedi la terzina
22	E come quei che con lena affannata	2 4 7 10		
23	uscito fuor del pelago a la riva	2 4 6 10	rel subord I	
24	si volge a l'acqua perigliosa e guata,	2 4 8 10		
Inf, I, v. 24	si volge a l'acqua perigliosa e guata,	2 4 8 10	rel subord I, rel coord I	Vedi la terzina
22	E come quei che con lena affannata	2 4 7 10		
23	uscito fuor del pelago a la riva	2 4 6 10		
24	si volge a l'acqua perigliosa e guata,	2 4 8 10	rel subord I, rel coord I	

Figure 5: an example of *TRIARS*' inaccuracy in showing the clauses in the tercets.

As this particular case shows, in the final output of the results there is a combination of different issues. To be specific, the major issues here are relevant to *Inf.*, I 24, where the tool points out the presence of one ‘rel subord I’ and one ‘rel coord I’, meaning that next to the (partial) relative clause *si volge a l’acqua perigliosa* there is the coordinated relative clause *e guata*, but they are not properly highlighted and separated by *TRIARS*, and to *Inf.*, I 23, where the temporal clause *uscito fuor dal pelago a la riva* is wrongly showed as a relative clause. Finally, a similar inaccuracy can be noticed in all the cases in which, contrarily to the example of *Inf.*, I 22-4, a clause is not included in the results for a researched syntactic type:

The screenshot displays three search results for syntactic analysis in *TRIARS*. Each result is presented in a box with a header and a table of results.

Example 1: Inf, I, v. 1
 Search: Nel mezzo del cammin di nostra vita (ID: 2 6 8 10, Type: dich princ)
 Results:
 1 Nel mezzo del cammin di nostra vita (ID: 2 6 8 10, Type: dich princ)
 2 mi ritrovai per una selva oscura (ID: 4 8 10)
 3 ché la diritta via era smarrita. (ID: 4 6 7 10)

Example 2: Inf, I, v. 2
 Search: mi ritrovai per una selva oscura (ID: 4 8 10, Type: dich princ)
 Results:
 1 Nel mezzo del cammin di nostra vita (ID: 2 6 8 10)
 2 mi ritrovai per una selva oscura (ID: 4 8 10, Type: dich princ)
 3 ché la diritta via era smarrita. (ID: 4 6 7 10)

Example 3: Inf, I, v. 3
 Search: ché la diritta via era smarrita. (ID: 4 6 7 10, Type: dich princ)
 Results:
 1 Nel mezzo del cammin di nostra vita (ID: 2 6 8 10)
 2 mi ritrovai per una selva oscura (ID: 4 8 10)
 3 ché la diritta via era smarrita. (ID: 4 6 7 10, Type: dich princ)

Figure 6: an example of a wrong output in *TRIARS*' results for syntactic analysis.

In *Inf.*, I 1-3 the causal clause *ché la diritta via era smarrita* is considered to be both a declarative sentence and, when a causal clause is researched, properly as a causal clause.

The problem with all these three kinds of issues is in the incompatibility between the XML encoding used by *DanteSearch* and *TRIARS*' search engine. Taking the last tercet as an example, it can be read as follows in *DanteSearch*'s annotation:

```
<s><cl type="dich" function="princ">Nel mezzo del cammin di nostra
vita<lb/> mi ritrovai per una selva oscura,<lb/> <cl type="caus"
function="subord I">ché; la diritta via era
smarrita.</cl></cl><lb/> </s>
```

It is evident that the end-tags of all the clauses, `</cl>`, are placed right before the end-tag of the sentence, `</s>`; this means that for *TRIARS* the start-tag of the declarative clause, *Nel mezzo del cammin di nostra vita / mi ritrovai per una selva oscura*, is only closed after the subordinate clause *ché la diritta via era smarrita*. The code, to be rightly interpreted by *TRIARS*, should be manually redefined adjusting the XML indexes, and pointing out, case by case, the exact boundaries between each clause.

It seems obvious that such an adjusting process for a poem like Dante's *Comedy* would require an incredible amount of working time, which, when the issue was firstly discovered with *TRIARS*, was not in line with the standard duration of a Ph.D.; at that time, the only viable possibility was a severe reduction of the database, confining it to a small portion of the *Comedy*, and to have it marked up in a more appropriate way for *TRIARS*. Anyway, this option did not seem suitable: since the main point of the project was to prove the possibility of a rhythmic and syntactic analysis, a large amount of data was necessary to do the preliminary research, which would not have been equally satisfactory on a small portion of the poem. For this reason, despite the awareness of the inaccuracy of several results and of the need of a manual analysis to confirm them, it has been opted to include these issues in the first version of *TRIARS*, thus postponing the debugging phase to a later point after the verification of the theories behind the entire project.

6. Possible improvements

Knowing all these issues, the research has nonetheless been conducted with the help of *TRIARS* and a later double check of all the results on the original sources of its database, *DanteSearch* and *Archivio Metrico Italiano*; accordingly, all the results have been manually reviewed and organized in a more readable way and discussed [9].

At this point it seems necessary to find a way to transform *TRIARS* into a proper working tool, and, later, to extend its database to include other poems in third rhyme. The first and most urgent improvements are related to the flaws that have been presented in the previous section; after this essential revision, several side aspects can benefit of the following minor improvements:

- 1) An important feature to allow more accurate queries would be the addition of a filter through which the user can decide which verses of the tercet have to be taken into account in every research, so that it would be possible to focus the attention, when needed, on a specific part of the third rhyme;
- 2) Another feature, which is already implemented on the *Archivio Metrico Italiano*, is to add a fourth option for the rhythmic research, giving the possibility to exclude verses with a specific ictus on any selected positions;
- 3) A third option, is to offer a text research limited to a specific sequence of alphabetical characters, instead of full words; for instance, this would allow to ease up the searching

process to find specific grammatical categories, such as adverbs of manner in *-mente* or gerund forms of verbs;

- 4) Another flaw of *TRLARS* is that it does not discern the indirect interrogative clauses from the direct ones, since they are both coded as ‘int’; of course, when the function of any interrogative in the sentence is specified, it is impossible to mix them up, but this often leads to some confusion and it needs to be fixed with an appropriate rewording, such as ‘inti’ and ‘intd’ for the two different types;
- 5) It would be necessary to offer the user the possibility to visualize the previous or next verses (with the relative rhythmic and syntactic analysis) when a result is expanded to show the whole tercet, since it may be crucial when a single clause is dislocated on many tercets;
- 6) An increase in the output legibility, with a more appropriate representation of the results with different styles in the instance of two results of the same type in a single hendecasyllable or tercet, is another important feature to be added;
- 7) Finally, the addition of Boolean operators is an expected improvement too, one that is already implemented on *DanteSearch*.

Besides these minor improvements, there is another option that has to be carefully taken into account, one whose implementation requires a slight reconsideration of the project’s theoretical basis. One of the most interesting prospects of *TRLARS* – something that emerged both from the first surveys on the *Comedy* and by scrutinising the most satisfactory results – is the introduction of the concept of intonational syntagm in the stylistic analysis of the third rhyme. As a matter of fact, some of the most interesting results on the *Comedy*’s tercets ([8]; [10]) show important similarities also when we consider them from the perspective of intonational syntagms, where the different intonational profiles seem to be comparable from one tercet to another. Indeed, this concept should be added beside the syntactic and rhythmic analysis, it should be intertwined with them when needed, but more research must be made in order to define an appropriate way of analysing and then encoding this new amount of data.

7. A further step

The key aspect of *TRLARS* is the extension of the database to other poems in third rhyme, and this can be done only after a final and more stable version of the tool is achieved. The importance of a comparative study of different poems in third rhyme has already been raised in other occasions; among the most recent and significant ones, Corrado Calenda [7] has strongly suggested to extend stylistic analysis to “altri minori affluenti”, pointing out the need of an “acquisizione di dati oggettivi fondati su rilevamenti tecnici”, as the one proposed by *TRLARS*.

Of course, due to the numerous issues that arose during the early stages of the research, no effort has still been made towards this direction, which is, however, essential to the whole project: as a matter of fact, the tool is designed for the analysis of the third rhyme and not only

to Dante’s *Comedy*. This is a crucial aspect of the research, because the general theory, which has already produced results of scientific interest regarding the *Comedy*, now has to be extended and tested on other poems to be truly productive. Starting from Dante’s poem and moving forward to other third rhyme works, the final objective is to provide a tool to help achieving a better understanding of the presence of Dante’s influence in different authors and, more generally, of the use of this meter in Italian literature throughout a comparison of its main occurrences.

To show a brief example of the attended results, one interesting case, very infrequent also in the *Comedy*, is when a consecutive clause introduced by *si che* is both in the last verse of a tercet and interrupted by an incidental clause. As already mentioned, Dante’s *Comedy* itself has no more than the following four examples:

Inf., XXI 52-4

Poi l’addentar con più di cento raffi,	1 4 6 8 10
disser: “Covertò convien che qui balli,	1 4 7 9 10
<i>si che</i> , se puoi, <i>nascosamente accaffi</i> ”.	1 4 8 10

Inf., XXII 70-2

E Libicocco “Troppo avem sofferto”,	4 6 8 10
disse; e preseli ’l braccio col runciglio,	1 3 6 10
<i>si che</i> , stracciando, <i>ne portò un lacerto</i> .	1 4 8 10

Inf., XXXII 88-90

“Or tu chi se’ che vai per l’Antenora,	2 4 6 10
percotendo”, rispuose, “altrui le gote,	3 6 8 10
<i>si che</i> , se fossi vivo, <i>troppo fora?</i> ”.	2 4 6 8 10

Par., III 61-3

però non fui a rimembrar festino;	2 4 8 10
ma or m’aiuta ciò che tu mi dici,	2 4 6 10
<i>si che</i> raffigurar <i>m’è più latino</i> .	1 6 10

An observation of these results tells us that the cases are not equally distributed in the poem, to the point that no example of this pattern can be found in the *Purgatorio*. Another interesting

point to be highlighted, which has to be directly connected to the syntactic use of the consecutive clauses, is that all four tercets end with a strong syntactic pause, one that coincides with the metrical limit of the third rhyme; moreover, with the exception of *Inf.*, XXII 88-90, all the tercets are part of a direct speech of one of the characters, which is introduced by a *verbum dicendi*, with one of the typical formulas often used by Dante in the poem in such cases and well analysed in [11]. Finally, from a rhythmic point of view, it is interesting to see that all the incidental clauses start from the third position of the verse, just after the *si che* and that they adopt several different rhythmic solutions, since just two out of four show the same 1 4 8 10 scheme, even though their intonational profiles are not dissimilar.

A closer look to other third rhyme works from the Italian fourteenth century literature reveals that such identical structures can be even found in shorter poems, like the *Caccia di Diana* by Giovanni Boccaccio and the *Trionfi* by Francesco Petrarca, which show one example each, in the following cases:

Caccia di Diana, XV, 28-30

per che lasciar l'astore allor le piacque,	4 6 8 10
il qual, montato, uno ne ferio,	2 4 6 10
<i>si che 'n sull'erba morendo si giacque,</i>	1 4 7 10

Triumphus Eternitatis, vv. 28-30

E le tre parti sue vidi ristrecte	4 6 7 10
ad una sola, e quella una esser ferma	4 7 8 10
<i>si che, come solea, più non s'affrette;</i>	1 6 7 10

The interesting aspect to highlight, and that should represent the main focus of the whole research as it has been intended, is related to the differences between Dante's style and other authors', evident even in the context of the analysis of such a short sample. Indeed, in these two cases it is clear how Boccaccio postpones the incidental clause to a later part of the hendecasyllable, even though he is still using the gerund, as Dante does in *Inf.*, XXII 70-2, and it is also evident how there is a rhythmic distance from the *Comedy*, since in both cases we have an ictus on the seventh position of the last hendecasyllables that is never accented in Dante's verses.

Of course, this is not enough data to make any definite assumption on these poets' styles and more research is to be expected to have a clearer idea of Dante's debt in the early Italian third rhyme. Even if we are not suggesting that both Boccaccio and Petrarca had had in mind these very tercets written by Dante when they composed their own verses, though, the main point is to see how similar structures can occur in different authors and how they can be compared in both their differences and similarities. The fact that even such a short sample can inspire

stylistic considerations shows how the extension of the database to other poems in third rhyme can be essential for a better understanding on one of the most successful meters of Italian literature.

8. Conclusions

Besides these first considerations on the third rhyme, the general idea one gets from these early observations on the *Comedy* is that such a study is possible and useful as to have a clearer vision of Dante's style. The most evident aspect is the coincidence in the utilisation of the same type of clauses in different contexts: as it may seem obvious, in the narrow space of three verses and in the metrical strictness of the hendecasyllable, there must be some recurrence in the sentences' structure, recurrences that are often stimulated by the poet's memory.

To provide a quick example of these coincidences, that, as it is important to note, are also related to the sentences' intonational profile, it is possible to notice an identical structure of the declarative sentences in the following tercets:

Inf., XXVIII 133-5

<i>E</i> perché tu di me novella porti,	3 4 6 8 10
<i>sappi</i> ch'ì son Bertram dal Bornio, quelli	1 4 6 8 10
che diedi al re giovane i ma' conforti.	2 4 5 10

Inf., XXXII 67-9

<i>E</i> perché non mi metti in più sermoni,	3 6 10
<i>sappi</i> ch'ì fu' il Camiscion de' Pazzi;	1 4 8 10
e aspetto Carlin che mi scagioni".	3 6 10

Moreover, interesting similarities can also be highlighted in the structure of the following occurrences of temporal clauses:

Inf., XXV 70-2

Già eran li due capi un divenuti,	2 6 7 10
<i>quando n'apparver due figure miste</i>	1 4 8 10
<i>in una faccia</i> , ov'eran due perduti.	4 6 8 10

Purg., XIX 25-7

Ancor non era sua bocca richiusa,	2 4 7 10
<i>quand'una donna apparve santa e presta</i>	4 6 8 10
<i>lunghezzo me per far colei confusa.</i>	2 4 8 10

More relevant coincidences though can also be found in the following relative clauses in rhyme, which even share the same rhyme word *resplende/risplende* in the last verse:

Par., XV 19-21

tale dal corno <i>che 'n destro si stende</i>	1 4 7 10
a piè di quella croce corse un astro	2 4 6 8 10
de la costellazion <i>che li resplende;</i>	1 4 6 8 10

Par., XX 4-6

lo ciel, <i>che sol di lui prima s'accende,</i>	2 4 6 7 10
subitamente si rifa parvente	4 8 10
per molte luci, <i>in che una risplende;</i>	2 4 7 10

This is just a quick summary of what results are to be expected with *TRIARS*, but it already shows how cases like these corroborate the idea that any syntactic type has its own preference when it comes to the usage of the structure of the “piccolo ritmo” of the hendecasyllable and, more importantly, of the “grande ritmo” of the third rhyme. These results seem to highlight what appear to be some composing habits that, even a poet like Dante, often resorts to: rhythm and syntax are both heavily influenced by these habits, and they are often structured in comparable ways that can tell us a lot about an author’s style.

References

- [1] Alighieri, Dante, “Commedia.” In Petrocchi, Giorgio (Ed.), *La Commedia secondo l'antica vulgata*, Firenze: Le Lettere, 1994.
- [2] Beccaria, Gian Luigi. 1975 (1970). “L'autonomia del significante. Figure dantesche.” In *L'autonomia del significante. Figure del ritmo e della sintassi. Dante, Pascoli, D'annunzio* (pp. 114-135), Torino: Einaudi.

- [3] Beltrami, Pietro. 2015 (1975). “Primi appunti sull’arte del verso nella Divina Commedia.” In *L’esperienza del verso. Scritti di metrica italiana* (pp. 5-32), Bologna: Il Mulino.
- [4] Beltrami, Pietro. 2015 (1976). “Sul metro della Divina Commedia: sondaggi per un’analisi sintattica.” In *L’esperienza del verso. Scritti di metrica italiana* (pp. 34-70), Bologna: Il Mulino.
- [5] Bertinetto, Pier Marco. 1973. “Ritmo e modelli ritmici. Analisi computazionale delle funzioni periodiche nella versificazione dantesca.” Torino: Rosenberg & Sellier.
- [6] Boccaccio, Giovanni. “La Caccia di Diana.” In Iocca, Irene (Ed.), *La Caccia di Diana*, Roma: Salerno Editrice, 2016.
- [7] Calenda, Corrado. 2014. “La terza rima tra Dante e Boccaccio.” In Azzetta, Luca, and Andrea Mazzucchi (Eds.). *Boccaccio editore e interprete di Dante. Atti del convegno internazionale di Roma, 28-30 ottobre 2013* (pp. 293-308), Roma: Salerno Editrice.
- [8] Conti, Ugo. 2020. “Prime osservazioni per un’analisi ritmico-sintattica della Commedia.” In Facini, Laura, and Jacopo Galavotti, and Arnaldo Soldani (Eds.). *Nuove prospettive sulla terza rima. Da Dante al Duemila* (pp. 29-48), Padova: libreriauniversitaria.it.
- [9] Conti, Ugo. 2021. “Per uno studio ritmico-sintattico della terza rima: prime osservazioni sulla Commedia dantesca.” PhD diss., Università di Firenze.
- [10] Conti, Ugo. 2022. “Per un’analisi ritmico-sintattica della terza rima: il caso delle dichiarative e delle interrogative dirette nella Commedia.” *Stilistica e Metrica Italiana* 22: 3-40.
- [11] Contini, Gianfranco. 1970 (1965). “Un’interpretazione di Dante.” In *Un’idea di Dante* (pp. 69-111), Torino: Einaudi.
- [12] Contini, Gianfranco. 2021 (1970). “An idea of Dante.” Translated by Stephen Sartarelli. New York: Agincourt Press.
- [13] Gigli, Sara. 2004. “Codifica sintattica della Commedia dantesca.” PhD diss., Università di Pisa.
- [14] Petrarca, Francesco. “Trionfi.” In Pacca, Vinicio, and Laura Paolino (Eds.), *Trionfi*, Milano: Mondadori, 1996.
- [15] Praloran, Marco, and Marco Tizi. 1988. “Narrare in ottave. Metrica e stile dell’innamorato.” Pisa: Nistri-Lischi.
- [16] Praloran, Marco (Ed.). 2004. “La metrica dei Fragmenta.” Roma-Padova: Antenore.

- [17] Praloran, Marco, and Arnaldo Soldani. 2003. “Teoria e modelli di scansione.” In Praloran, Marco (Ed.). *La metrica dei Fragmenta* (pp. 3-123), Roma-Padova: Antenore.
- [18] Robey, David. (2000). “Sound and structure in the Divine Comedy.” Oxford: Oxford University Press.
- [19] TEI Consortium, “TEI P5: Guidelines for Electronic Text Encoding and Interchange.” *Text Encoding Initiative*, 2022-04-19. <https://tei-c.org/Guidelines/P5/>.