Quantifying Greek Rhyme

ISTAL 23

April 2 2017

Nina Topintzi
Konstantinos Avdelidis @ Aristotle University of Thessaloniki
Theodora Valkanou
What is (poetic) rhyme?

- **Rhyme**: sound correspondence between one or more syllables at – usually – the ends of poetic lines, e.g. krína - elafína

- Although related, poetic rhyme and phonological rime are different. We’ll be mostly talking about the former here
Why rhyme?

- Rather under-studied in comparison to other metrical components (Köhnlein & van Oostendorp 2014)
  - In Greek, virtually unexplored [but see Κοκόλης (1993), for a philological study]
- It’s phonologically interesting too!
  - Patterns commonly attested in various poetic traditions correspond to patterns languages employ in their general phonological systems (Fabb 2010)
  - Rhyme is reminiscent of phonological reduplication
This project

- Funded by AUTH Research Committee: grant to first author
- Also with the support of The Centre for the Greek Language (ΚΕΓ), and especially Dr. Vasilis Vasileiadis
- Construction of a pilot database with a sample of rhymes as they appear in the poetic works of diverse Greek poets, including Karyotakis, Palamas, Solomos, Valaoritis, Varnalis among others
- Database URL: http://greek-rhyme.web.auth.gr/index.php/home
  - Not fully accessible to the public yet!
Project initiatives

- Online repository of rhymes
  - Web site
  - Library of poems
  - Provision of statistics

- Theoretical rhyme analysis

- Database population tools
  - Analyst expert knowledge integration productivity GUI (graphical user interface)
  - Rhyme detection/classification algorithm

- Expandability
  - Independent application and rules expression
  - Future goals
Our aims

- Descriptive: to better understand rhyme in Greek → which patterns are found and which are common
- Theoretical and/or typological
  - how does Greek rhyme fit in the typology of rhyme in general? For instance, Holtman (1996: 32) suggests – based on Middle English – that languages with rich inflectional morphology prefer feminine over masculine rhymes. Does Greek corroborate this claim?
  - how common are certain patterns, such as rich or imperfect rhyme?
Outline of the talk

- Methodological Issues
  - Rhyme Classification & Basic Patterns
  - Algorithms & Meta-information
- Concrete examples of RPs
- Phonological Implications
- Further issues and future work
Rhyme Patterns & Algorithms
Patterns to identify

Three main distinctions

- **Rhyme Type**: Feminine (=non-final) or Masculine (=final)
  - Feminine-3σ (προπαροξύτονη): πέρασα - γέρασα / Feminine-2σ (παροξύτονη): ελάφι - χωράφι
  - Masculine (οξύτονη): κοινός - καλός
    - Mosaic refers to rhymes that span word boundaries, e.g. δώς μου - φώς μου
- **Rich**: when the onset(s) of the stressed syllable match across rhyme pair (RP)
  - For some poets, e.g. Solomos, it is common to get rich rhymes, when the directly previous vowel matches too
- **Imperfect**: a vowel or consonant within the rhyme differs across RP
  - αγέρι - λογάρι: vowel differs
  - ξαφνίζει - τεχνίτη: consonant differs
Initial Rhyme Classification chart

- Line Rhyme
  - TYPE
    - Feminine-2σ
      - Mosaic
        - YES
    - Feminine-3σ
      - Mosaic
        - YES
    - Masculine-1σ
      - Mosaic
        - YES
  - RICH
    - YES
    - NO
  - IMPERFECT
    - YES
    - NO

Consonants
- Previous vowel is same?
  - YES
  - NO

Which
- Which (POA-MOA-Voice)?
  - Which?
Actual Rhyme Classification chart

- **Masculine**
  - Rich Consonant variants

- **Feminine 2s**
  - Imperfect Other V/C variants
  - Pre-rhyme identical V

- **Feminine 3s**
  - Mosaic
  - Copy

- **Line Rhyme**

  - Across words

Only for Feminine 2s and Feminine 3s
Algorithm outline

Three step process

- Poem pre-processing
  - Rule-based syllabification
  - Rule-based orthographic to phonetic transcription (use of SAMPA; see Appendix for SAMPA-IPA correspondence)
  - Per-line Synchronous multi layered representation (word, syllable, cluster, phoneme)

- Line analysis
  - Standardized syntax of hierarchical rhyme detection rules
  - Rule-based line pair rhyme detection

- Rhyme post-processing
  - Line pair characterization (πλεχτή, ζευγαρωτή, κτλ.)
  - Database wide statistics
Multi layered representation

<table>
<thead>
<tr>
<th>Line</th>
<th>Word</th>
<th>Syllable</th>
<th>Cluster</th>
<th>Phoneme</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>'Cere</td>
<td>'o</td>
<td>'Cere</td>
<td>elefTe'ri\a</td>
</tr>
<tr>
<td></td>
<td>Ce re</td>
<td>'o</td>
<td>Ce re</td>
<td>Te 'ri\a</td>
</tr>
<tr>
<td></td>
<td>C e r e o</td>
<td>e l e f</td>
<td>C e r e e l e f</td>
<td>Te r 'ri\a</td>
</tr>
<tr>
<td></td>
<td>C e r e o</td>
<td>C e r e e l e f</td>
<td>C e r e e l e f</td>
<td>Te r 'ri\a</td>
</tr>
</tbody>
</table>

Index: 17 16 15 14 13 12 11 10 9 8 7 6 5 4 3 2 1 0

Representation node 2:

Word: ortho $\rightarrow$ ελευθεριά, phono $\rightarrow$ elefTe'ri\a
Syllable: ortho $\rightarrow$ ριά, phono $\rightarrow$ 'ri\a, stressed $\rightarrow$ yes
Cluster: ortho $\rightarrow$ ρι, phono $\rightarrow$ 'ri\, type $\rightarrow$ consonant
Phoneme: ortho $\rightarrow$ ρ, phono $\rightarrow$ r

Layer relative indices
Rhyme detection rules

- Queues of successive comparison steps of respective line representation nodes in reverse order
- Node level property in index position operants → Operator condition syntax
- Suitable repertoire of comparison operators
- Hierarchy support by rules inheritance
- Match if all queue comparison steps are true

Comparison step definition

$1: \text{syllable.r\_index.line} \rightarrow 0$
$\text{cluster.type} \rightarrow V$
$\text{cluster.phono} \rightarrow \text{op:eq}$
$\text{cluster.f\_index.syllable} \rightarrow \text{op:and:lte:1}$

Interpretation

$1$: Each line last syllable
cluster types are vowels and are phonemically same and are both first or second in syllable,

Example:
Step match: εντροπαλή, πεί
Step mismatch: εντροπαλή, ακαρτερούσες
Some concrete examples
- F2 = Feminine rhyme on penult
- M = Masculine/final rhyme

29 8/1 Τότε εσήκωνες το βλέμμα
31 8/3 και εις το ρούχο σου έστας' αίμα,
F2
30 8/2 μες στα κλάιματα θολό,
32 8/4 πλήθος αίμα Ελληνικό.
M
31 8/3 και εις το ρούχο σου έστας' αίμα,
32 8/4 πλήθος αίμα Ελληνικό.

(1) 'to-te e-'si-ko-nes 'to 'vle-ma
(1) 'mes 'sta 'klai-ma-ta To-'lo
(0) 'ce 'is 'to 'ru-xo 'su 'e-sta-'ze-ma
(0) 'pli-Tos 'e-ma e-li-ni-'ko
- **M (TR-S, IDV)**
  - **M** = masculine/final rhyme
  - **TR-S** = singleton rich rhyme
  - **IDV** = pre-rhymal V is identical

```
1  1/1  Στο περιγιάλι το κρυφό       (1) 'sto pe-ri-\'j\a-li 'to kri-'fo
4  1/4  μα το νερό γλυφό.       M
       M ( TR-S, IDV )
2  1/2  κι άσπρο σαν περιστέρι       (1) 'ci 'a-spro 'san pe-ri-'ste-ri
3  1/3  διψάσαμε το μεσημέρι       (0) Di-'psa-sa-me 'to me-si-'me-ri
4  1/4  μα το νερό γλυφό.       (0) 'ma 'to ne-'ro Gli-'fo
```
- **TR-S** = singleton rich rhyme
- **PR-CC2** = partial rich rhyme with clusters; C2 is same

---

5  2/1  Αγάλματα θεών ζωντανεμένα

8  2/4  Τα στήθια θα χαρούν τα πονεμένα.
        F2 (TR-S, IDV)

6  2/2  θ' αγναντέψουν στη Νίμπρο εκεί την πρώτη

7  2/3  της λεφτερίας αστραφτερή λαμπρότη.
        F2 (PR-CC2)

7  2/3  της λεφτερίας αστραφτερή λαμπρότη.

8  2/4  Τα στήθια θα χαρούν τα πονεμένα.

---

(1) a-'Gal-ma-ta Te-'on zo-da-ne-'me-na

F2

(1) T' a-Gna-'de-psun 'stli 'ni-bro e-'ci 'tin 'pro-li

F2

(0) 'tis le-fte-'rj'as a-strf-te-'ri la-'bro-li

(0) 'la 'sli-TCa 'Ta xa-'run 'la po-ne-'me-na

---

**Mavilis – Excelsior!**
- F2 = feminine rhyme on penult
- IDV (IDV-2W) = pre-rhymal V is identical and found in previous word

5 2/1 Βαρύ μέσ’ στο[ν] πρασινισμένο δρόμο, (1) va-'ri 'me-ston pra-si-ni-'zme-no 'Dro-mo
8 2/4 σαν του Χριστού τον σταυρωμένον ώμο. F2
   F2 ( IDV ( IDV-2W ) )
6 2/2 ποι εγώ μονάχος ἔχω καθαρίσει, (1) 'pu e-'Go mo-'na-xos 'e-xo ka-Ta-'ri-si
7 2/3 το δάκρυ το χορτάρι ἔχει λυγίσει F2
7 2/3 το δάκρυ το χορτάρι ἔχει λυγίσει (0) 'to 'Da-kri 'to xor-'ta-ri 'e-Ci li-]i-si
8 2/4 σαν του Χριστού τον σταυρωμένον ώμο. (0) 'san 'tu xri-]stu 'ton sta-vro-'me-non 'o-mo

V-identity
again

Varnalis - Proi
- **IMP-C** = within the RP, onset Cs differ, viz. [r] vs. [ð]
- **IDV** = pre-rhymal V is identical

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2/5</td>
<td>και γιομίστη μαξιλάρες</td>
<td>(1) 'ce jlo-'mi-ste ma-ksi-'la-res</td>
</tr>
<tr>
<td>11</td>
<td>2/6</td>
<td>να πλαγιάσουν οι κυράδες</td>
<td>F2 (IMP-C, IDV)</td>
</tr>
<tr>
<td>11</td>
<td>2/6</td>
<td>να πλαγιάσουν οι κυράδες</td>
<td>(0) 'na pla-'jla-sun 'i ci-'ra-Des</td>
</tr>
</tbody>
</table>

**Valaoritis – Prolegomena (i Kyra Frosini)**
- F3 (IMP-C, IMP-V, PR-C2, IDV (IDV-2W))
  - F3 = feminine rhyme on antepenult
  - IMP-C = C within RP differs, viz. [k] vs. [n]
  - IMP-V = stressed V of RP differs, viz. [e] vs. [i]
  - PR-C2 = partially rich rhyme; one half of RP has cluster, the other has singleton; they agree on C2
  - IDV (IDV-2W) = pre-rhymal V is identical and found in previous word

Complex cases

Palamas – o Dodekalogos tou Gyftou (Logos A’)

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>54</td>
<td>4/19</td>
<td>Καὶ ἦταν ὡς νὰ πλέκονταν</td>
</tr>
<tr>
<td>55</td>
<td>4/20</td>
<td>καὶ ἦταν ὡς νὰ λύνονταν</td>
</tr>
</tbody>
</table>

F3 (IMP-C, IMP-V, PR-C2, IDV (IDV-2W))

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>55</td>
<td>4/20</td>
<td>καὶ ἦταν ὡς νὰ λύνονταν</td>
</tr>
</tbody>
</table>

(1) 'ce 'i-tan 'os 'na 'ple-ko-dan

(0) 'ce 'i-tan 'os 'na 'li-no-dan
Phonological Implications
Source

- The following observations are primarily based on
  - Σολωμός – Ύμνος εις την Ελευθερίαν (abbreviated as Sol. + Line No)
  - The poem is written in verses of 4 lines each, that alternate between 8 and 7 syllables, i.e. Verse1: [8-7-8-7], Verse 2: [8-7-8-7], ...
  - This rather rigid pattern allows us to form some generalizations with relative certainty
On hiatus

- In current SMG, many [ia] sequences present a single form with either hiatus [i.a] or its avoidance with a glide [ja] and subsequent glide strengthening (see Baltazani et al. 2016) based on whether they receive the traditional label «λόγιο» and «μη λόγιο», respectively, thus [anisićia] for ανησυχία with hiatus and [kormiā] for κορμία with hiatus avoidance.

- Solomos often contravenes this distinction and accepts alternations much more freely!

- For instance:
  - κορμία – κορμία
  - ελευθερία - ελευθερία

- Similarly, although we don’t have the full pairs, we get forms like ανησυχία (instead of ανησυχία, Sol. 558), επιθύμια (instead of επιθυμία, Sol. 517)
On /ria/ vs. /rja/ sequences (1)

- Solomos provides evidence that he accepts the onset sequence [rja] even initially
  - Pay attention to line 416

  413 104/1 Την αισθάνονται, και αφρίζουν  
  414 104/2 τα νερά, καὶ τ’ αγρικώ  
  415 104/3 δυνατά να μουρμουρίζουν  
  416 104/4 σαν να ρυάζετο θηριό.  

  (1) 'tin e-'sTa-no-de 'ce a-'fri-zun  
  (1) 'ta ne-'ra 'ce ta-Gri-'ko  
  (0) Di-na-'ta 'na mur-mu-'ri-zun  
  (0) 'san 'na 'rjá-ze-to Ti-'rjó

- ρυάζετο needs to be syllabified as [rjá.ze.to] with an onset cluster. The alternative [ri.á.zeto] is impossible, because it would render the line octasyllabic, while it must be heptasyllabic because of metrical restrictions
On /ria/ vs. /rja/ sequences (2)

- Focus on line 580 of verse 145

  - Can be transcribed as: (a) i’se 'Da-kri-a Tli-ve-ra or (b) i’se 'Da-kri\a Tli-ve-ra
  - Line 580 is in a position that must be 7σ-long \(\rightarrow\) (b) must be correct! (a) is wrong
  - But ['ďa.kri.a] is what Soultatis (2013: 276) argues for, as a sole possibility
  - Interestingly, he uses the alleged lack of [Crj] sequences as one of his basic arguments against Topintzi (2011) and the existence of underlying glides

- Upshot: Poetry may put the validity of phonological proposals to the test!
On palatalization (1)

- Focus on lines 189 and 191 of verse 48

- The words 'ίσκιοι - νεανίσκοι form a rhyme pair

- Since Solomos considers it a rhyme, then 'ίσκιοι should be: [i-sci] and not [i-sci-i]. Syllable counting of line suggests the same (189, 191 = 8σ; 190, 192 = 7σ)

- The comparison between pairs like [i-sci] and [nea-ni-sci] is phonologically telling with respect to palatalization
On palatalization (2)

- Baltazani & Topintzi (2012) distinguish between 3 types of PAL(atalization)
  - Simple PAL (triggered by front Vs) as in /ke'ros/ \(\rightarrow\) [ce'ros] ‘weather’
  - Extreme PAL (triggered by pal. glide; glide palatalizes preceding C and gets absorbed by it)
    - in underived environments (EUP), e.g. /xjoni/ \(\rightarrow\) ['çoni] ‘snow’
    - in morphologically derived environments (EDP), e.g. /'nixj+a/ \(\rightarrow\) ['niç+a] ‘nails’
- B & T (2012) present preliminary evidence that EUP and EDP are phonetically different
  - EDP: longer transition duration from pal.C to V and greater stability of the position of the transition point in the F1XF2 space \(\rightarrow\) maybe less absorption of the palatalizing trigger?
  - EUP: shorter transition duration and more variable position of the transition point \(\rightarrow\) maybe more absorption of the palatalizing trigger?
- But are EDP and EUP phonologically different too? Solomos’ lines above suggests not, since [i-sci] and [nea-ni-sci] are treated phonologically on a par
Considerations & Future Work
Considerations

- Fix glitches, e.g. algorithms may identify more RPs than the poem’s structure suggests.
- Mavilis’s Paliokastritsa is a sonnet; here RPs should be lines 1-4 and 2-3. RPs 1-2, 1-3 are also – incorrectly? – identified as imperfect rhymes. But what to do when the poem has no clear structure (as in Dodekalogos for example)?

```
1 1/1 Σαν πεθάνω εδώ θάρησ τα μύρια... (3) 'san pe-'Ta-no e-'Do 'Tar-To 'me 'ta 'mi-η)a
2 1/2 φαντάσματα άνων μέσα σε άνηλα γνέφια,... F2 ( IMP-C, IMP-V, IDV (IDV-2W ) )
3 1/3 ή σε ασθμοβολής μακά σεντέφια... F2 ( IMP-C, IMP-V )
4 1/4 τ’άνγις της νύχτας να χαρώ μυστήρια... F2 ( IMP-0F )
```

```
2 1/2 φαντάσματα άνων μέσα σε άνηλα γνέφια,... (2) fa-'de-zma-ta 'a-i-pna 'me-sa 'se 'a-i-la 'Gne-fCa
3 1/3 ή σε ασθμοβολής μακά σεντέφια... (1) 'i 'se a-si-mo-vo-'lis ma-i-ka se-'de-fCa
4 1/4 τ’άνγις της νύχτας να χαρώ μυστήρια... (0) 'a-j'a 'tis 'ni-xtas 'na xa-'ro mi-'si-η)a
```
On GUI considerations...

Decisions on GUI are not as simple as they seem! Compare the two examples below

- Kariotakis - Gala

  Τ’ αστέρια τρεμούλιαζουν καθώς
to μάτι ανοιγκλεί προτού δακρύσει.
O κόσμος τω δεντρώνε ρέβει όρθος.
Κλαίει παρακάτω η βρύση.

  (1) ta-ste-ri'a tre-mu-li-a-zou-ne ka-Tos
  (1) 'to 'ma-ti a-ni-Go-kli pro-tu Da-'kri-si
  (0) 'o 'ko-zmos 'to De-'dro-ne 're-vi or-'Tos
  (0) 'klei pa-ra-'ka-tu 'i 'vri-si

- Varnalis – Portreto se rimes

  Νούρι νούρι μελανούρι
έχει κούτελο και μύρη
η Μιράντα το μωρό:
dε μπορεί να πει το ρο!

  (1) 'nu-ri 'nu-ri me-la-'nu-ri
  (0) 'e-Ci 'ku-te-lo 'ce 'mu-ri
  (1) 'i mi-'ra-da 'to mo-'ro
  (0) 'De bo-'ri 'na 'pi 'to 'ro
...On GUI considerations

- Rhyme representation only appears once (the first time it is encountered), thus the RP 1-3 is noted, but not the RP 3-1. Why?
- Rhyme representation only on first line benefits
  - Intuitive inference of rhyme type
    - in examples above: Kariotakis $\rightarrow$ πλεχτή (1-3, 2-4), but Varnalis $\rightarrow$ ζευγαρωτή (1-2, 3-4)
  - Easy detection of redundant/false detected rhymes
  - Single appearance in statistics
  - In case corrections are needed, there is only a single editing position
Future Work

- Expand database
- Optimize GUI
- Current project does not cover this, but the current structure and rules allow for future integration of metric scheme attributes according to rhyme type characterization (πλεχτή, σταυρωτή, ζευγαρωτή), poem type (e.g. sonnets) or syllabification issues (recognition of synaloepha)
- Extend to other phenomena, e.g. alliteration
- Make this more useful/accessible to philology/poetry scholars and school teachers. How? Ideas?
Thank you!
Bibliography


- Κοκόλης, Ξ.Α. 1993. Η ομοιοκαταληξία. Τύποι και λειτουργικές διαστάσεις. Αθήνα: Στιγμή


<table>
<thead>
<tr>
<th>SAMPA</th>
<th>IPA</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>p</td>
<td>p</td>
<td>voiceless bilabial plosive</td>
</tr>
<tr>
<td>b</td>
<td>b</td>
<td>voiced bilabial plosive</td>
</tr>
<tr>
<td>t</td>
<td>t</td>
<td>voiceless alveolar plosive</td>
</tr>
<tr>
<td>d</td>
<td>d</td>
<td>voiced alveolar plosive</td>
</tr>
<tr>
<td>ts</td>
<td>ts</td>
<td>voiceless alveolar affricate</td>
</tr>
<tr>
<td>dz</td>
<td>dz</td>
<td>voiced alveolar affricate</td>
</tr>
<tr>
<td>c</td>
<td>c</td>
<td>voiceless palatal plosive</td>
</tr>
<tr>
<td>θ</td>
<td>j</td>
<td>voiced palatal plosive</td>
</tr>
<tr>
<td>k</td>
<td>k</td>
<td>voiceless velar plosive</td>
</tr>
<tr>
<td>g</td>
<td>g</td>
<td>voiced velar plosive</td>
</tr>
<tr>
<td>f</td>
<td>f</td>
<td>voiceless labiodental fricative</td>
</tr>
<tr>
<td>v</td>
<td>v</td>
<td>voiced labiodental fricative</td>
</tr>
<tr>
<td>θ</td>
<td>θ</td>
<td>voiceless dental fricative</td>
</tr>
<tr>
<td>δ</td>
<td>δ</td>
<td>voiced dental fricative</td>
</tr>
<tr>
<td>s</td>
<td>s</td>
<td>voiceless alveolar fricative</td>
</tr>
<tr>
<td>z</td>
<td>z</td>
<td>voiced alveolar fricative</td>
</tr>
<tr>
<td>c</td>
<td>c</td>
<td>voiceless palatal fricative</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>voiced palatal fricative</td>
</tr>
<tr>
<td>x</td>
<td>x</td>
<td>voiceless velar fricative</td>
</tr>
<tr>
<td>γ</td>
<td>γ</td>
<td>voiced velar fricative</td>
</tr>
<tr>
<td>m</td>
<td>m</td>
<td>bilabial nasal</td>
</tr>
<tr>
<td>f</td>
<td>m</td>
<td>labiodental nasal</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>alveolar nasal</td>
</tr>
<tr>
<td>j</td>
<td>p</td>
<td>palatal nasal</td>
</tr>
<tr>
<td>n</td>
<td>n</td>
<td>velar nasal</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>alveolar lateral approximant</td>
</tr>
<tr>
<td>l</td>
<td>l</td>
<td>palatal lateral approximant</td>
</tr>
<tr>
<td>r</td>
<td>r</td>
<td>alveolar trill</td>
</tr>
<tr>
<td>j</td>
<td>j</td>
<td>palatal approximant</td>
</tr>
</tbody>
</table>

Appendix

- Sampa – IPA correspondence for Greek consonants