

# Detección automática de rasgos métricos y encabalgamiento: aplicaciones

*Seminario LEMH, 04/2018*

Pablo Ruiz ([pablo.ruiz@linhd.uned.es](mailto:pablo.ruiz@linhd.uned.es))

LINHD - Laboratorio de Innovación en Humanidades Digitales, UNED

SCC - Sistemas de Comunicación y Control, UNED

Miembro del proyecto POSTDATA - Poetry Standardization and Linked Open Data

# Plan

- Proyecto POSTDATA
- Corpus
- Herramientas de análisis
  - Métrica
  - Encabalgamiento
- Resultados
  - Métrica
  - Encabalgamiento
- Discusión

# Proyecto POSTDATA



**POSTDATA**

Poetry Standardization  
and Linked Open Data

IP: Elena González-Blanco

Director Técnico: Salvador Ros



# Proyecto POSTDATA

- **Modelo conceptual común** para describir poesía en las diferentes tradiciones europeas
  - Según criterios lingüísticos
- Representación de los poemas como **datos enlazados** (web semántica)
- **Automatización** de las anotaciones lingüísticas
  - Procesamiento del lenguaje natural

# Plan

- Proyecto POSTDATA
- **Corpus**
- Herramientas de análisis
  - Métrica
  - Encabalgamiento
- Resultados
  - Métrica
  - Encabalgamiento
- **Discusión**

# Corpus (sonetos)

- ADSO (Navarro et al., 2016)

*Análisis Distante del Siglo de Oro*

– <https://github.com/bncolorado/CorpusSonetosSigloDeOro>

- DISCO (Ruiz et al., 2017)

*Diachronic Spanish Sonnet Corpus*

– <https://github.com/postdataproject/disco>

# Corpus (sonetos)

- ADSO (Navarro et al., 2016)
  - 5000 sonetos del Siglo de Oro, TEI
  - Anotaciones métricas
  - Autores canónicos
- DISCO (Ruiz et al., 2017)
  - 4000 sonetos de los siglos XV al XIX, TEI
  - Anotaciones métricas y encabalgamiento
  - Autores canónicos y no, Europa y América

# Corpus (sonetos)

- ADSO (Navarro et al., 2016)

<https://github.com/bncolorado/CorpusSonetosSigloDeOro>



**ADSO** *Distant reading approach to Golden Age Spanish Sonnets*

ADSO project ▾ Corpus of

## ADSO project

Search .

The purpose of this Project is to develop a macroanalysis and distant reading of Spanish Golden Age Sonnets, from the Renaissance Era (Garcilaso de la Vega) to the poetry from the end of the Baroque period (Sor Juana Inés de la Cruz). Computational methods will be



# Corpus (sonetos)

- DISCO (Ruiz et al., 2017)
  - Dentro de nuestro proyecto
  - <https://github.com/postdataproject/disco>

# Corpus DISCO

Periodo	Número de sonetos	Número de autores	Fuentes	
Siglo de Oro	1088	477	Mujeres 31	García (2006a)
			Hombres 446	
			América 12	
			Europa 458 (+7)	
Siglo XVIII	323	42	Mujeres 1	García (2005) Wikisource
			Hombres 41	
			América 6	
			Europa 36	
Siglo XIX	2676	685	Mujeres 48	García (2006b)
			Hombres 637	
			América 334	
			Europa 348 (+3)	

# Corpus DISCO

Periodo	Número de sonetos	Número de autores	Fuentes	
Siglo de Oro	1088	Mujeres	31	García (2006a)
		Hombres	446	
		América	12	
		Europa	458 (+7)	
Siglo XVIII	323	Mujeres	1	García (2005) Wikisource
		Hombres	41	
		América	6	
		Europa	36	
Siglo XIX	2676	Mujeres	48	García (2006b)
		Hombres	637	
		América	334	
		Europa	348 (+3)	

## △▽ **Carvajal y Mendoza, Luisa de**

Jaraicejo (Cáceres). 1566 - Londres. 1614

Religiosa.

### **A la ausencia de su Dulcísimo Señor en la Sagrada Comunión**

¡Ay, soledad amarga y enojosa,  
causada de mi ausente y dulce Amado!  
¡Dardo eres en el alma atravesado,  
dolencia penosísima y furiosa!

# Corpus DISCO: HTML original

△▽ **Carvajal y Mendoza, Luisa de**

Jaraicejo (Cáceres). 1566 - Londres. 1614

Religiosa.

## **A la ausencia de su Dulcísimo Señor en la Sagrada Comunión**

¡Ay, soledad amarga y enojosa,  
causada de mi ausente y dulce Amado!  
¡Dardo eres en el alma atravesado,  
dolencia penosísima y furiosa!

# Corpus DISCO: HTML original

```
<h2><a name="I_99_" id="I_99_"></a><a href="#I_98_"></a><a href="#I_100_"></a>Carvajal y Mendoza, Luisa de</h2></div>
```

```
<a name="1431"></a>
```

```
<p style="text-align: center;">Jaraicejo (C&acute;ceres). 1566 - Londres. 1614</p>
```

```
<a name="1432"></a>
```

```
<p style="text-align: center;">Religiosa.</p>
```

```
<a name="1433"></a><br clear="all" />
```

```
<table cellpadding="0" cellspacing="0" align="center" width="462">
```

```
<tr>
```

```
<td colspan="5"><br />
```

```
<br />
```

```
<div align="center"><span class="h3"><strong>A la ausencia de su Dulc&iacute;simo Se&ntilde;or en la Sagrada Comuni&oacute;n</strong></span><br />
```

```
<br /></div>
```

```
</td>
```

```
</tr>
```

# Corpus DISCO: Extracción de metadatos

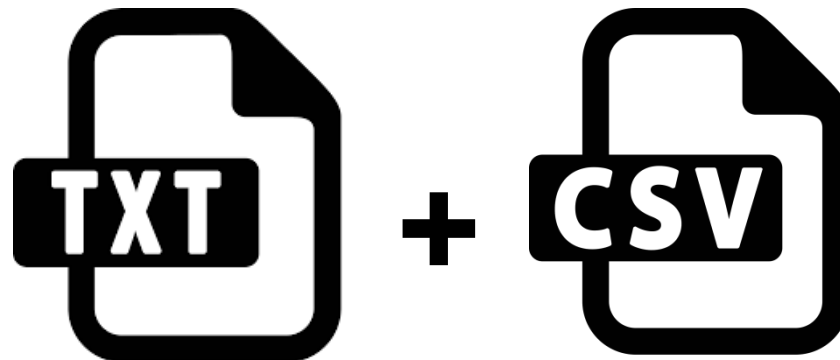
- Género
- Fechas de nacimiento y fallecimiento
- Origen: lugar, país, continente
- Identificador de VIAF

# Corpus DISCO: Extracción de metadatos





# Corpus DISCO: Extracción de metadatos



# Corpus DISCO: Metadatos

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
</persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
  <sex property="foaf:gender" value="F"/>
```

```
  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
        <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
        <country>España</country>
```

```
        <bloc>Europa</bloc>
```

```
      </placeName>
```

# Corpus DISCO: Metadatos

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1 <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
<persName type="short">Carvajal y Mendoza</persName>
```

```
<persName type="full">
```

```
  <forename property="foaf:givenName">Luisa</forename>
```

```
  <nameLink>de</nameLink>
```

```
  <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
</persName>
```

```
<persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
<sex property="foaf:gender" value="F"/>
```

```
<birth>
```

```
  <location property="schema:birthPlace">
```

```
    <placeName typeOf="schema:Place">
```

```
      <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
      <country>España</country>
```

```
      <bloc>Europa</bloc>
```

```
    </placeName>
```

# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
</persName>
```

```
<persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
<sex property="foaf:gender" value="F"/>
```

```
<birth>
```

```
  <location property="schema:birthPlace">
```

```
    <placeName typeOf="schema:Place">
```

```
      <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
      <country>España</country>
```

```
      <bloc>Europa</bloc>
```

```
    </placeName>
```

# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
2  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
  <sex property="foaf:gender" value="F"/>
```

```
  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
        <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
        <country>España</country>
```

```
        <bloc>Europa</bloc>
```

```
      </placeName>
```

# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
    property="dc:creator"
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
2  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
3  <sex property="foaf:gender" value="F"/>
```

```
  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
        <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
        <country>España</country>
```

```
        <bloc>Europa</bloc>
```

```
      </placeName>
```

# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
  <sex property="foaf:gender" value="F"/>
```

```
  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
        <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
        <country>España</country>
```

```
        <bloc>Europa</bloc>
```

```
      </placeName>
```

1

2

3

4



# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
2  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
3  <sex property="foaf:gender" value="F"/>
```

```
4  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
5      <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
      <country>España</country>
```

```
      <bloc>Europa</bloc>
```

```
    </placeName>
```



# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
2  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
3  <sex property="foaf:gender" value="F"/>
```

```
4  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
5      <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
      <country>España</country>
```

```
      <bloc>Europa</bloc>
```

```
    </placeName>
```

# Corpus DISCO: Metadatos

VIAF

Fichier d'autorité international virtuel

```
<listPerson>
```

```
  <person xml:id="disco_241g"
```

```
    property="dc:creator"
```

```
    typeOf="foaf:Person schema:Person dc:Agent">
```

```
1  <idno cert="high" resource="https://viaf.org/viaf/54173778"/>
```

```
2  <persName type="short">Carvajal y Mendoza</persName>
```

```
  <persName type="full">
```

```
    <forename property="foaf:givenName">Luisa</forename>
```

```
    <nameLink>de</nameLink>
```

```
    <surname property="foaf:familyName">Carvajal y Mendoza</surname>
```

```
  </persName>
```

```
  <persName type="source" property="foaf:name">Luisa de Carvajal y Mendoza</persName>
```

```
3  <sex property="foaf:gender" value="F"/>
```

```
4  <birth>
```

```
    <location property="schema:birthPlace">
```

```
      <placeName typeOf="schema:Place">
```

```
5      <settlement property="schema:name">Jaraicejo (Cáceres)</settlement>
```

```
      <country>España</country>
```

```
      <bloc>Europa</bloc>
```

```
    </placeName>
```

# Plan

- Proyecto POSTDATA
- Corpus
- Herramientas de análisis
  - Métrica
  - Encabalgamiento
- Resultados
  - Métrica
  - Encabalgamiento
- Discusión

# Análisis métrico (ADSO, Navarro 2017)

- Navarro Colorado et al. (2017) (*DSH*)
- Sistema para metro fijo (11 sílabas)
- Resolución de ambigüedades métricas (p. ej. varias sinalefas posibles para 11 sílabas)
  - Base de frecuencias de patrones métricos
  - Acepta el candidato de mayor frecuencia
- 92% escansiones correctas (N=100 sonetos)

# Ambigüedad métrica (Navarro 2017)

(1) cuando el padre Hebrero nos enseña

*[12 sílabas léxicas]*

(Hurtado de Mendoza)

(2) cuan-do\_el pa-dre\_He-bre-ro nos en-se-ña

*[2 sinalefas posibles para 11 síl métricas]*

(3) cuan-**do**\_**el** pa-dre He-bre-ro nos en-se-ña

*[3-6-10]*

(4) cuan-do el pa-**dre**\_He-bre-ro nos en-se-ña

*[4-6-10]*

# Ambigüedad métrica (Navarro 2017)

(1) cuando el padre Hebrero nos enseña

*[12 sílabas léxicas]*

(Hurtado de Mendoza)

(2) cuan-do\_el pa-dre\_He-bre-ro nos en-se-ña

*[2 sinalefas posibles para 11 síl métricas]*

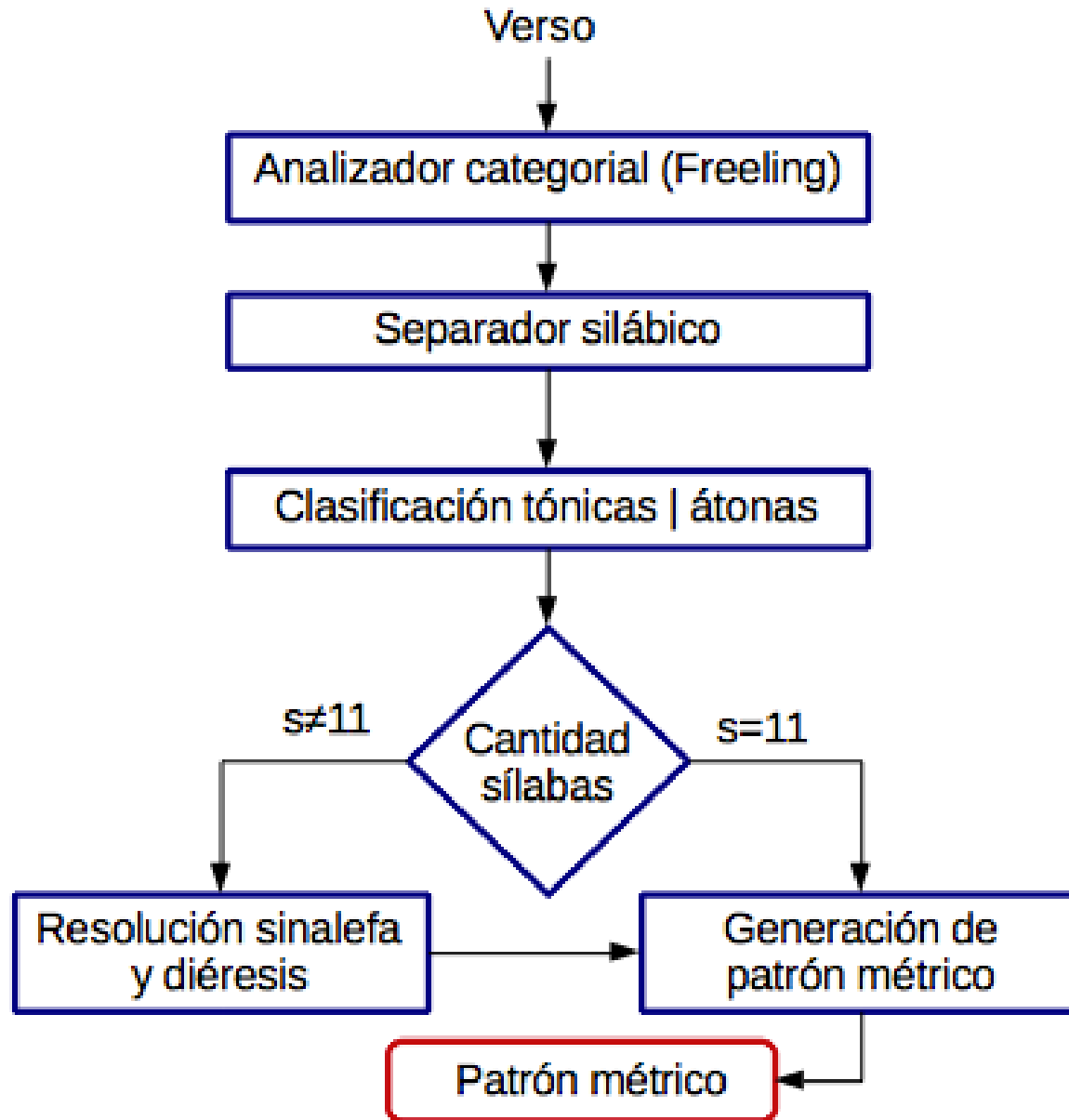
(3) cuan-do\_el pa-dre He-bre-ro nos en-se-ña

*[3-6-10] (más frecuente que 4-6-10)*

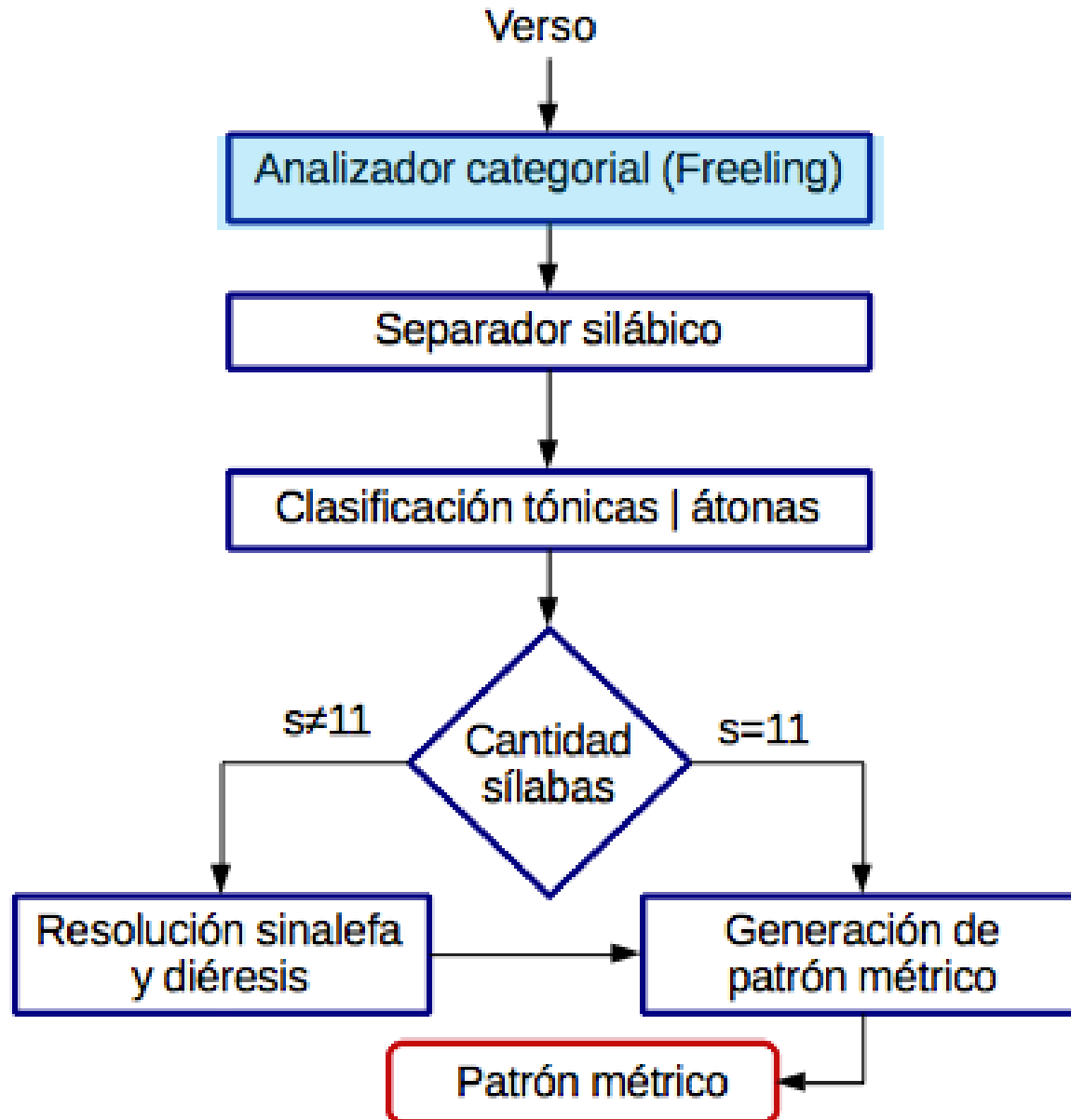
(4) cuan-do el pa-dre\_He-bre-ro nos en-se-ña

*[4-6-10]*

# Análisis métrico (ADSO, Navarro 2017)

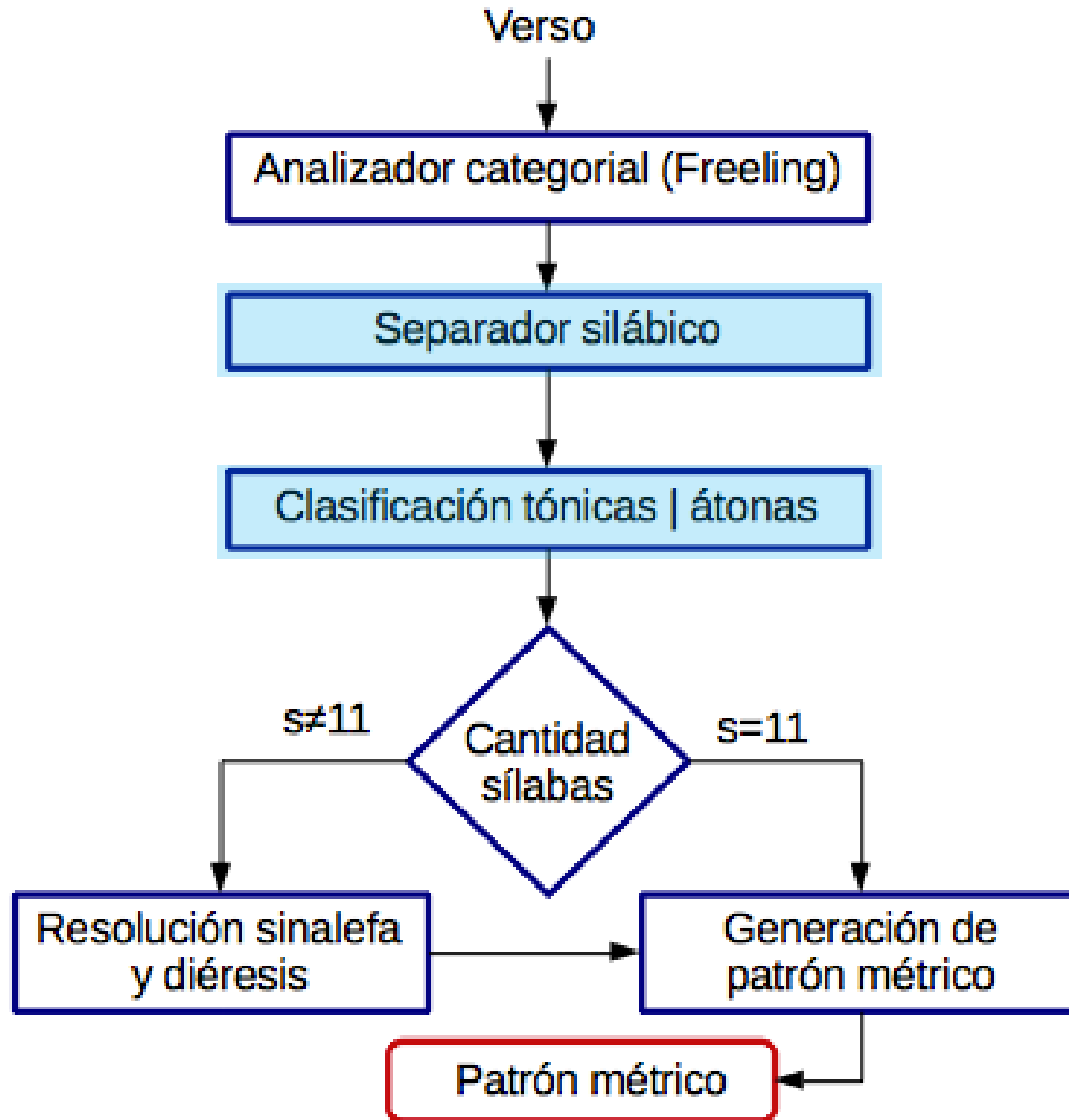


# Análisis métrico (ADSO, Navarro 2017)

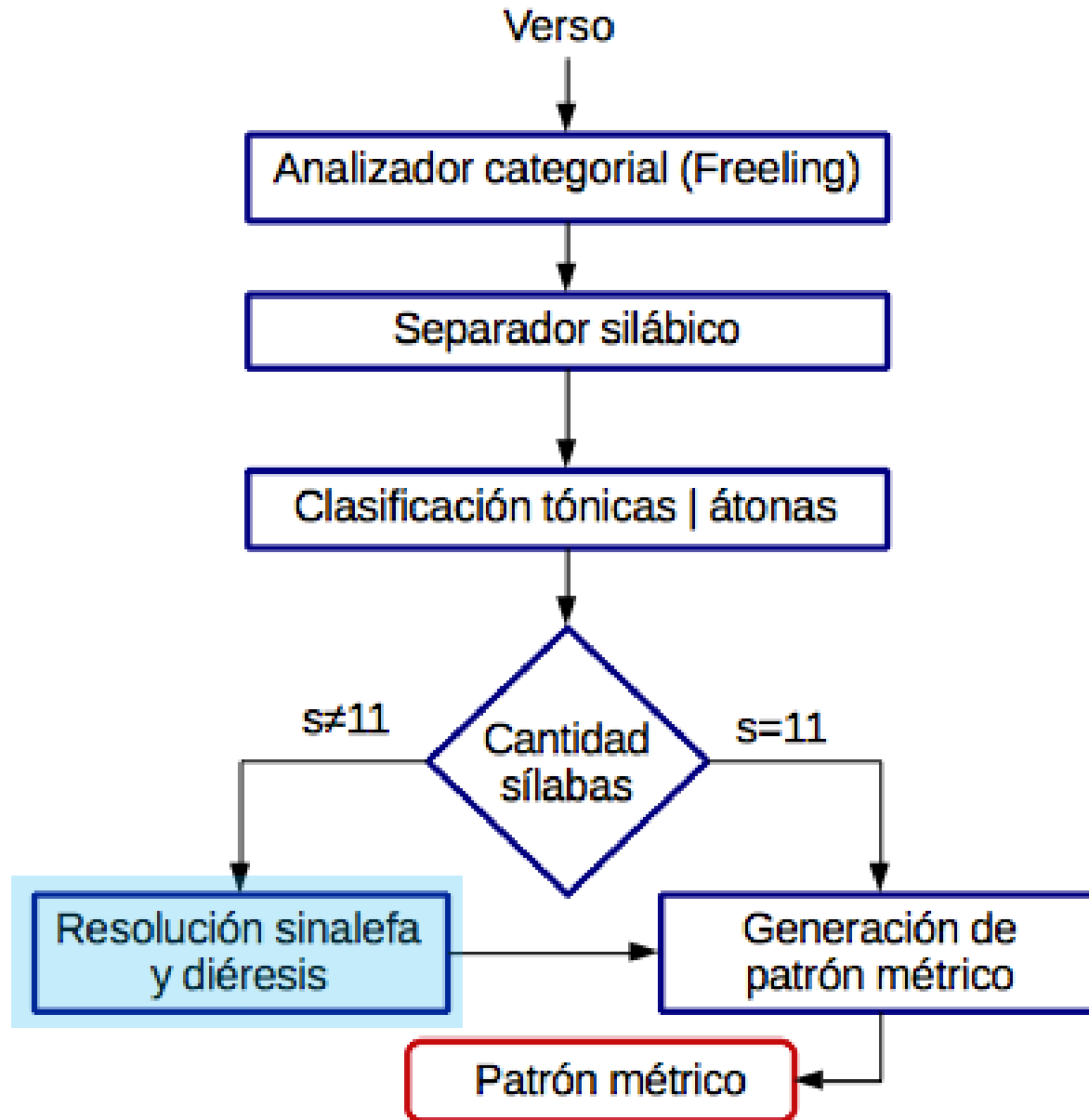




# Análisis métrico (ADSO, Navarro 2017)



# Análisis métrico (ADSO, Navarro 2017)



# Análisis métrico (ADSO, Navarro 2017)

```
<lg type="cuarteto">
```

```
<l n="1" met="+-+--+---+-">Este es el propio tiempo de mudarse</l>
```

```
<l n="2" met="--+--+---+-">cuando el padre Hebrero nos enseña,</l>
```

```
<l n="3" met="+---+--+---+-">ora mostrando cara halagüeña,</l>
```

```
<l n="4" met="+---+--+---+-">ora mostrando al cielo de enojarse.</l>
```

```
</lg>
```

# Trabajo en curso sobre escansión

#	Text	1	2	3	4	5	6	7	8	9	10	11
1	Ya se doran las hojas, ya las tardes	Ya	se	do	ran	las	ho	jas	ya	las	tar	des
2	Son lo mejor del día	Son	lo	me	jor	del	dí	a				
3	Bajo su luz madura de sosiego,	Ba	jo	su	luz	ma	du	ra	de	so	sie	go

# Plan

- Proyecto POSTDATA
- Corpus
- Herramientas de análisis
  - Métrica
  - Encabalgamiento
- Resultados
  - Métrica
  - Encabalgamiento
- Discusión

# Encabalgamiento: definición

- Discordancia entre unidades métricas y sintácticas
- Escisión de una unidad sintáctica a través del final del verso (o una pausa métrica)
- Un “corte” no natural resulta de esto, que puede dar lugar a efectos expresivos (énfasis, contraste, dobles sentidos)

# Encabalgamiento: definición

## EJEMPLO

1. Prefiero
2. quedarme **desnuda //**
3. **de amigo insincero**

## INTERPRETACIONES

- a) desnuda (*sin ropa*)
- b) desnuda de amigo insincero (*sin tal persona*)

# Encabalgamiento: definición

## EJEMPLO

1. Prefiero
2. quedarme **desnuda //**
3. **de amigo insincero**

## INTERPRETACIONES

a) desnuda (*sin ropa*)

b) desnuda de amigo insincero (*sin tal persona*)



# Encabalgamiento: definición

## EJEMPLO

1. Prefiero
2. quedarme **desnuda //**
3. **de amigo insincero**

## INTERPRETACIONES

a) desnuda (*sin ropa*)

b) desnuda de amigo insincero (*sin tal persona*)

# Encabalgamiento: debates actuales

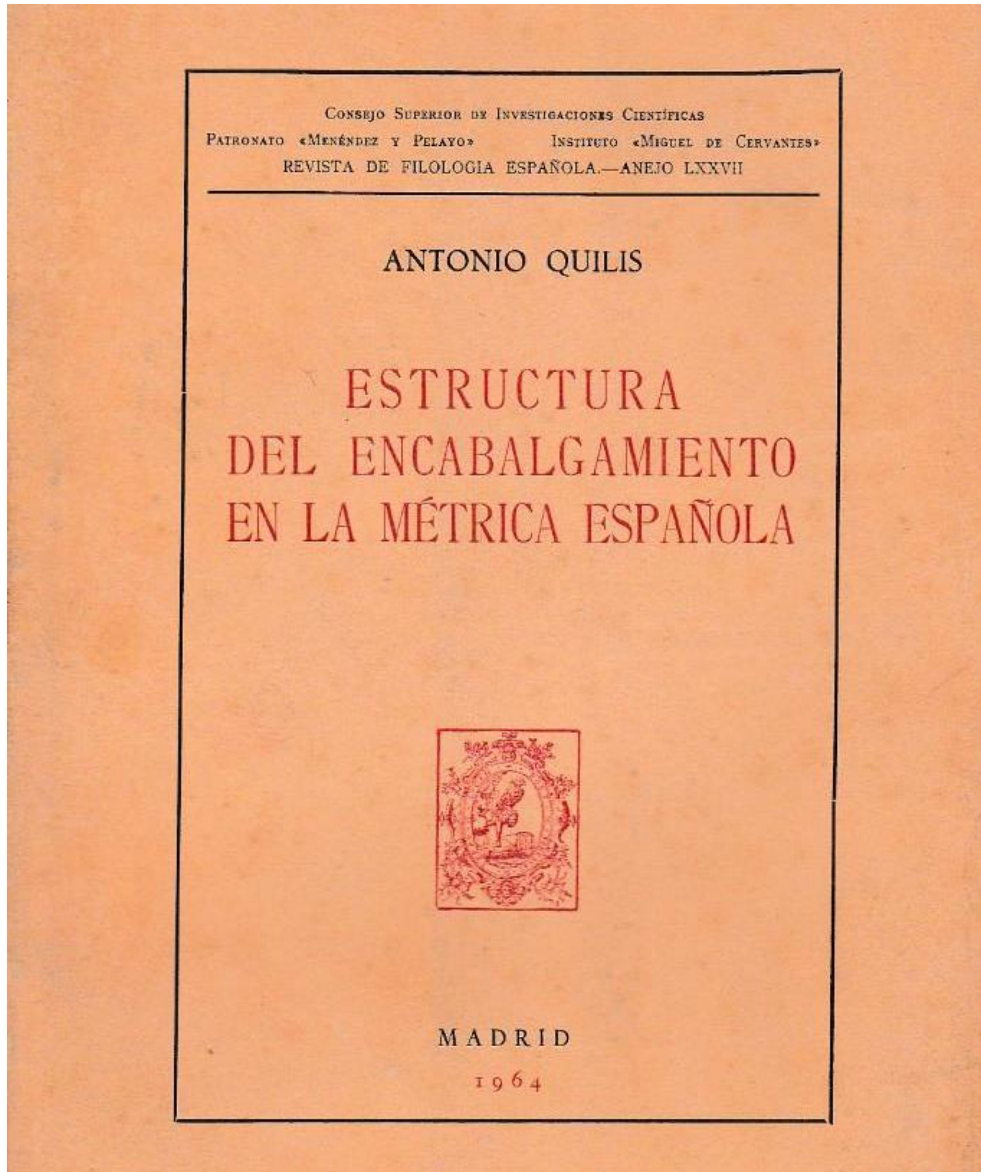
- ¿En que unidades sintácticas se pueden percibir efectos “no naturales” como los citados, cuando se escinden?
- ¿Cuáles son los valores expresivos posibles?
- ¿Es suficiente con criterios sintácticos para definir el encabalgamiento?

# Contribución de la automatización

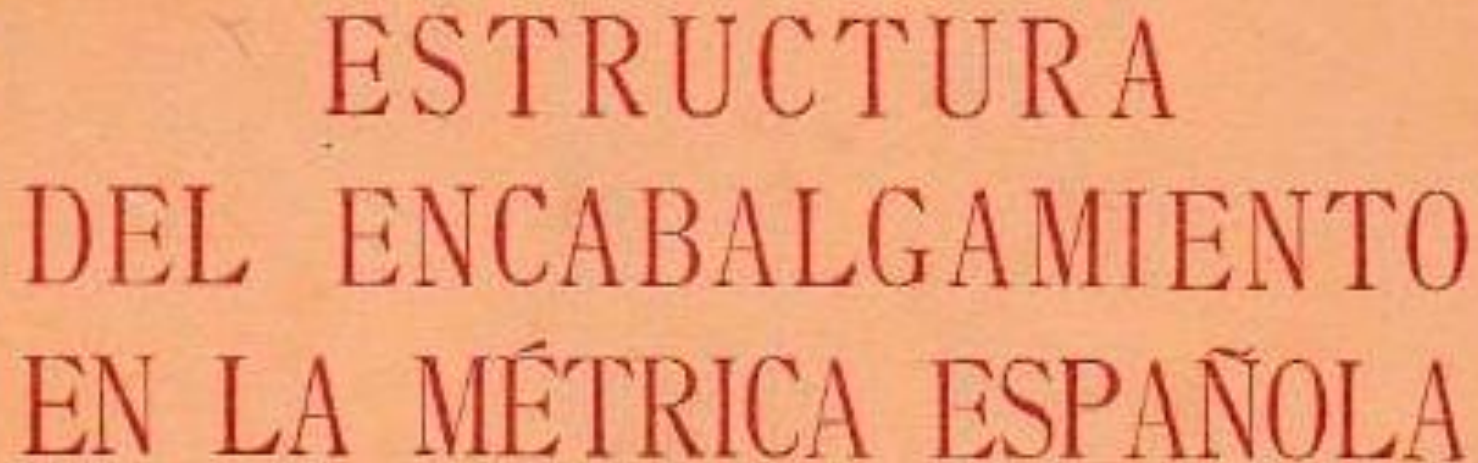
*Ruiz et al (2017); Martínez et al. (2018)*

- Recopilar ejemplos sistemáticamente, a partir de grandes corpus
- Ejemplos que puedan dar luz sobre los debates actuales en torno al fenómeno
- Proporcionar evidencia cuantitativa sobre el uso del fenómeno en distintos géneros y periodos

# El encabalgamiento en español



# El encabalgamiento en español



ESTRUCTURA  
DEL ENCABALGAMIENTO  
EN LA MÉTRICA ESPAÑOLA

- Quilis (1964). Estructura del encabalgamiento en la métrica española. Madrid, CSIC.
- Martínez Cantón, C. (2011). Ch. 4 « La Pausa » *Métrica y estética de Antonio Colinas*. Sevilla. Padilla Libros.

# Encabalgamiento según Quilis (1964)

- Sucede cuando sintagmas que tienen una **cohesión particular** se ven divididos por una frontera métrica (versal o de hemistiquio)
- Cohesión determinada experimentalmente:
  - **lecturas de prosa literaria**, midiendo pausas
  - **las unidades en las que no se hace pausa** se consideran muy cohesivas
  - Escindir las puede dar lugar a un efecto expresivo, son posibles contextos de encabalgamiento

# Encabalgamiento: tipología

## Algunos tipos *sirremáticos*

### Sintagmas nominales o adjetivos

- determinante + sustantivo
- sustantivo o adjetivo + complem. preposicional
- adj + sustantivo
- adj + adverbio

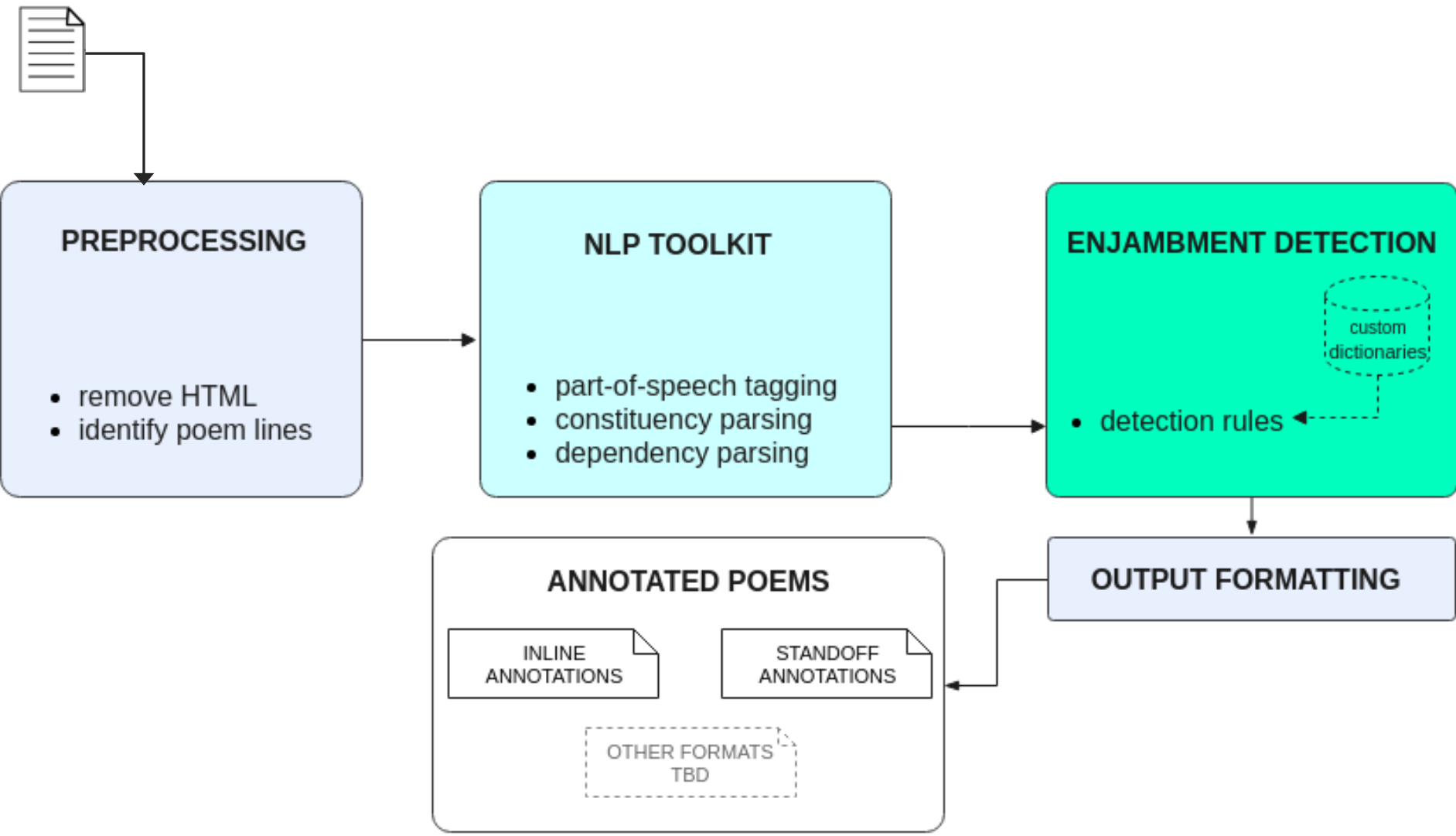
### Sintagmas verbales

- verbo + adverbio
- clítico + verbo
- tiempos compuestos
- verbo + suplemento

## TIPOLOGÍA COMPLETA

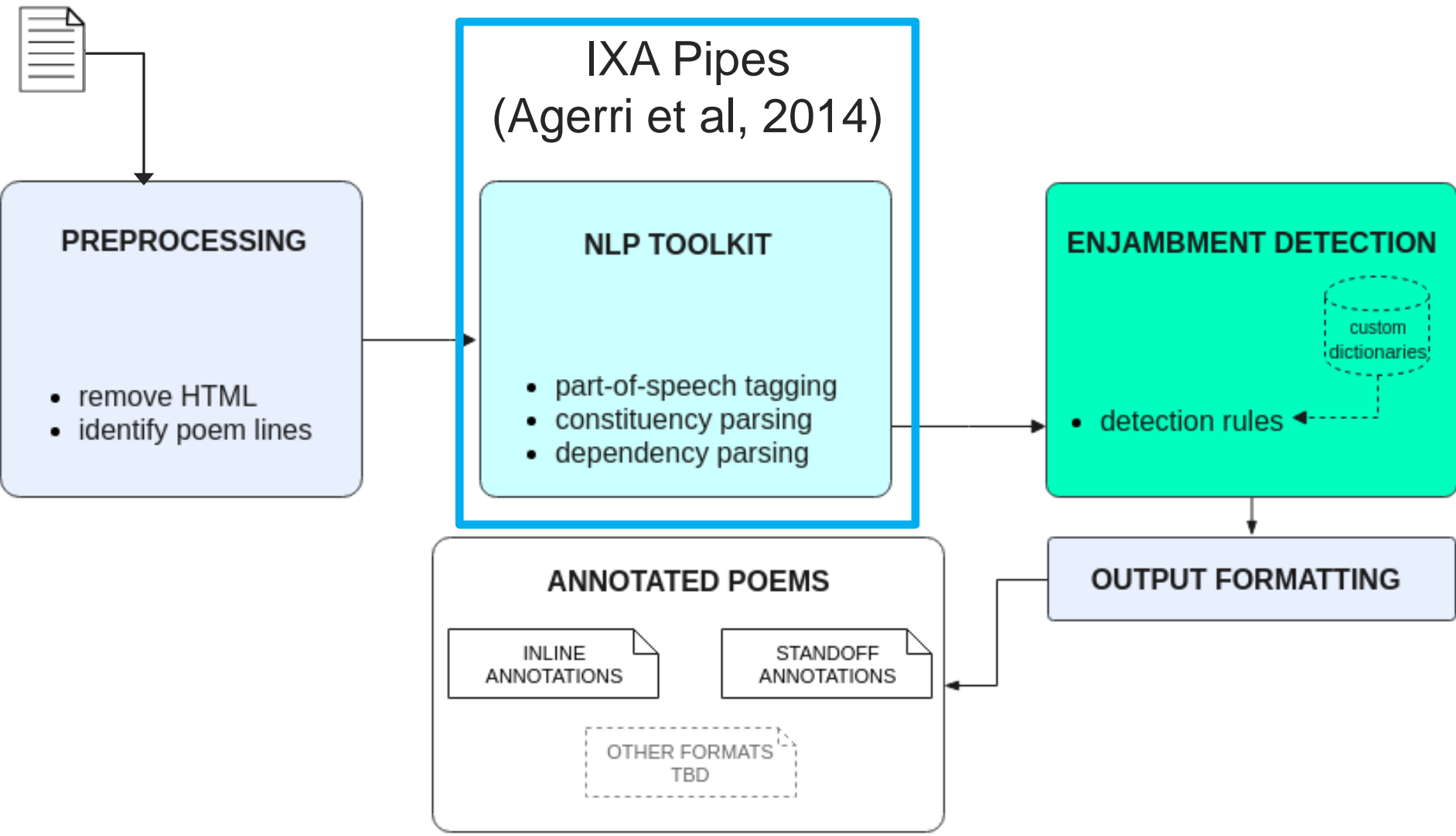
- <https://sites.google.com/site/spanishenjambment>

# Encabalgamiento: detección automática

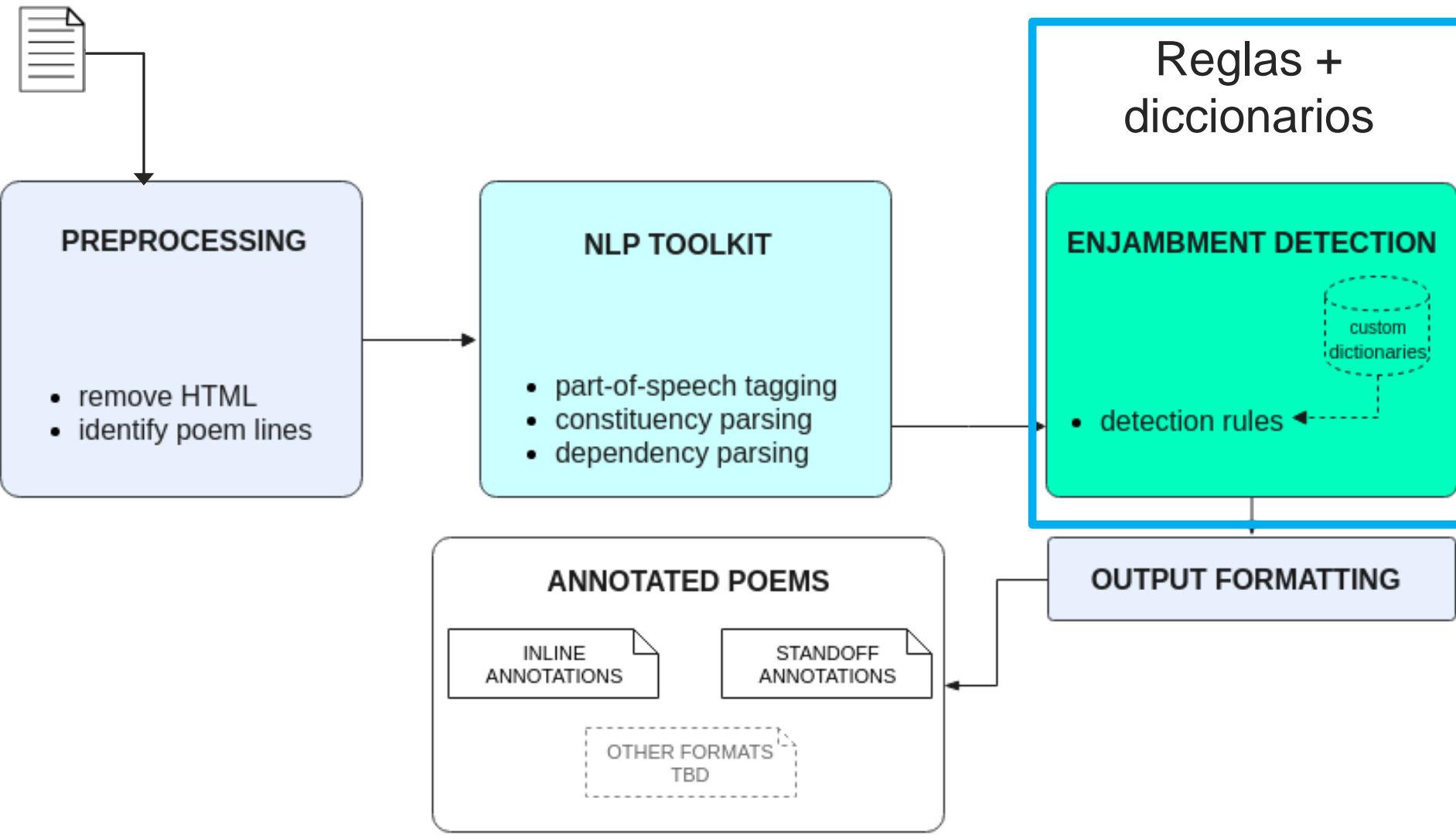




# Encabalgamiento: detección automática



# Encabalgamiento: detección automática



# Encabalgamiento: reglas

- Aproximadamente 30 reglas
- Basadas en **secuencias de categorías gramaticales**
- Su aplicación puede estar restringida a ciertos elementos léxicos (**lemas o formas**)
- Pueden estar basadas en **funciones sintácticas** (dependencias)

# Encabalgamiento: ejemplo de regla

- 1 si, en posición adyacente a la pausa versal, se encuentra
  - el **lema** de un verbo del diccionario de restricciones léxicas *Y*
  - la **preposición** a través de la cual ese verbo rige un suplemento
- 2 añadir el tipo **sirrem\_verb\_supl** al par de versos

# DetECCIÓN AUTOMÁTICA: EVALUACIÓN

- **Corpus** de referencia:
  - 1000 versos del siglo XX
  - 100 sonetos (1400 versos) de los siglos XV a XIX (con 30% del XIX)
  - Cubre una variedad de tipos de encabalgamiento
- **Acuerdo** entre anotadores?
  - En pares de versos con encabalgamiento: 75.4%
  - En pares de versos sin él: 95.7%

# Evaluación: tareas

- **Detección**

- Los pares de versos propuestos por el sistema deben estar en la referencia

- **Clasificación**

- Tanto los pares de versos propuestos por el sistema como el tipo propuesto deben coincidir con la referencia

# Evaluación: resultados

		N	P	R	F1
S. XX	detección	277	84.19	88.45	<b>86.27</b>
	clasificac.	277	78.50	83.03	<b>80.70</b>

		N	P	R	F1
Dia- crónico	detección	260	74.18	87.64	<b>80.35</b>
	clasificac.	260	61.24	72.31	<b>66.31</b>

Número de pares de versos, **P**recisión, **R**ecall, **F1**

# Evaluación: resultados

		N	P	R	F1
S. XX	detección	277	84.19	88.45	<b>86.27</b>
	clasificac.	277	78.50	83.03	<b>80.70</b>

		N	P	R	F1
Dia- crónico	detección	260	74.18	87.64	<b>80.35</b>
	clasificac.	260	61.24	72.31	<b>66.31</b>

Número de pares de versos, **P**recisión, **R**ecall, **F1**



# Fuentes de error

- Diferentes patrones de error en el siglo XX y en el corpus diacrónico
- **S. XX**: Además de errores por las reglas, errores comunes del motor de PLN fueron:
  - Análisis de sintagmas preposicionales
  - Confusión entre el sujeto y el objeto directo
- **Diacrónico**: Más variedad de errores, p. ej.
  - Etiquetado gramatical
  - Las reglas sobreaplican por el hipérbaton

**ANJA**

**A**utomatic en**NJ**ambment  
**A**nalysis

<http://prf1.org/anja/index/>

# Plan

- Proyecto POSTDATA
- Corpus
- Herramientas de análisis
  - Métrica
  - Encabalgamiento
- **Resultados**
  - Métrica
  - Encabalgamiento
- Discusión

# Resultados

- Comparación y enriquecimiento de los resultados de:

Navarro-Colorado (2016). Hacia un análisis distante del endecasílabo áureo: Patrones métricos, frecuencias y evolución histórica. *Rhythmica*, 14: pp. 89-118.

# Resultados

- Comparación y enriquecimiento de los resultados de:

Navarro-Colorado (2016). Hacia un análisis distante del endecasílabo áureo: Patrones métricos, frecuencias y evolución histórica. *Rhythmica*, 14: pp. 89-118.

[BN16]

# Resultados

- Mismos tipos de análisis que BN16
- Aplicados al corpus DISCO
- Aplicados a:
  - métrica
  - encabalgamiento
  - la interacción entre los dos
- Según permiten los metadatos de DISCO
  - fechas

# Resultados: interés

- Mismos tipos de análisis que BN16
- Aplicados al corpus **DISCO**
- Aplicados a:
  - métrica
  - **encabalgamiento**
  - la interacción entre los dos
- Según permitan los metadatos de DISCO
  - ¿fechas?
- Autores **conocidos vs. poco estudiados**

# Resultados: plan

## Comparación entre DISCO y ADSO

- Frecuencia de patrones métricos
- Frecuencia de bigramas métricos
  - Combinaciones de patrones binarios y ternarios
  - Paralelos (mismo patron en dos versos seguidos)
- Patrones por estrofa
- Interacción con encabalgamiento



# ADSO: evolución en el tiempo

- **Renacimiento**

- Basados en **6-10** más frecuentes
- **Generaciones**
  - Primera: 2-6-10 y 3-6-10
  - Segunda: 2-4-6-10 y 2-4-6-8-10 (más apoyos métricos)

- **Barroco**

- **2-4-8-10** y **2-6-10** más frecuentes
- **Generaciones**
  - Primera: 2 4 8 10
  - Segunda: 2 6 10 y 3 6 10 (vuelve el gusto renacentista)

# Comparaciones posibles

- En DISCO\_SO tenemos el *SIGLO*, no fechas exactas de vida, para una proporción importante de los autores
- Comparaciones temporales detalladas no son posibles
- Es posible comparar todo DISCO\_SO con todo ADSO y DISCO\_XIX

# Comparaciones posibles

- En DISCO\_SO tenemos el *SIGLO*, no fechas exactas de vida, para una proporción importante de los autores
- Comparaciones temporales detalladas no son posibles
- Es posible comparar todo DISCO\_SO con todo ADSO y DISCO\_XIX

# Comparaciones posibles

- Comparar todo DISCO\_SO con todo ADSO y DISCO\_XIX
  - unigramas y bigramas métricos
  - combinaciones de patrones binarios y ternarios
  - encabalgamiento

# Patrones métricos por (sub)corpus

Patrón	DISCO_SO		ADSO		DISCO_XIX	
3-6-10	809	<b>6.45</b>	5031	<b>7.07</b>	3540	<b>9.34</b>
2-6-10	748	<b>5.96</b>	5568	<b>7.83</b>	2695	<b>7.11</b>
2-4-6-10	711	<b>5.67</b>	4365	6.14	1322	3.49
2-4-8-10	697	5.55	5506	<b>7.74</b>	2854	<b>7.53</b>
4-6-10	541	4.31	3342	4.70	1585	4.18
2-6-8-10	492	3.92	3151	4.43	935	2.47
3-6-8-10	455	3.63	2572	3.62	1112	2.93
1-4-6-10	442	3.52	2717	3.82	1052	2.78
2-4-6-8-10	429	3.42	2776	3.90	484	1.28
1-4-8-10	392	3.12	2763	3.88	1739	4.59
4-8-10	304	2.42	2776	3.90	1582	4.17

# Patrones métricos por (sub)corpus

Patrón	DISCO_SO		ADSO		DISCO_XIX	
3-6-10	809	1	5031	3	3540	1
2-6-10	748	2	5568	1	2695	3
2-4-6-10	711	3	4365	6.14	1322	3.49
2-4-8-10	697	5.55	5506	2	2854	2
4-6-10	541	4.31	3342	4.70	1585	4.18
2-6-8-10	492	3.92	3151	4.43	935	2.47
3-6-8-10	455	3.63	2572	3.62	1112	2.93
1-4-6-10	442	3.52	2717	3.82	1052	2.78
2-4-6-8-10	429	3.42	2776	3.90	484	1.28
1-4-8-10	392	3.12	2763	3.88	1739	4.59
4-8-10	304	2.42	2776	3.90	1582	4.17

# Patrones métricos por (sub)corpus

Patrón	DISCO_SO		ADSO		DISCO_XIX	
3-6-10	809	1	5031	3	3540	1
2-6-10	748	2	5568	1	2695	3
2-4-6-10	711	3	4365	4	1322	3.49
2-4-8-10	697	4	5506	2	2854	2
4-6-10	541	4.31	3342	4.70	1585	4.18
2-6-8-10	492	3.92	3151	4.43	935	2.47
3-6-8-10	455	3.63	2572	3.62	1112	2.93
1-4-6-10	442	3.52	2717	3.82	1052	2.78
2-4-6-8-10	429	3.42	2776	3.90	484	1.28
1-4-8-10	392	3.12	2763	3.88	1739	4
4-8-10	304	2.42	2776	3.90	1582	4.17

# Patrones métricos por (sub)corpus

Patrón	DISCO_SO		ADSO		DISCO_XIX	
3-6-10	809	1	5031	3	3540	1
2-6-10	748	2	5568	1	2695	3
2-4-6-10	711	3	4365	4	1322	7
2-4-8-10	697	4	5506	2	2854	2
4-6-10	541	4.31	3342	4.70	1585	4.18
2-6-8-10	492	3.92	3151	4.43	935	2.47
3-6-8-10	455	3.63	2572	3.62	1112	2.93
1-4-6-10	442	3.52	2717	3.82	1052	2.78
2-4-6-8-10	429	3.42	2776	3.90	484	1.28
1-4-8-10	392	3.12	2763	3.88	1739	4
4-8-10	304	2.42	2776	3.90	1582	4.17



# Correlaciones: unigramas métricos

<b>Corpus 1</b>	<b>Corpus 2</b>	<b>rho</b>
DISCO_SO	ADSO	0.792
DISCO_SO	DISCO_XIX	0.527
ADSO	DISCO_XIX	0.624

# Correlaciones: unigramas métricos

<b>Corpus 1</b>	<b>Corpus 2</b>	<b>rho</b>
DISCO_SO	ADSO	0.792
DISCO_SO	DISCO_XIX	0.527
ADSO	DISCO_XIX	0.624

# Bigramas métricos por subcorpus

<b>DISCO_SO</b>		<b>ADSO</b>	
3-6-10   3-6-10	0.55	2-4-8-10   2-4-8-10	0.87
2-6-10   3-6-10	0.52	2-6-10   2-6-10	0.73
2-6-10   2-6-10	0.46	2-6-10   3-6-10	0.67
3-6-10   2-6-10	0.45	3-6-10   3-6-10	0.66
2-4-8-10   2-4-8-10	0.44	3-6-10   2-6-10	0.66
2-6-10   2-4-6-10	0.42	2-4-8-10   2-6-10	0.60
2-4-8-10   2-4-6-10	0.42	2-6-10   2-4-8-10	0.59
2-4-6-10   2-4-8-10	0.41	3-6-10   2-4-8-10	0.56
2-4-6-10   2-6-10	0.35	2-6-10   2-4-6-10	0.54
2-4-6-10   2-4-6-10	0.35	2-4-8-10   3-6-10	0.54

# Bigramas métricos por subcorpus

DISCO_SO		ADSO	
<b>3-6-10   3-6-10</b>	0.55	<b>2-4-8-10   2-4-8-10</b>	0.87
2-6-10   3-6-10	0.52	<b>2-6-10   2-6-10</b>	0.73
<b>2-6-10   2-6-10</b>	0.46	2-6-10   3-6-10	0.67
3-6-10   2-6-10	0.45	<b>3-6-10   3-6-10</b>	0.66
<b>2-4-8-10   2-4-8-10</b>	0.44	3-6-10   2-6-10	0.66
2-6-10   2-4-6-10	0.42	2-4-8-10   2-6-10	0.60
2-4-8-10   2-4-6-10	0.42	2-6-10   2-4-8-10	0.59
2-4-6-10   2-4-8-10	0.41	3-6-10   2-4-8-10	0.56
2-4-6-10   2-6-10	0.35	2-6-10   2-4-6-10	0.54
<b>2-4-6-10   2-4-6-10</b>	0.35	2-4-8-10   3-6-10	0.54

# Bigramas métricos por subcorpus

**DISCO\_SO**

**ADSO**

<b>3-6-10   3-6-10</b>	0.55	<b>2-4-8-10   2-4-8-10</b>	0.87
2-6-10   3-6-10	0.52	<b>2-6-10   2-6-10</b>	0.73
<b>2-6-10   2-6-10</b>	0.46	2-6-10   3-6-10	0.67
3-6-10   2-6-10	0.45	<b>3-6-10   3-6-10</b>	0.66
<b>2-4-8-10   2-4-8-10</b>	0.44	3-6-10   2-6-10	0.66
2-6-10   2-4-6-10	0.42	2-4-8-10   2-6-10	0.60
2-4-8-10   2-4-6-10	0.42	2-6-10   2-4-8-10	0.59
2-4-6-10   2-4-8-10	0.41	3-6-10   2-4-8-10	0.56
2-4-6-10   2-6-10	0.35	<b>2-6-10   2-4-6-10</b>	0.54
<b>2-4-6-10   2-4-6-10</b>	0.35	2-4-8-10   3-6-10	0.54

# Bigramas métricos por subcorpus

**DISCO\_SO**

**ADSO**

<b>3-6-10   3-6-10</b>	0.55	<b>2-4-8-10   2-4-8-10</b>	<b>+0.39</b>
2-6-10   3-6-10	0.52	<b>2-6-10   2-6-10</b>	+0.37
<b>2-6-10   2-6-10</b>	0.46	2-6-10   3-6-10	+0.21
3-6-10   2-6-10	0.45	<b>3-6-10   3-6-10</b>	+0.11
<b>2-4-8-10   2-4-8-10</b>	0.44	3-6-10   2-6-10	+0.22
2-6-10   2-4-6-10	0.42	2-4-8-10   2-6-10	
2-4-8-10   2-4-6-10	0.42	2-6-10   2-4-8-10	
2-4-6-10   2-4-8-10	0.41	3-6-10   2-4-8-10	
2-4-6-10   2-6-10	0.35	<b>2-6-10   2-4-6-10</b>	<b>+0.12</b>
<b>2-4-6-10   2-4-6-10</b>	0.35	2-4-8-10   3-6-10	

# Bigramas métricos por subcorpus

**DISCO\_SO**

**DISCO\_XIX**

3-6-10   3-6-10	0.55	3-6-10   3-6-10	1.41
2-6-10   3-6-10	0.52	2-4-8-10   2-4-8-10	0.96
2-6-10   2-6-10	0.46	3-6-10   2-6-10	0.94
3-6-10   2-6-10	0.45	2-6-10   3-6-10	0.92
2-4-8-10   2-4-8-10	0.44	2-6-10   2-6-10	0.87
2-6-10   2-4-6-10	0.42	2-4-8-10   3-6-10	0.68
2-4-8-10   2-4-6-10	0.42	3-6-10   2-4-8-10	0.65
2-4-6-10   2-4-8-10	0.41	1-4-8-10   2-4-8-10	0.55
2-4-6-10   2-6-10	0.35	2-4-8-10   4-8-10	0.53
2-4-6-10   2-4-6-10	0.35	2-4-8-10   2-6-10	0.51

# Bigramas métricos por subcorpus

## DISCO\_SO

## DISCO\_XIX

3-6-10   3-6-10	0.55	3-6-10   3-6-10	<b>+0.86</b>
2-6-10   3-6-10	0.52	2-4-8-10   2-4-8-10	+0.52
2-6-10   2-6-10	0.46	3-6-10   2-6-10	+0.49
3-6-10   2-6-10	0.45	2-6-10   3-6-10	+0.46
2-4-8-10   2-4-8-10	0.44	2-6-10   2-6-10	<b>+0.41</b>
2-6-10   2-4-6-10	0.42	2-4-8-10   3-6-10	
2-4-8-10   2-4-6-10	0.42	3-6-10   2-4-8-10	
2-4-6-10   2-4-8-10	0.41	1-4-8-10   2-4-8-10	
2-4-6-10   2-6-10	0.35	2-4-8-10   4-8-10	
2-4-6-10   2-4-6-10	0.35	2-4-8-10   2-6-10	



# Bigramas métricos por subcorpus

## ADSO

## DISCO\_XIX

2-4-8-10   2-4-8-10	0.87	3-6-10   3-6-10	1.41
2-6-10   2-6-10	0.73	2-4-8-10   2-4-8-10	0.96
2-6-10   3-6-10	0.67	3-6-10   2-6-10	0.94
3-6-10   3-6-10	0.66	2-6-10   3-6-10	0.92
3-6-10   2-6-10	0.66	2-6-10   2-6-10	0.87
2-4-8-10   2-6-10	0.60	2-4-8-10   3-6-10	0.68
2-6-10   2-4-8-10	0.59	3-6-10   2-4-8-10	0.65
3-6-10   2-4-8-10	0.56	1-4-8-10   2-4-8-10	0.55
2-6-10   2-4-6-10	0.54	2-4-8-10   4-8-10	0.53
2-4-8-10   3-6-10	0.53	2-4-8-10   2-6-10	0.51

# Bigramas métricos por subcorpus

**ADSO**

**DISCO\_XIX**

2-4-8-10   2-4-8-10	0.87	3-6-10   3-6-10	<b>+0.54</b>
2-6-10   2-6-10	0.73	2-4-8-10   2-4-8-10	+0.23
2-6-10   3-6-10	0.67	3-6-10   2-6-10	+0.28
3-6-10   3-6-10	0.66	2-6-10   3-6-10	+0.25
3-6-10   2-6-10	0.66	2-6-10   2-6-10	+0.14
2-4-8-10   2-6-10	0.60	2-4-8-10   3-6-10	+0.15
2-6-10   2-4-8-10	0.59	3-6-10   2-4-8-10	+0.09
3-6-10   2-4-8-10	0.56	1-4-8-10   2-4-8-10	
2-6-10   2-4-6-10	0.54	2-4-8-10   4-8-10	
2-4-8-10   3-6-10	0.53	2-4-8-10   2-6-10	<b>-0.09</b>

# Correlaciones: bigramas métricos

<b>Corpus 1</b>	<b>Corpus 2</b>	<b>rho</b>
DISCO_SO	ADSO	0.619
DISCO_SO	DISCO_XIX	0.48
ADSO	DISCO_XIX	0.640

# Correlaciones: bigramas métricos

<b>Corpus 1</b>	<b>Corpus 2</b>	<b>rho</b>
DISCO_SO	ADSO	0.619
DISCO_SO	DISCO_XIX	0.48
ADSO	DISCO_XIX	0.640

# Secuencias ternario-binario

- BN16: *“llama la atención cómo los patrones de ritmo ternario (con acento en tercera y sexta) tienden a resolver en patrones de ritmo binario (con acentos en segunda y sexta o en cuarta y octava)”*
- A examinar:
  - Si este patrón es típicamente áureo
  - Si está relacionado con ciertas posiciones en el soneto

# Secuencias ternario-binario

- BN16: *“llama la atención cómo los patrones de ritmo **ternario** (**con acento en tercera y sexta**) tienden a resolver en patrones de ritmo **binario** (con acentos en **segunda y sexta** o en **cuarta y octava**)”*

# Secuencias ternario-binario

- Binario
  - (2 y 6) o (4 y 8 pero no 1)
- Ternario
  - (3 y 6) y (ni 2 ni 4)
- No clasificado

# Secuencias ternario-binario

- Binario
  - (2 y 6) o (4 y 8 **pero no 1**)
- Ternario
  - (3 y 6) y (ni 2 ni 4)
- No clasificado



# Secuencias ternario-binario

	<b>DISCO_SO</b>		<b>ADSO</b>		<b>DISCO_XIX</b>	
ter – bin	834	7.14	5810	8.79	2994	8.49
ter – ter	485	4.15	2633	3.98	2507	7.11
ter – noclas	858	7.35	4007	6.06	2932	8.32

# Secuencias ternario-binario

	<b>DISCO_SO</b>		<b>ADSO</b>		<b>DISCO_XIX</b>	
<b>ter – bin</b>	834	7.14	5810	8.79	2994	8.49
<b>ter – ter</b>	485	4.15	2633	3.98	2507	7.11
<b>ter – noclas</b>	858	7.35	4007	6.06	2932	8.32

# Secuencias ternario-binario

	DISCO_SO		ADSO		DISCO_XIX	
<b>ter – bin</b>	834	<b>7.14</b>	5810	<b>8.79</b>	2994	<b>8.49</b>
<b>ter – ter</b>	485	<b>4.15</b>	2633	<b>3.98</b>	2507	<b>7.11</b>
ter – noclas	858	7.35	4007	6.06	2932	8.32

- Tendencia clara en Siglo de Oro de que los versos con patrón ternario vayan seguidos de un verso con patrón binario
- Posible problema: versos sin clasificar

# Secuencias ternario-binario: posición

- No he encontrado efectos de posición
- Encuentro
  - aprox. 30% en cada cuarteto
  - aprox 20% en cada terceto
- Es lo que cabría esperar si no hay efectos de posición

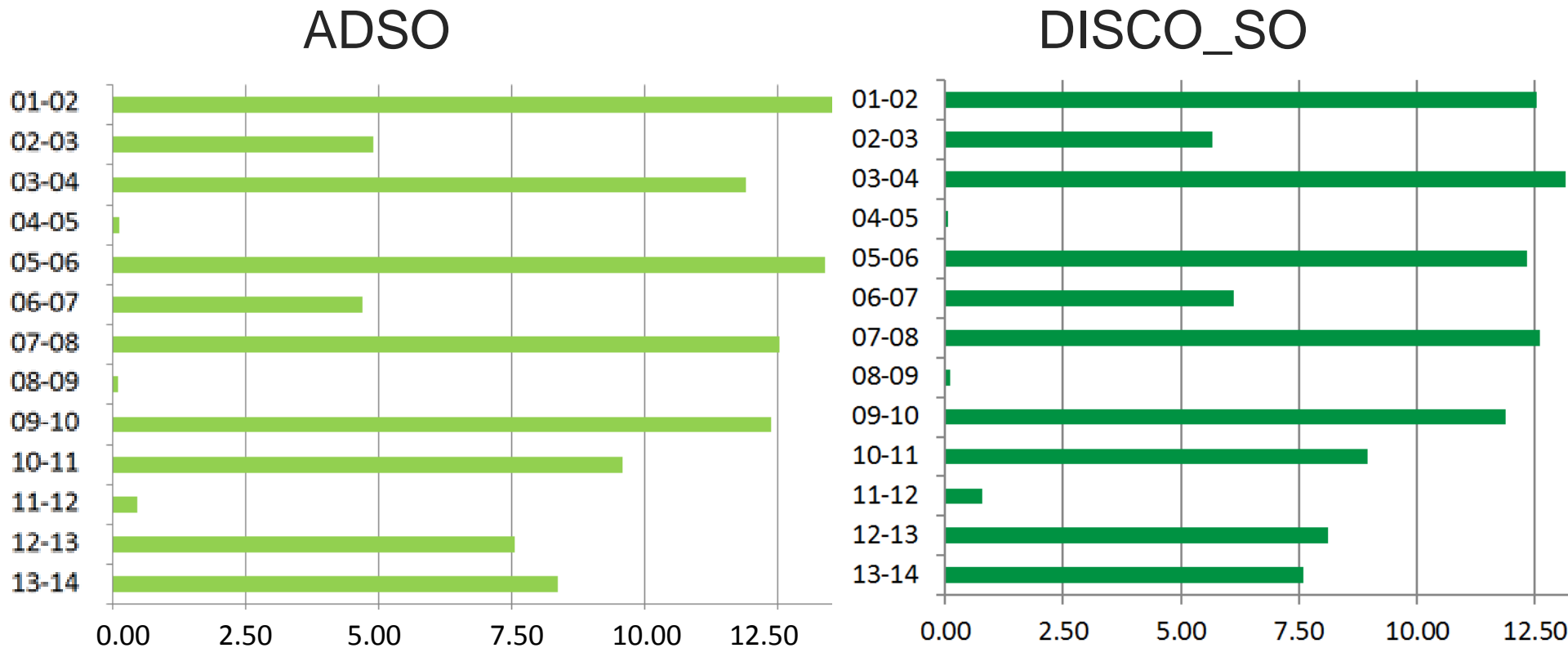
# Resumen características métricas

- Patrón ternario seguido de patrón binario  
tendencia fuerte en DISCO\_SO y ADSO
- Es una diferencia con DISCO\_XIX
- Patrones de DISCO se corresponden a los de ADSO, pero los bigramas de DISCO no predicen bien los bigramas del S. XIX

# Encabalgamiento

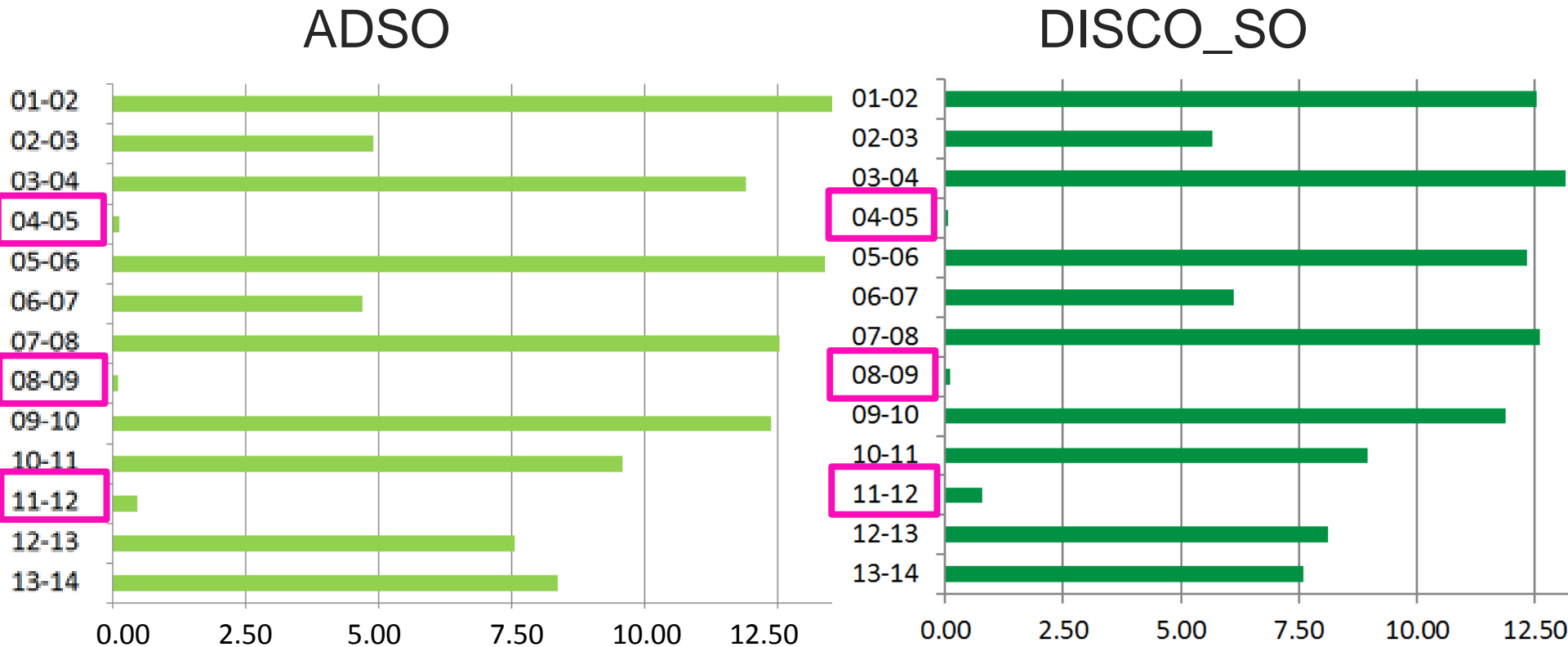
- Distribución del encabalgamiento según la posición en el soneto
  - ADSO
  - DISCO\_SO
  - DISCO\_XIX

# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

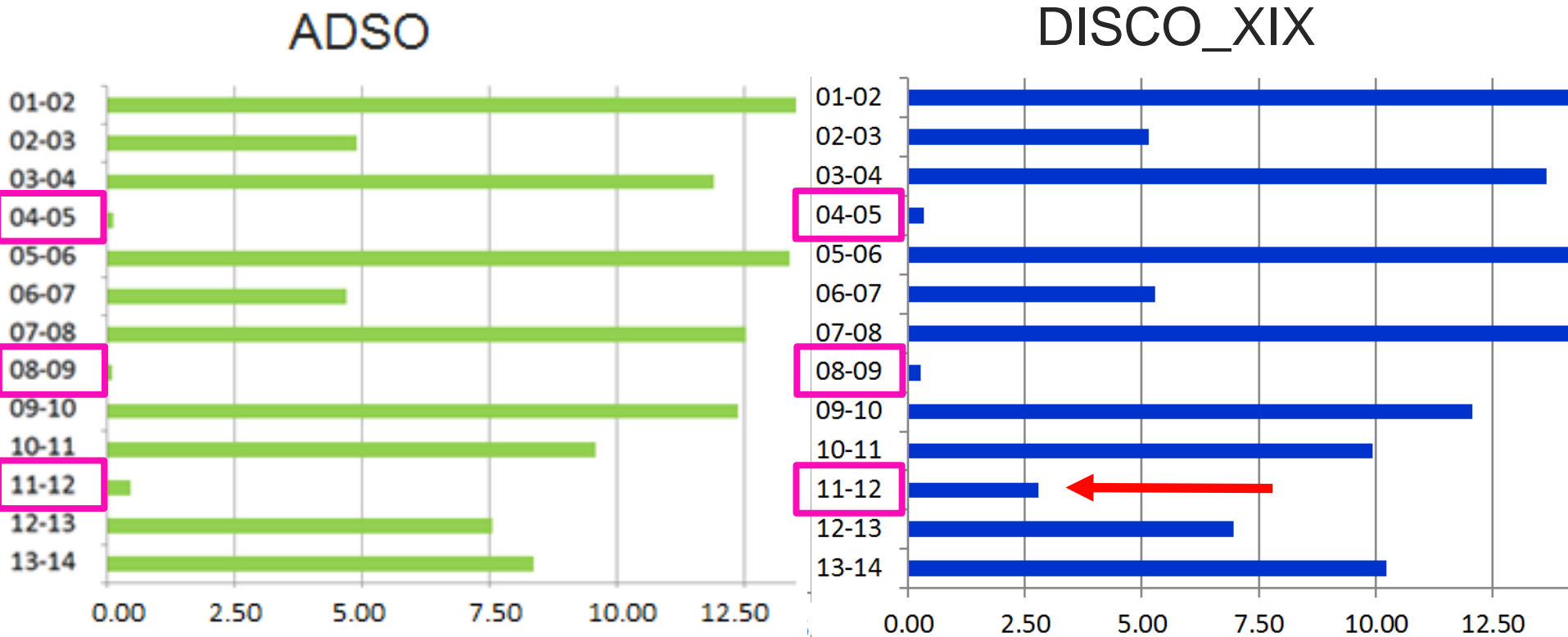
# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

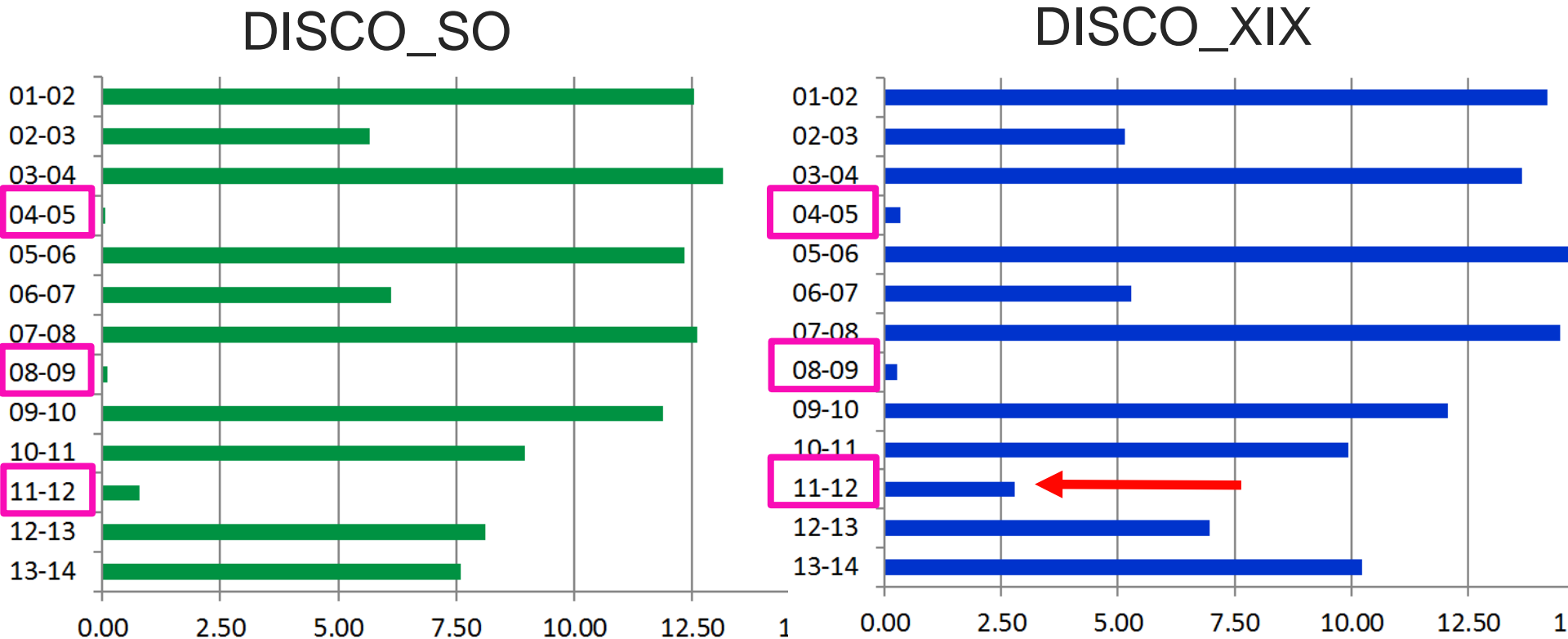


# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

# Encabalgamiento: Herrera sobre Garcilaso

dencias de las Juperiores. no dexare de traer esta aduersion, pues se ofrece lugar para ello; que cortar el verso en el Soneto, como,

Quien me dixera, cuando en las passadas  
oras —

no es vicio si no virtud, i uno de los caminos principales para alcanzar l' alteza

alteza i hermosura del estilo; como en el Eroto latino, que romper el verso es grandeza del modo de dezir: refiero esto porque se persuaden algunos, que nunca dizen mejor, que cuando siempre acaban la sentencia con la rima: i óso afirmar, que ninguna mayor falta se puede casi hallar en el soneto; que terminar los versos deste modo. porque aunque seán compuestos de letras sonantes, i de silabas llenas casi todas, parecen de muy umilde estilo i simplicidad, no por flaqueza i desmayo de letras, si no por sola esta igual manera de passo, no apartando algun verso; que siendo todo entero a acabarse en su fin, no puede tener alguna cumplida gravedad, ni alteza, ni hermosura de estilo; si bien concurriessen todas las otras partes. Pero cuando quiere alguno acompañar el estilo confor-

# Encabalgamiento: Herrera sobre Garcilaso

dencias de las Juperiores. no dexare de traer esta aduersion, pues se ofre  
ce lugar para ello; que cortar el verso en el Soneto, como,

Quien me dixera, cuando en las passadas  
oras —

no es vicio si no virtud, i uno de los caminos principales para alcanzar l  
alteza

alteza i hermosura del estilo; como en el Eroto latino, que romper el vers

**no es vicio sino virtud, y uno de los  
caminos principales para alcanzar la  
alteza y hermosura del estilo**

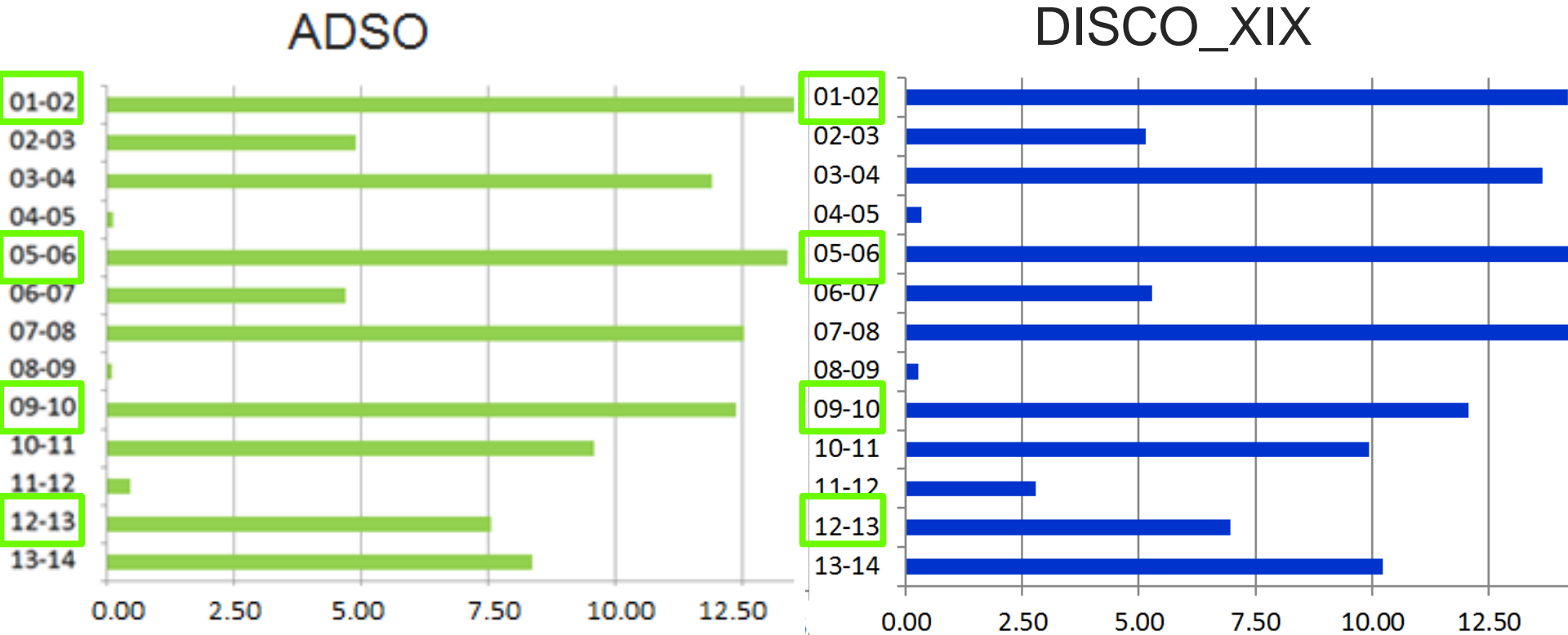


# Encabalgamiento: Herrera sobre Garcilaso

*continuo, porque engendra fastidio la perpetua semejança. quieren algunos de los que siguen esta observacion; que en el primer verso de los cuarteles i de los tercetos no tenga lugar esta incision, que la juzgan por vicio indigno de perdonar; i son ellos los que no merecen de sculpa en esto; porque antes se alcanza hermosura i variedad i grandeza. i desta suerte lo vemos en todos los que an escrito con mas arte i cuidado. Deve*

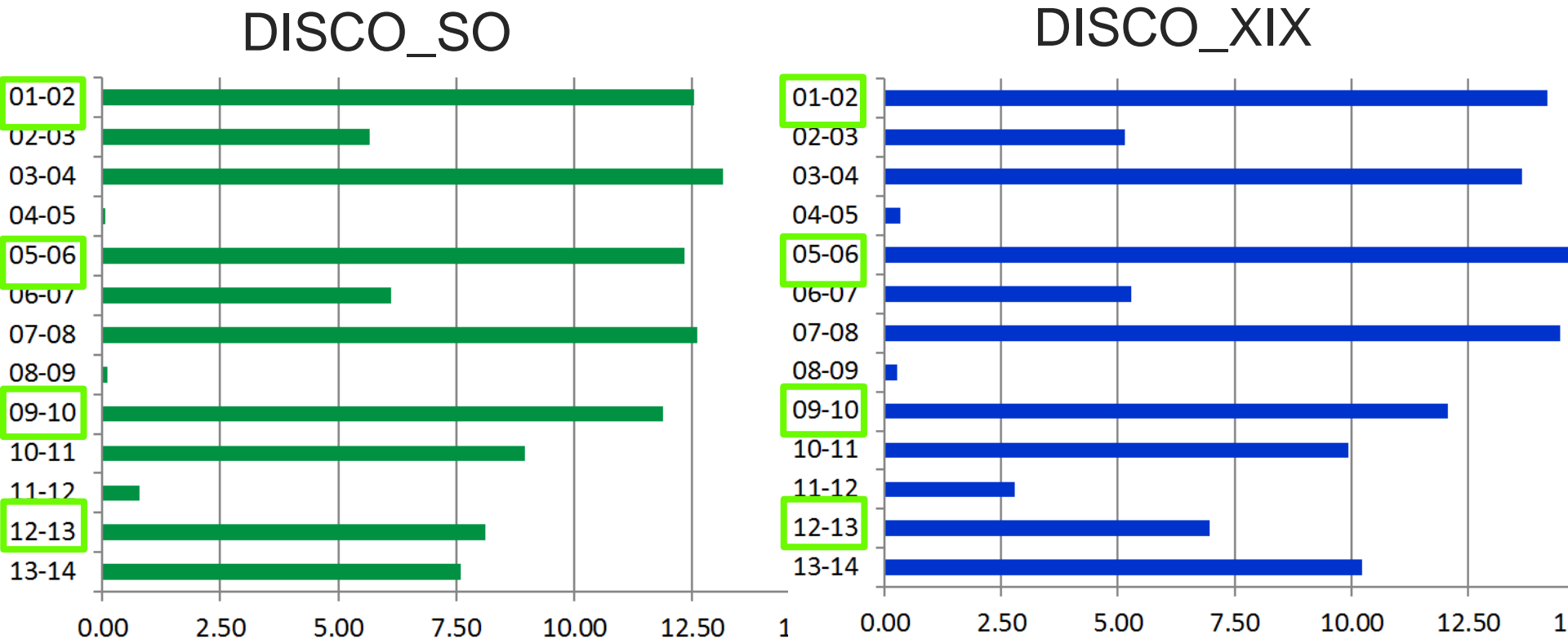
Quiere algunos de los que siguen esta observación que **en el primer verso de los cuarteles y de los tercetos** no tenga lugar esta **incisión**, que la juzgan por vicio indigno de perdonar; y **son ellos los que no merecen disculpa en esto**; porque antes se alcanza **hermosura y variedad y grandeza**, y de esta suerte lo vemos en todos los que han escrito con más arte y cuidado

# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

# Distribución de versos con encabalgamiento por período en sonetos



Porcentaje de encabalgamientos por par de versos en sonetos

# ¿Interacción métrica/encabalgamiento?

- Ejemplos en BN16 sobre patrones ternarios seguidos de binarios sugieren que puede haber una interacción ...
- Veamos dos ejemplos 3-6-10 | 2-6-10



# ¿Interacción métrica/encabalgamiento?

“De la vida asaltadas, **se estremecen atónitas** las mudas sepulturas;”

(ArgensolaBartolome\_9.xml)

“que me mata el temor, y **el desengaño me tiene** de la muerte temeroso.”

(GutierreDeCetina\_41.xml)

# ¿Interacción métrica/encabalgamiento?

- No he encontrado interacción en secuencias ternario-binario
- Proporción similar de pares con y sin encabalgamiento en estas secuencias
- Sería interesante continuar con esto

# Resumen

- Aplicamos anotación automática de métrica y encabalgamiento sobre sonetos
  - Siglo de Oro canónicos (ADSO)
  - Siglo de Oro poco conocidos (DISCO)
  - Siglo XIX (DISCO)
- Generalizaciones sobre
  - “unigramas” y “bigramas” métricos más frecuentes
  - combinación de ritmo ternario y binario
  - distribución de encabalgamiento
- Limitaciones

# Resumen: generalizaciones

- “Unigramas y bigramas métricos” más frecuentes
  - Unigramas: Similares en ADSO y DISCO\_SO
  - Bigramas:
    - ADSO refleja mejor que DISCO\_SO los bigramas métricos que se usarán después, en el Siglo XIX

# Resumen: generalizaciones

- Encabalgamiento:
  - poco usado a través de frontera de estrofa
  - si bien en el Siglo XIX esto se empieza a ver
- Interacciones entre patrones métricos y posición
  - no se han encontrado
  - sería interesante continuar

# Resumen: limitaciones

- Definición de patrón ternario y binario (que deja casos sin clasificar)
- Límite de la calidad de detección del encabalgamiento
- Poca discusión de ejemplos que ilustren bien los datos cuantitativos

**¡Gracias!**



[pablo.ruiz@linhd.uned.es](mailto:pablo.ruiz@linhd.uned.es)  
[@prf1z](https://twitter.com/prf1z)

<http://postdata.linhd.es/>  
<http://linhd.uned.es>