Segmentation strategies for inflection class inference

Sacha Beniamine (LLF), Benoît Sagot (Alpage)
Université Paris Diderot

Décembrettes 9, Toulouse, 2015

- Concept of Inflection Classes widely used to analyse inflectional systems
  - The definition of IC is crucial for many linguistic and psycholinguistic studies, yet they are often taken for granted.

- Concept of Inflection Classes widely used to analyse inflectional systems
  - The definition of IC is crucial for many linguistic and psycholinguistic studies, yet they are often taken for granted.
- **No consensus** on how to obtain the classification

- Concept of Inflection Classes widely used to analyse inflectional systems
  - The definition of IC is crucial for many linguistic and psycholinguistic studies, yet they are often taken for granted.
- **No consensus** on how to obtain the classification
- **We explore the concept through computational means**: Brown and Evans, 2012; Lee and Goldsmith, 2013; Bonami, 2014
  - Formal definitions of the concept
  - Large datasets
  - Reproducible classifications
  - Commensurable across languages
  - Basis for theoretical and typological comparisons

Groups of lexemes that inflect alike.

	INF	PRES.3.SG	PRES.3.PL	PP
TENIR ‘hold’	təniʁ	tjɛ̃	tjɛn	təny
FINIR ‘finish’	finiʁ	fini	finis	fini
HAÏR ‘hate’	aiʁ	ɛ	ais	ai
PELER ‘peel’	pəle	pɛl	pɛl	pəle
LAVER ‘wash’	lave	lav	lav	lave
TASSER ‘press’	tase	tas	tas	tase

Groups of lexemes that inflect alike .

	INF	PRES.3.SG	PRES.3.PL	PP
TENIR ‘hold’	tənir	tjɛ̃	tjɛn	təny
FINIR ‘finish’	finir	fini	finis	fini
HAÏR ‘hate’	air	ɛ	ais	ai
PELER ‘peel’	pəle	pɛl	pɛl	pəle
LAVER ‘wash’	lave	lav	lav	lave
TASSER ‘press’	tase	tas	tas	tase

What is needed to infer IC from paradigmatic data*.

1. What form should an IC system take?
2. What Inflectional Realisations should we infer from the data?
3. How do we measure which lexemes inflect alike?
4. How do we find the best classes among all possible ones?

Table of Contents

Inflection classes: Cohesive or distinctive?

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:
  - ***Cohesive**: Maximal homogeneity within classes*

Inflection classes: Cohesive or distinctive?

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:
  - ***Cohesive***: *Maximal homogeneity within classes*
  - ***Distinctive***: *Maximal heterogeneity between classes*

Inflection classes: Cohesive or distinctive?

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:
  - ***Cohesive***: *Maximal homogeneity within classes*
  - ***Distinctive***: *Maximal heterogeneity between classes*
- In most languages, each of these criteria leads to different partitions:

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR 'hold'	təniʁ	tjɛ̃	tjɛn	təny
FINIR 'finish'	finiʁ	fini	finis	fini
HAÏR 'hate'	aiʁ	ɛ	ais	ai
PELER 'peel'	pəle	pɛl	pɛl	pəle
LAVER 'wash'	lave	lav	lav	lave
TASSER 'press'	tase	tas	tas	tase

Inflection classes: Cohesive or distinctive?

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:
  - ***Cohesive***: *Maximal homogeneity within classes*
  - ***Distinctive***: *Maximal heterogeneity between classes*
- In most languages, each of these criteria leads to different partitions:
  - **favouring cohesion**: *numerous small, similar classe*

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR 'hold'	tənir	tjɛ̃	tjɛn	təny
FINIR 'finish'	finir	fini	finis	fini
HAÏR 'hate'	air	ɛ	ais	ai
PELER 'peel'	pəle	pɛl	pɛl	pəle
LAVER 'wash'	lave	lav	lav	lave
TASSER 'press'	tase	tas	tas	tase

Inflection classes: Cohesive or distinctive?

- Insight from Canonical Typology (Corbett, 2009).
  An ideal inflection class system is a partition of the set of lexemes that is:
  - ***Cohesive***: *Maximal homogeneity within classes*
  - ***Distinctive***: *Maximal heterogeneity between classes*
- In most languages, each of these criteria leads to different partitions:
  - **favouring cohesion**: *numerous small, similar classe*
  - **favouring distinction**: *fewer large classes with exceptions*

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR 'hold'	tənir	tjɛ̃	tjɛn	təny
FINIR 'finish'	finir	fini	finis	fini
HAÏR 'hate'	air	ɛ	ais	ai
PELER 'peel'	pəle	pɛl	pɛl	pəle
LAVER 'wash'	lave	lav	lav	lave
TASSER 'press'	tase	tas	tas	tase

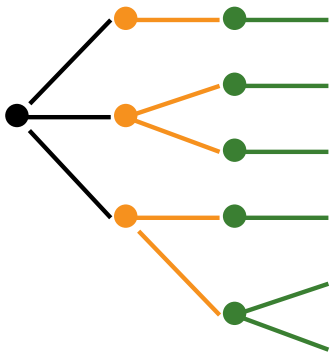
Inflection classes: Macro and microclasses?

- Dressler and Thornton's terminology (1996):
- **Micro-classes**
  - *Numerous small, similar classes.*
- **Macro-classes**
  - *Fewer large classes with exceptions.*

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR 'hold'	təniʁ	tjɛ̃	tjɛn	təny
FINIR 'finish'	finiʁ	fini	finis	fini
HAÏR 'hate'	aiʁ	ɛ	ais	ai
PELER 'peel'	pəle	pɛl	pɛl	pəle
LAVER 'wash'	lave	lav	lav	lave
TASSER 'press'	tase	tas	tas	tase

Inflection classes: Macro and microclasses?

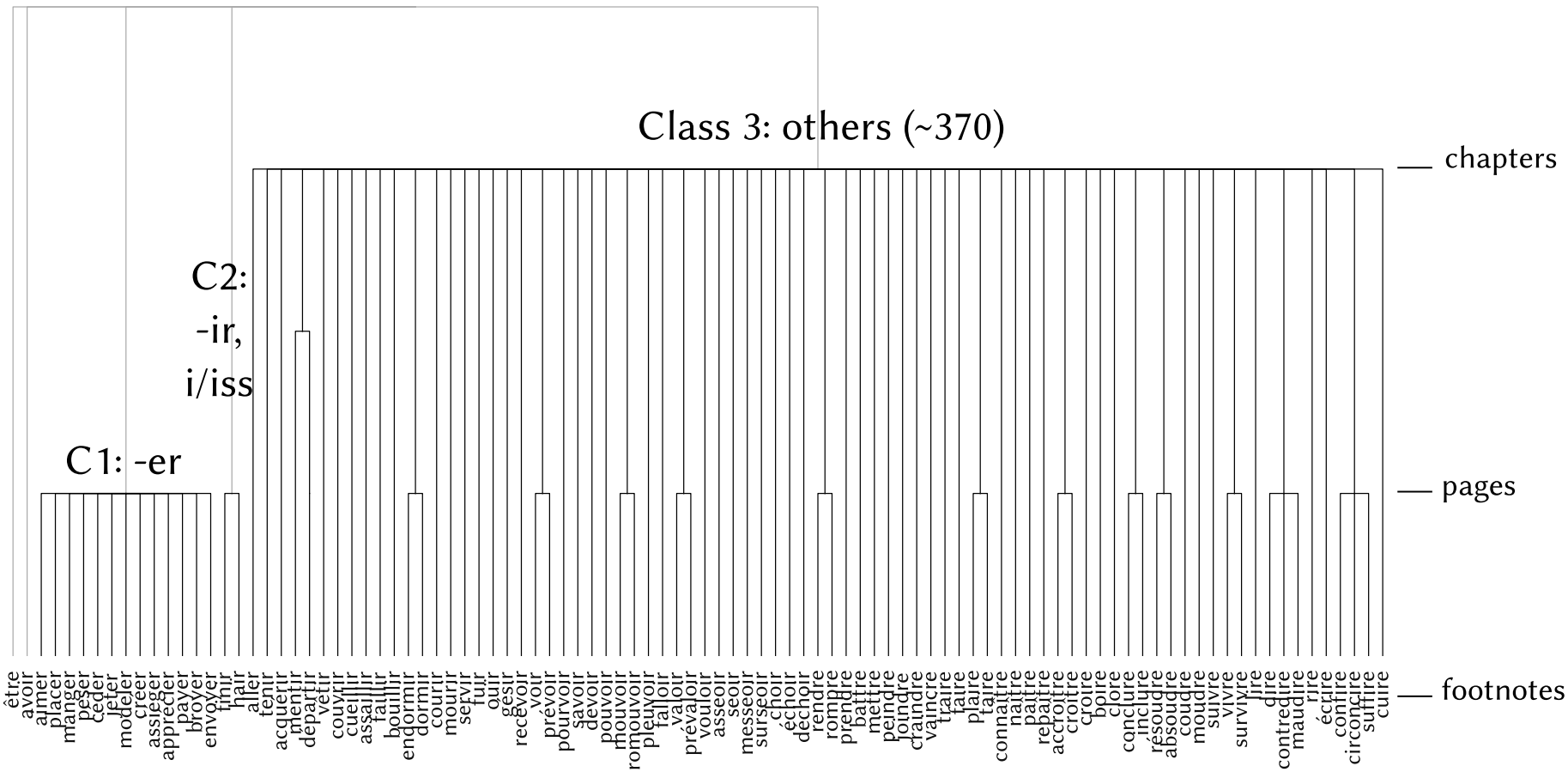
- Dressler and Thornton's terminology (1996):
- **Micro-classes**
  - *Numerous small, similar classes.*
- **Macro-classes**
  - *Fewer large classes with exceptions.*
- Combined in a hierarchy. (Corbett and Fraser, 1993; Dressler and Thornton, 1996; Brown and Evans, 2012)



Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR 'hold'	tənir	tjɛ̃	tjɛn	təny
FINIR 'finish'	finir	fini	finis	fini
HAÏR 'hate'	air	ɛ	ais	ai
PELER 'peel'	pəle	pɛl	pɛl	pəle
LAVER 'wash'	lave	lav	lav	lave
TASSER 'press'	tase	tas	tas	tase

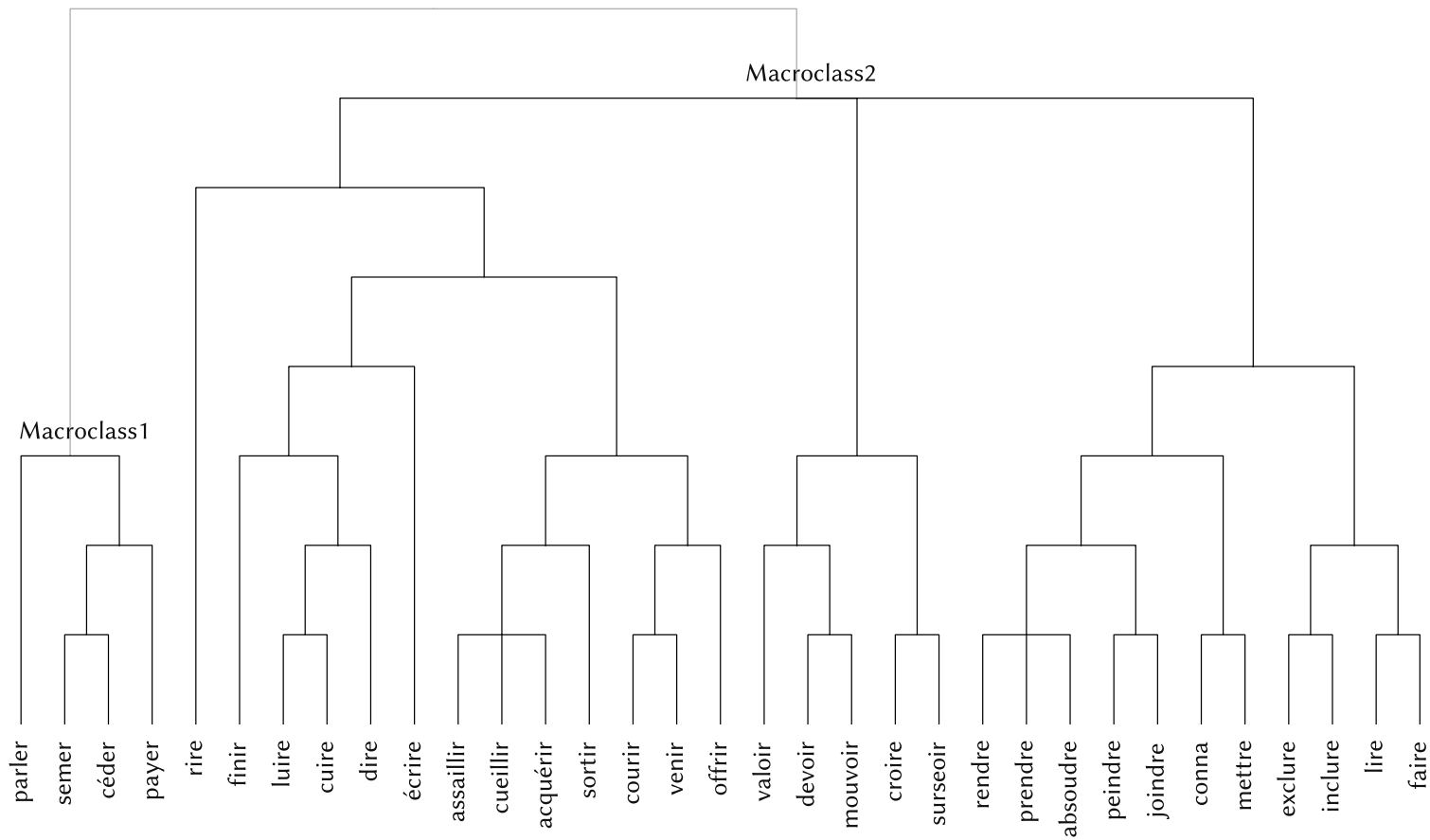
The example of french verbal inflection

- School grammar (Bescherelle) :



The example of french verbal inflection

- School grammar (Bescherelle)
- Kilani-Schoch and Dressler, 2005: different microclasses, some dropped, two macroclasses (dual route).



Inflection classes: Macro and microclasses?

- **Micro-classes**
  - Homogenous: *Numerous small, similar classes.*
  - Inventories vary across accounts.
  - Empirically motivated
- **Macro-classes**
  - Heterogenous: *Fewer large classes with "exceptions".*
  - High variation across accounts.
  - Empirical motivation in question:

- **Micro-classes**
  - Homogenous: *Numerous small, similar classes.*
  - Inventories vary across accounts.
  - Empirically motivated
- **Macro-classes**
  - Heterogenous: *Fewer large classes with "exceptions".*
  - High variation across accounts.
  - Empirical motivation in question:

Are macroclasses a descriptive artefact?

Table of Contents

- **Stem and exponents**
  - *Captures differences between cells under the assumption of a* ***constant stem****.*
  - *cf. (Blevins, 2006)'s notion of constructive approach.*

- **Stem and exponents**
  - *Captures differences between cells under the assumption of a* ***constant stem****.*
  - *cf. (Blevins, 2006)'s notion of constructive approach.*

- **Binary alternation patterns**
  - *Captures the* ***implicative relation*** *between each pair of cells.*
  - *cf. (Blevins, 2006)'s notion of abstractive approach.*

- **Stem and exponents**
  - *Captures differences between cells under the assumption of a* ***constant stem****.*
  - *cf. (Blevins, 2006)'s notion of constructive approach.*

- **Binary alternation patterns**
  - *Captures the* ***implicative relation*** *between each pair of cells.*
  - *cf. (Blevins, 2006)'s notion of abstractive approach.*
- Both rely on a **segmentation** of forms.
  - global segmentation over the whole paradigm.
  - local segmentation over pairs of forms.

- **Global**: On the basis of a whole paradigm.
- **Local**: On each pair of cells.

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR ‘hold’	tənir	tjɛ̃	tjɛn	təny
FINIR ‘finish’	finir	fini	finis	fini
HAÏR ‘hate’	air	ɛ	ais	ai
PELER ‘peel’	pəle	pɛl	pɛl	pəle
LAVER ‘wash’	lave	lav	lav	lave
TASSER ‘press’	tase	tas	tas	tase

- Global: On the basis of a whole paradigm.
- **Local**: On each pair of cells.

Lexeme	INF	PRS.3SG	PRS.3PL	PST.PTCP
TENIR ‘hold’	Xəniʁ	Xjɛ̃	Xjɛn	Xəny
FINIR ‘finish’	Xʁ	X	Xs	X
HAÏR ‘hate’	aiʁ	ɛ	ais	ai
PELER ‘peel’	X_1əX_2e	X_1ɛX_2	X_1ɛX_2	X_1əX_2e
LAVER ‘wash’	Xe	X	X	Xe
TASSER ‘press’	Xe	X	X	Xe

- **Global**: On the basis of a whole paradigm.
- Local: On each pair of cells.

Lexeme	INF⇌ PRS.3SG	INF⇌ PRS.3PL	INF⇌ PST.PTCP	…
TENIR ‘hold’	Xəniʁ ⇌ Xjɛ̃	Yəniʁ ⇌ Yjɛn	Ziʁ ⇌ Zy	
FINIR ‘finish’	Xʁ ⇌ X	Yʁ ⇌ Ys	Zʁ ⇌ Z	
HAÏR ‘hate’	aiʁ ⇌ ɛ	Yʁ ⇌ Ys	Zʁ ⇌ Z	…
PELER ‘peel’	X_1əX_2e ⇌ X_1ɛX_2	Y_1əY_2e ⇌ Y_1ɛY_2	Z ⇌ Z	
LAVER ‘wash’	Xe ⇌ X	Ye ⇌ Y	Z ⇌ Z	
TASSER ‘press’	Xe ⇌ X	Ye ⇌ Y	Z ⇌ Z	

- In general, grouping elements into classes is a clustering problem.
- There are many well-known solutions in computer science to address such problems.
- All of them require two things:
  - A criterion to evaluate the quality of clusters (classes).
  - An algorithm to explore the search space of all possible groupings.

- In general, grouping elements into classes is a clustering problem.
- There are many well-known solutions in computer science to address such problems.
- All of them require two things:
  - A criterion to evaluate the quality of clusters (classes).
    $\rightarrow$ *Minimum description length*
  - An algorithm to explore the search space of all possible groupings.

- In general, grouping elements into classes is a clustering problem.
- There are many well-known solutions in computer science to address such problems.
- All of them require two things:
  - A criterion to evaluate the quality of clusters (classes).
    $\rightarrow$ *Minimum description length*
  - An algorithm to explore the search space of all possible groupings.
    $\rightarrow$ *Greedy bottom-up algorithm*

Table of Contents

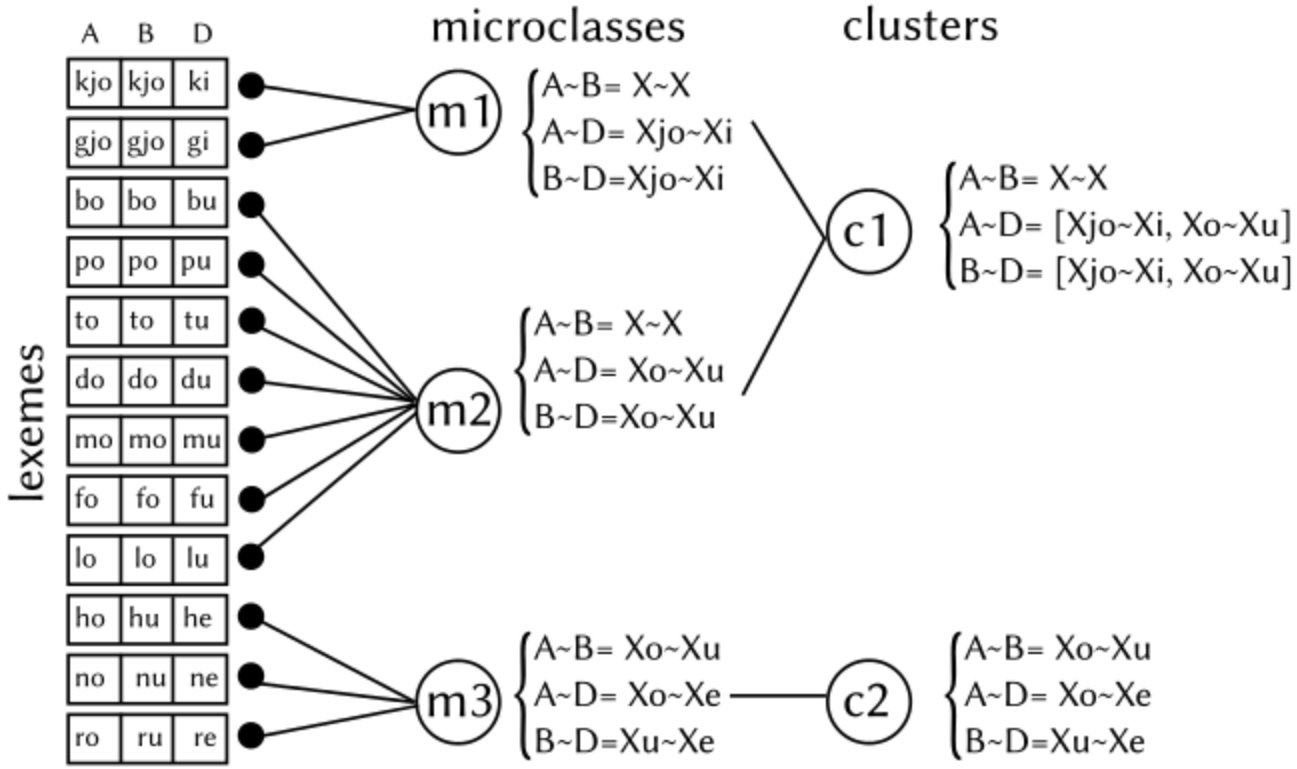
- **Minimum description length** (Rissanen, 1984): Choose the model allowing for the shortest description of the data.
- A partition of the set of lexemes is better than another one if it leads to a more economical description of the system. (Sagot and Walther, 2011; Walther, 2013)

$$DL(system) =$$

$$number\ of\ symbols \times \underbrace{- \sum_{x \in symbols} P(x) \times log_2\left(P(x)\right)}_{\text{entropy}}$$

Description length of a partition of the set of lexemes

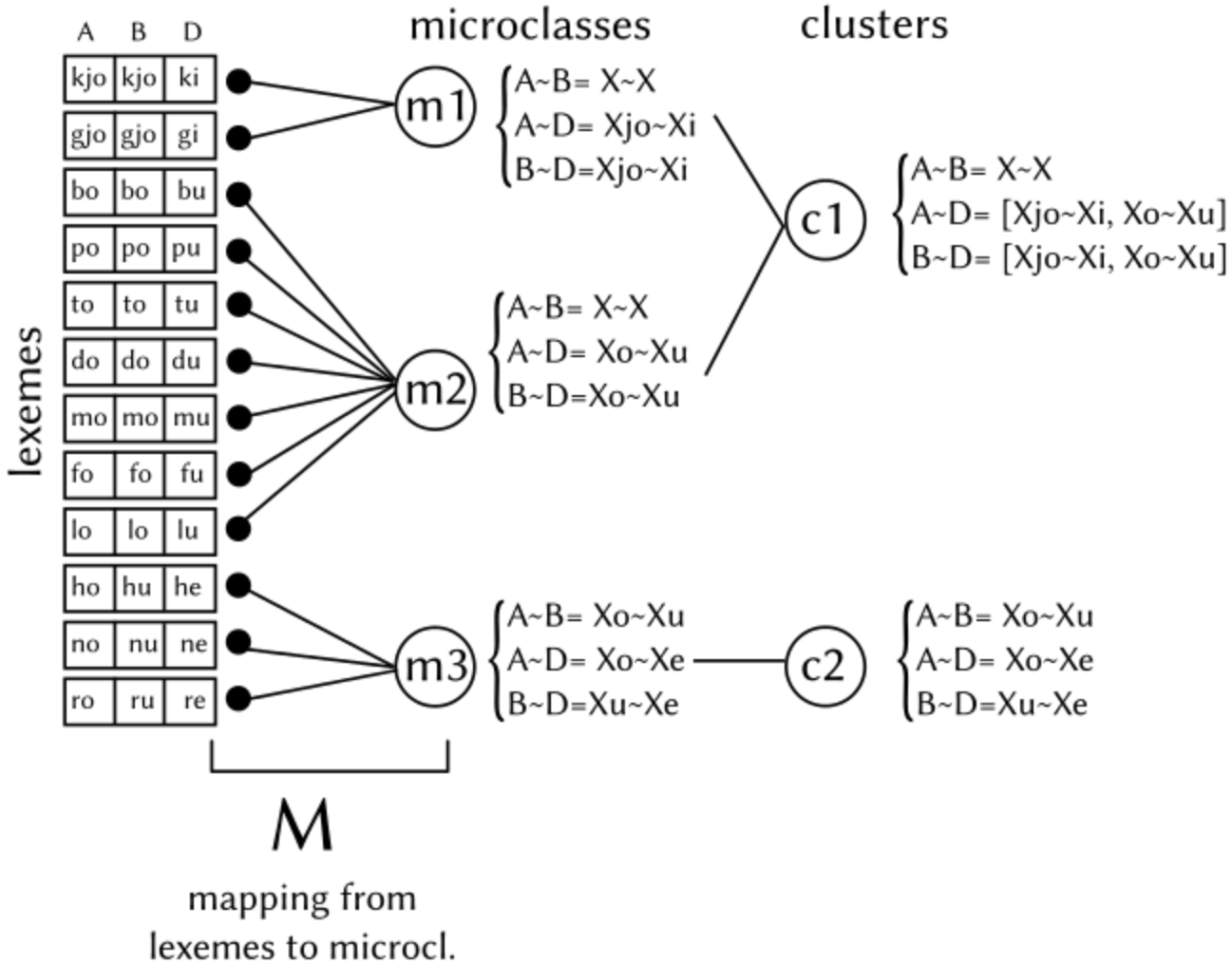
- We break down the description length into four components:



Toy imaginary dataset with three cells A, B and D.

Description length of a partition of the set of lexemes

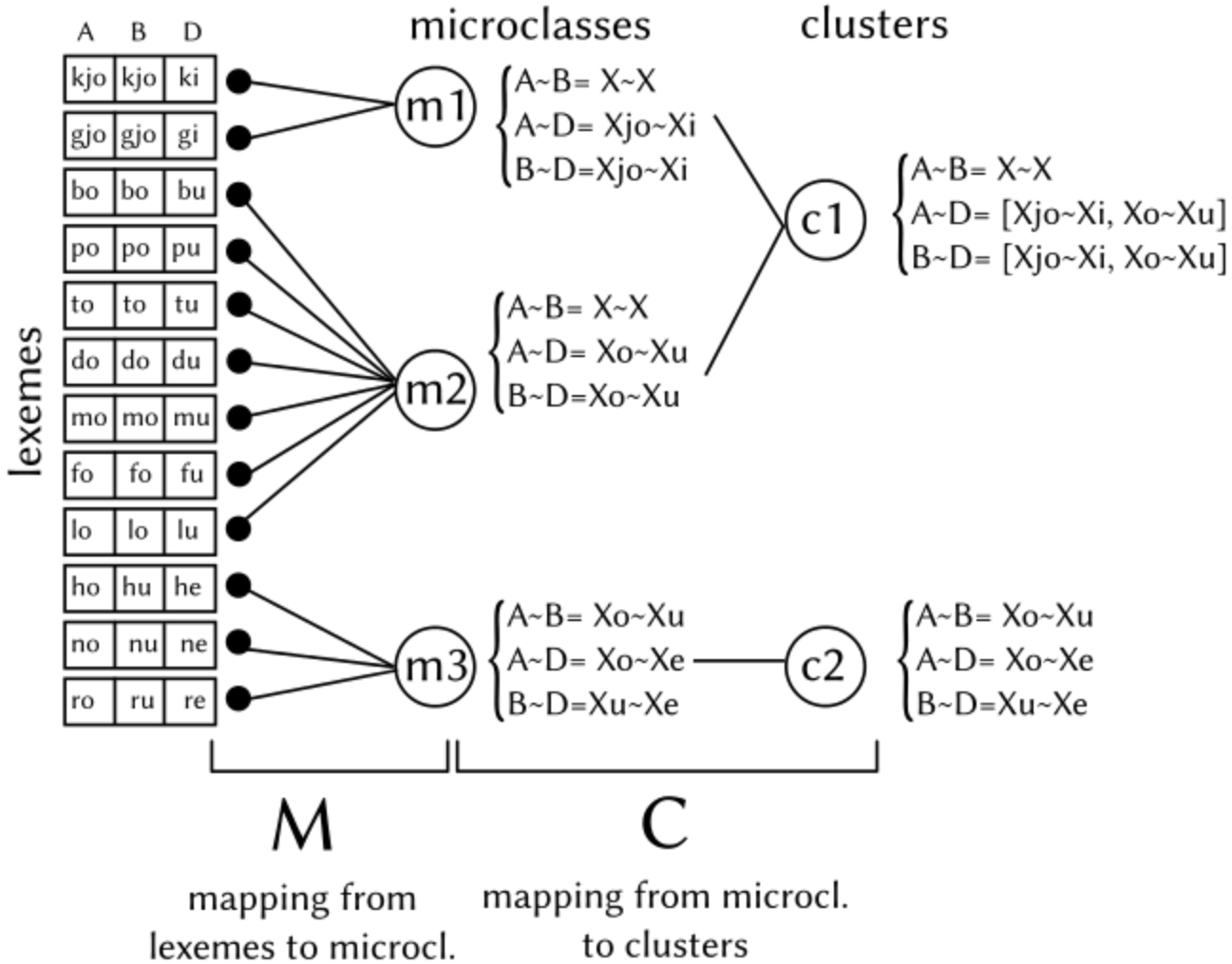
- We break down the description length into four components:



Toy imaginary dataset with three cells A, B and D.

Description length of a partition of the set of lexemes

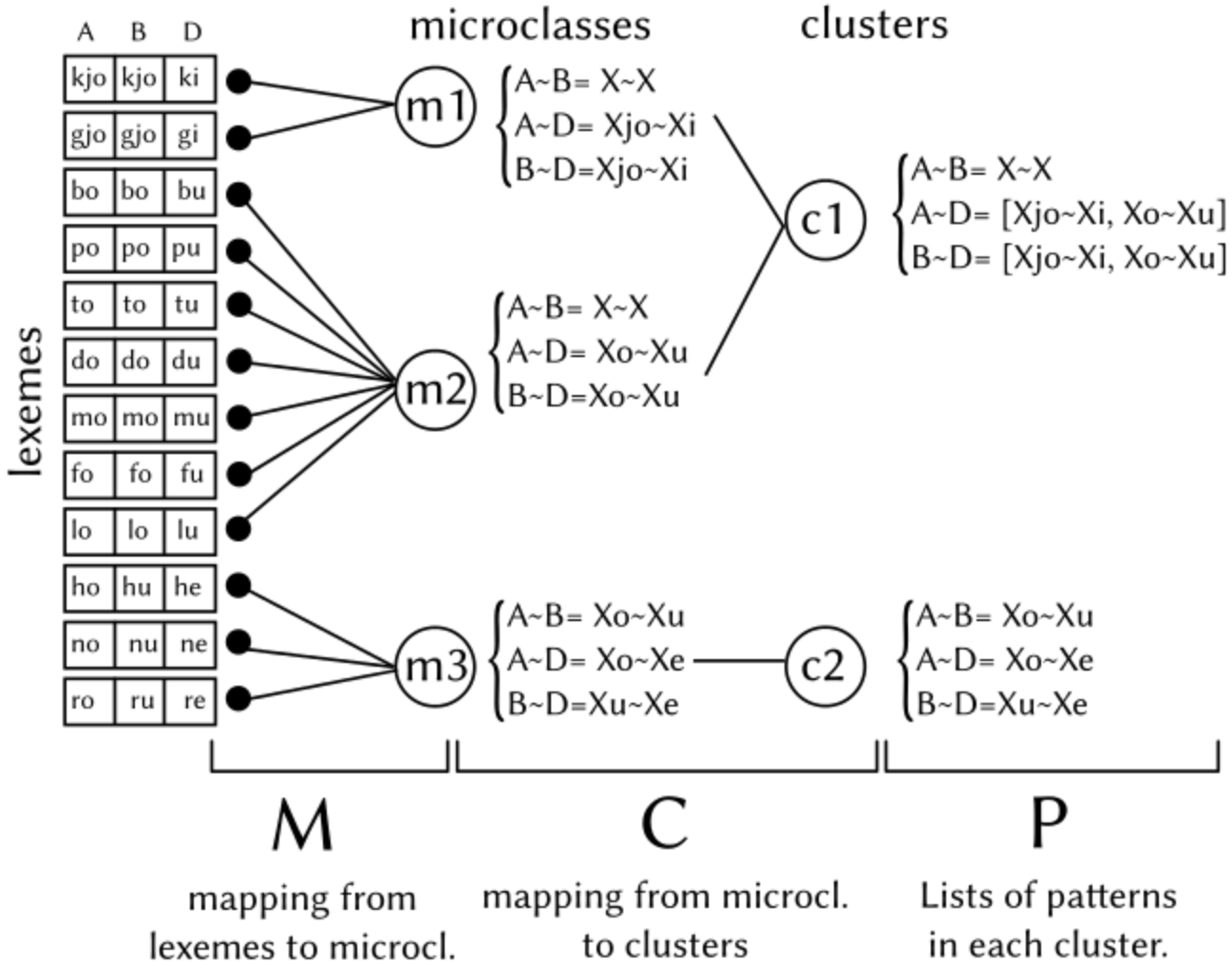
- We break down the description length into four components:



Toy imaginary dataset with three cells A, B and D.

Description length of a partition of the set of lexemes

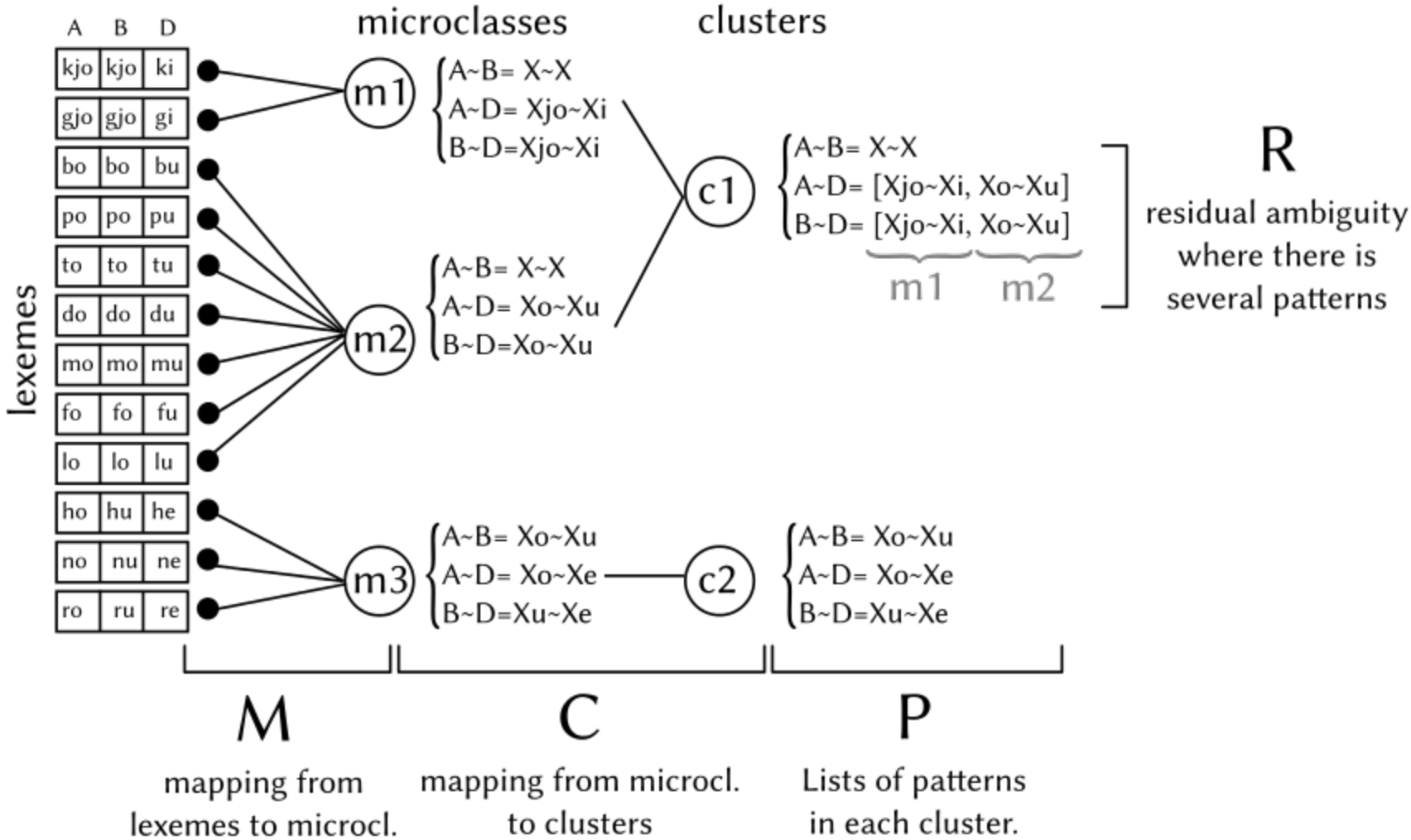
- We break down the description length into four components:



Toy imaginary dataset with three cells A, B and D.

Description length of a partition of the set of lexemes

- We break down the description length into four components:



Toy imaginary dataset with three cells A, B and D.

Description length of a partition of the set of lexemes

- We break down the description length into four components:

$$DL = M + C + P + R$$

Table of Contents

Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.

passar	pagar	chegar	achar	abandonar	voar	desempenhar	nomear	levar	jogar	ficar	ganhar
(155)	(1)	(20)	(3)	(12)	(17)	(4)	(53)	(162)	(177)	(911)	(1)

Clustering algorithm, ex. on European Portuguese conjugation.

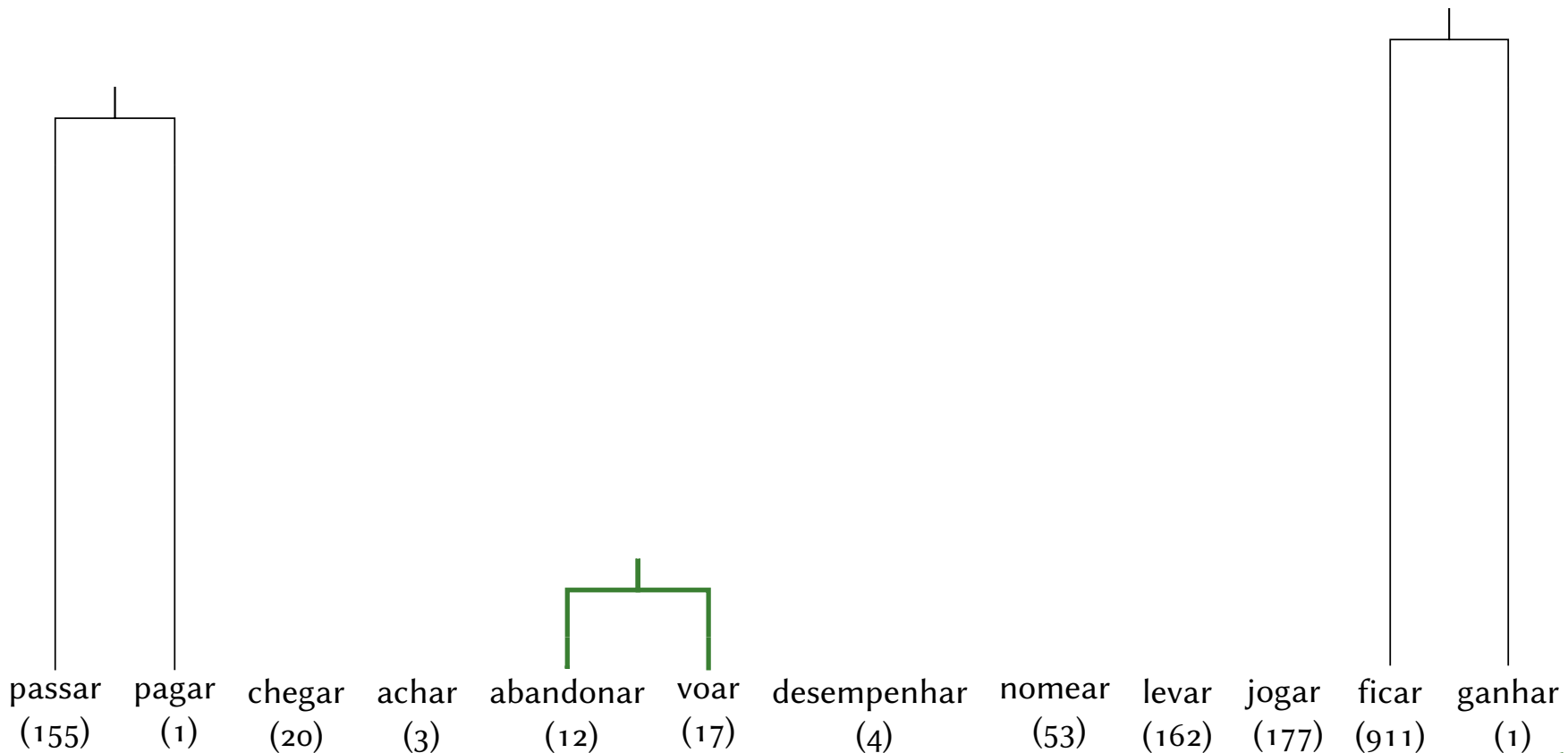
(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.

Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.

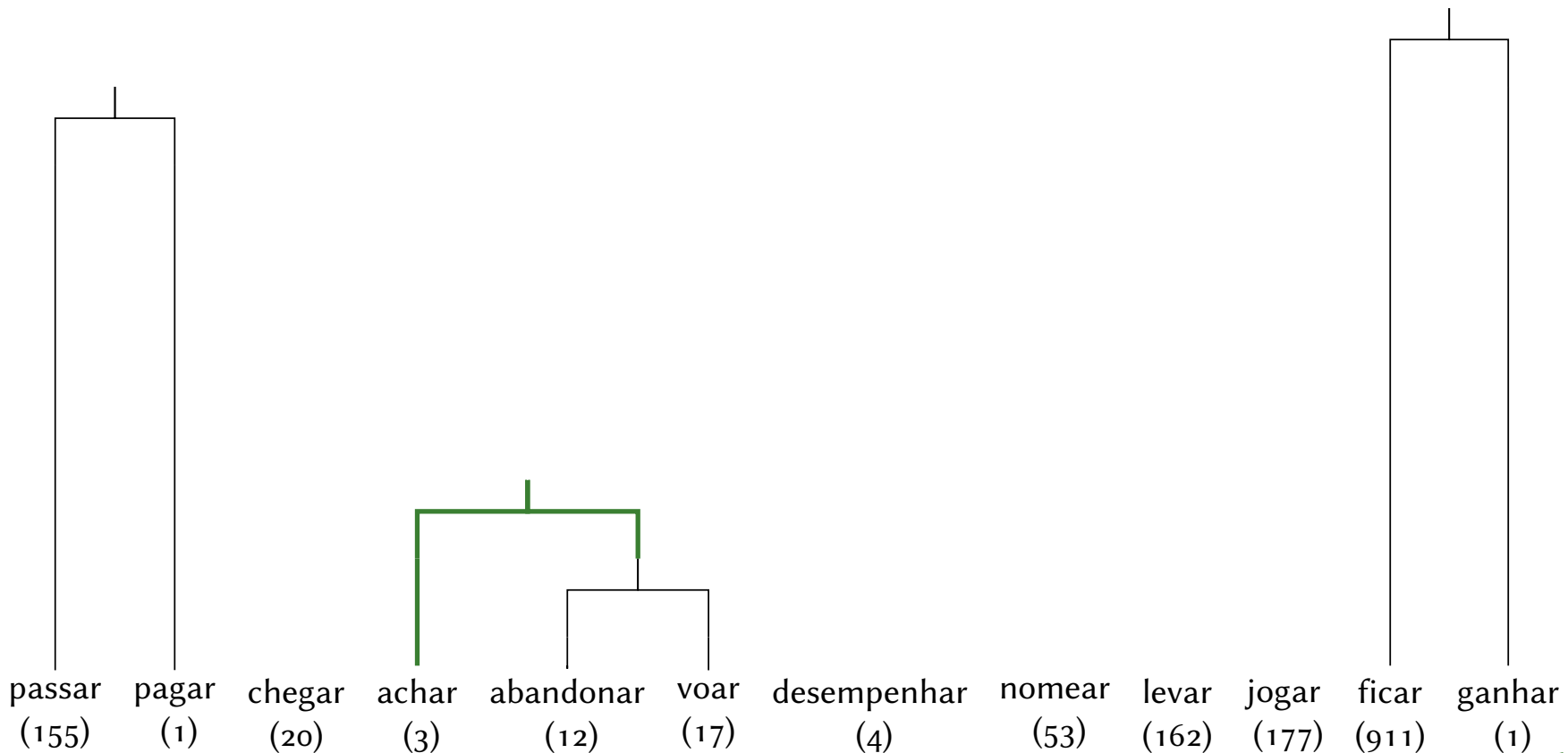
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



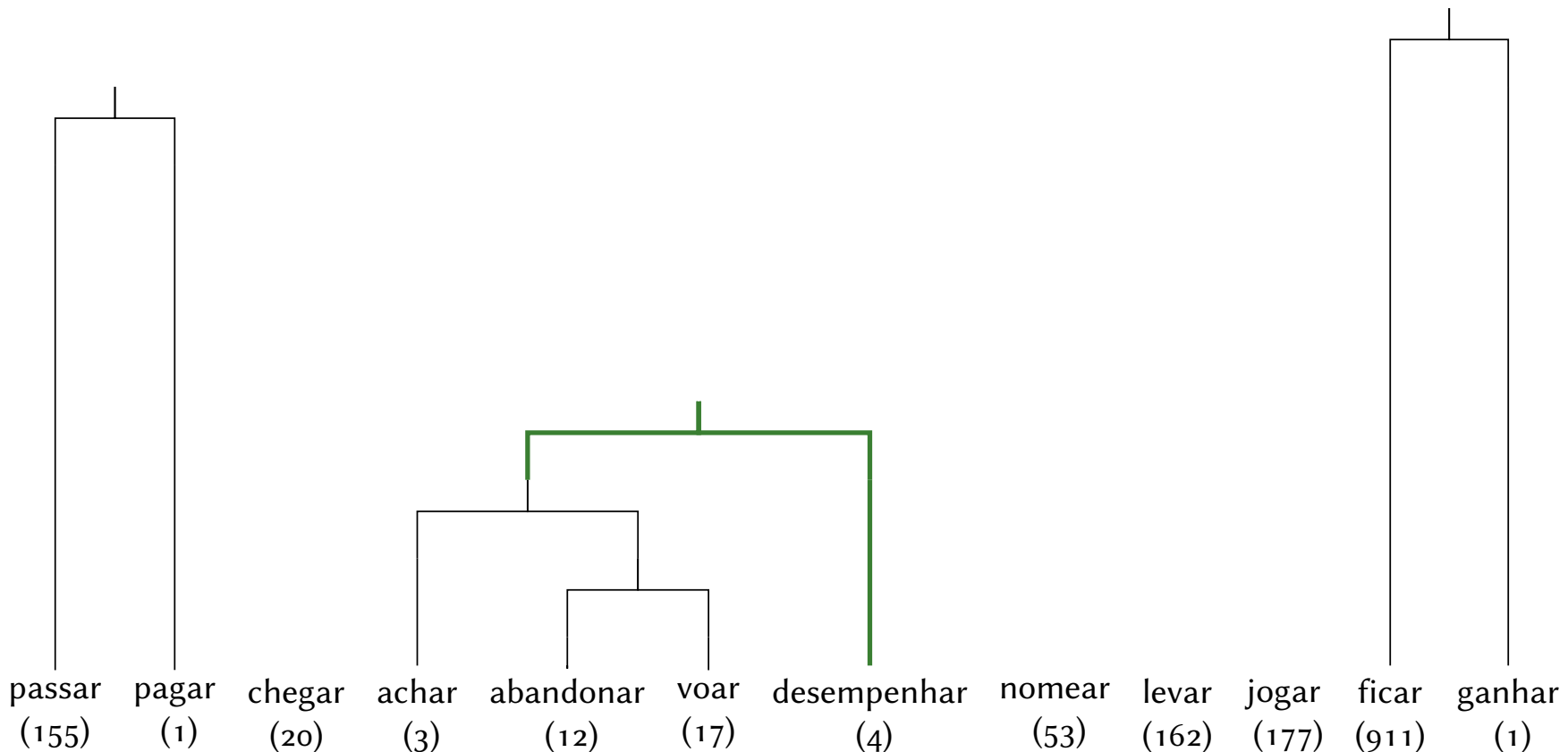
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



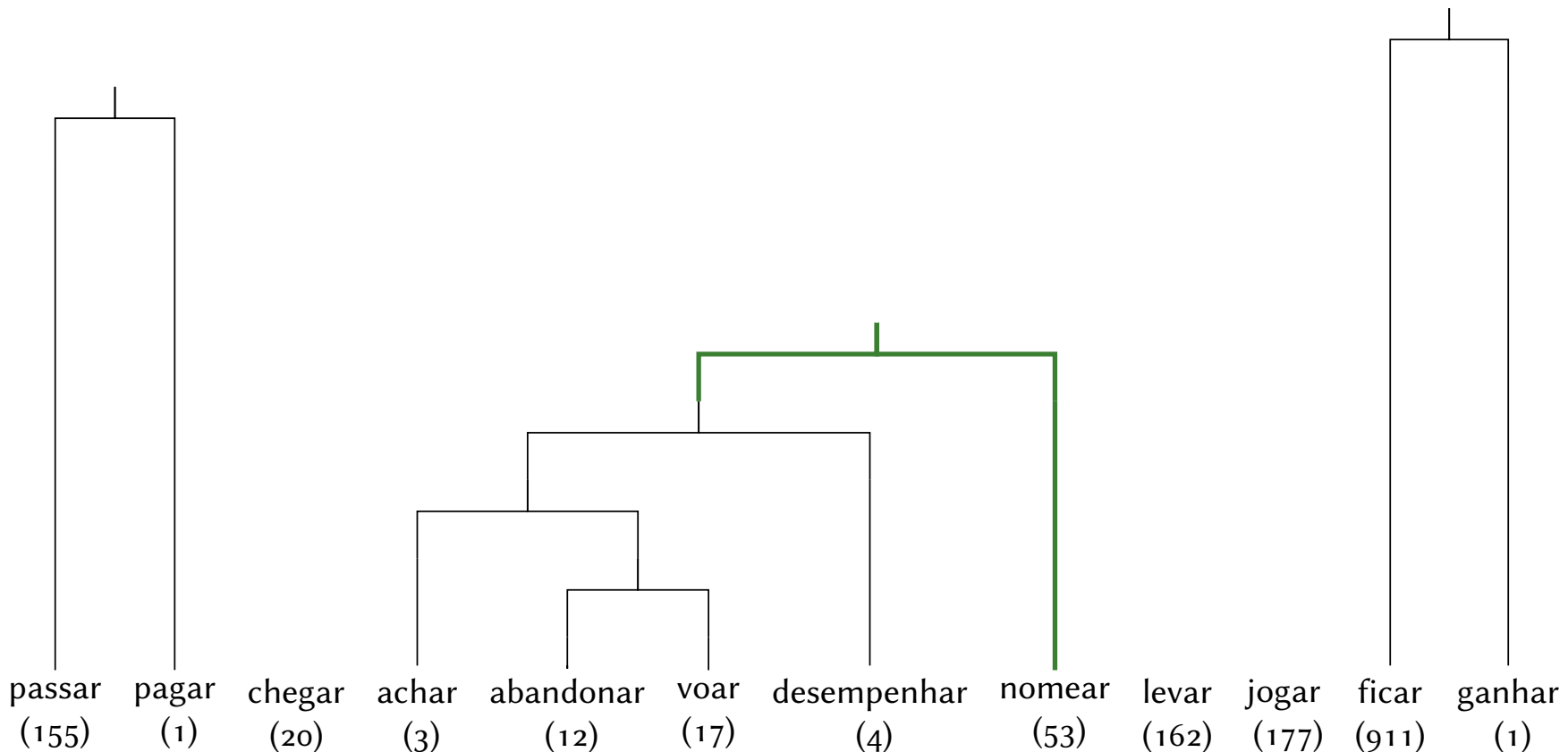
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



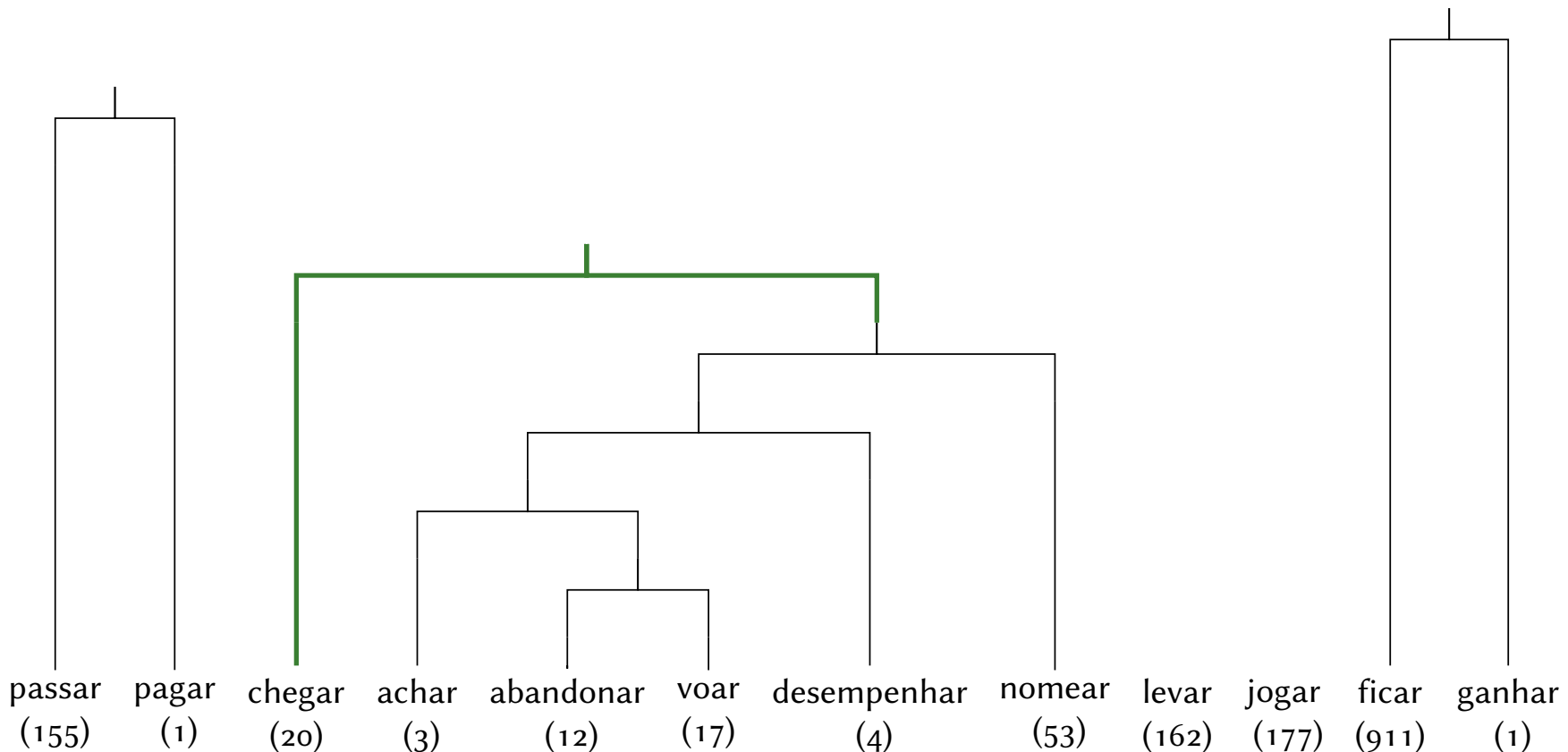
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



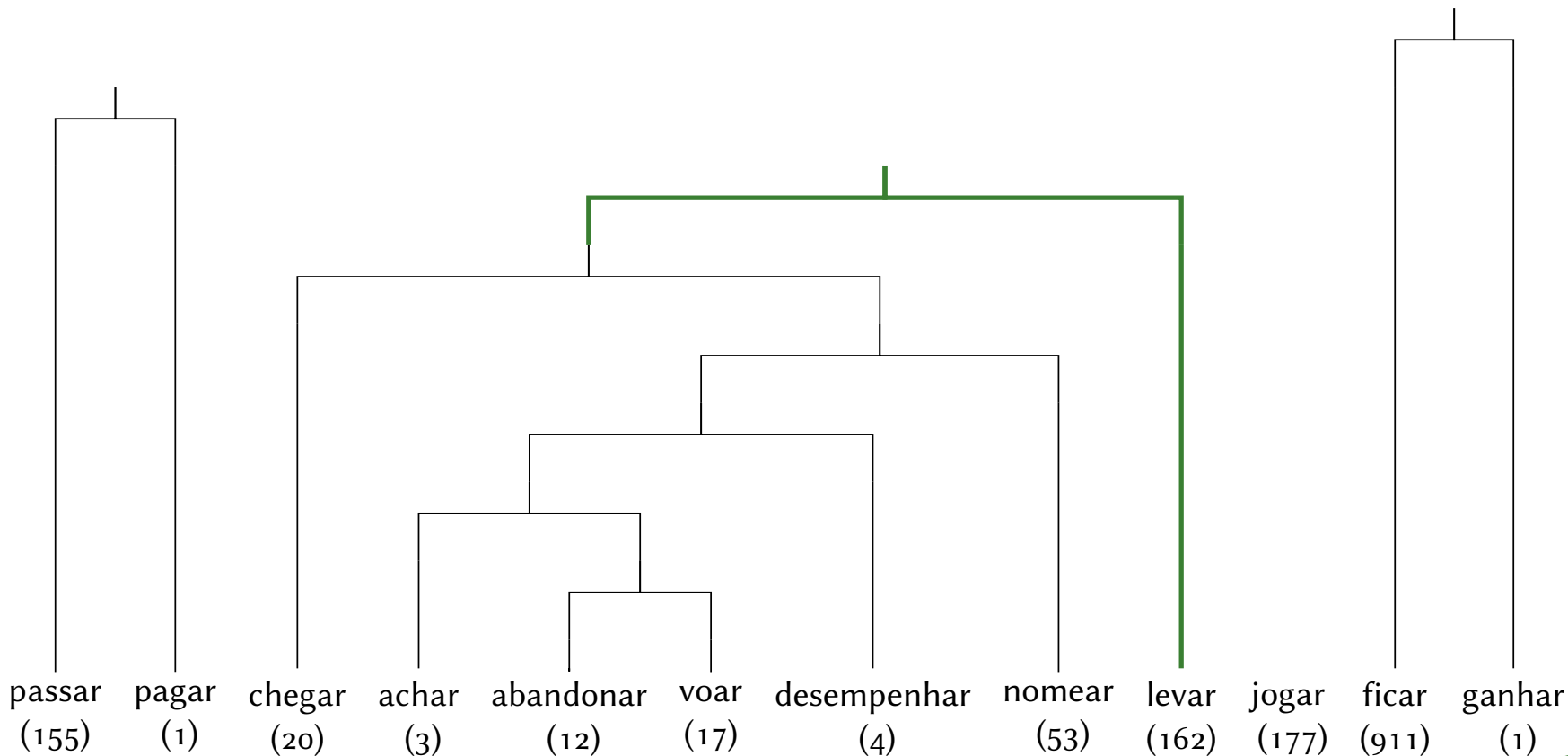
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



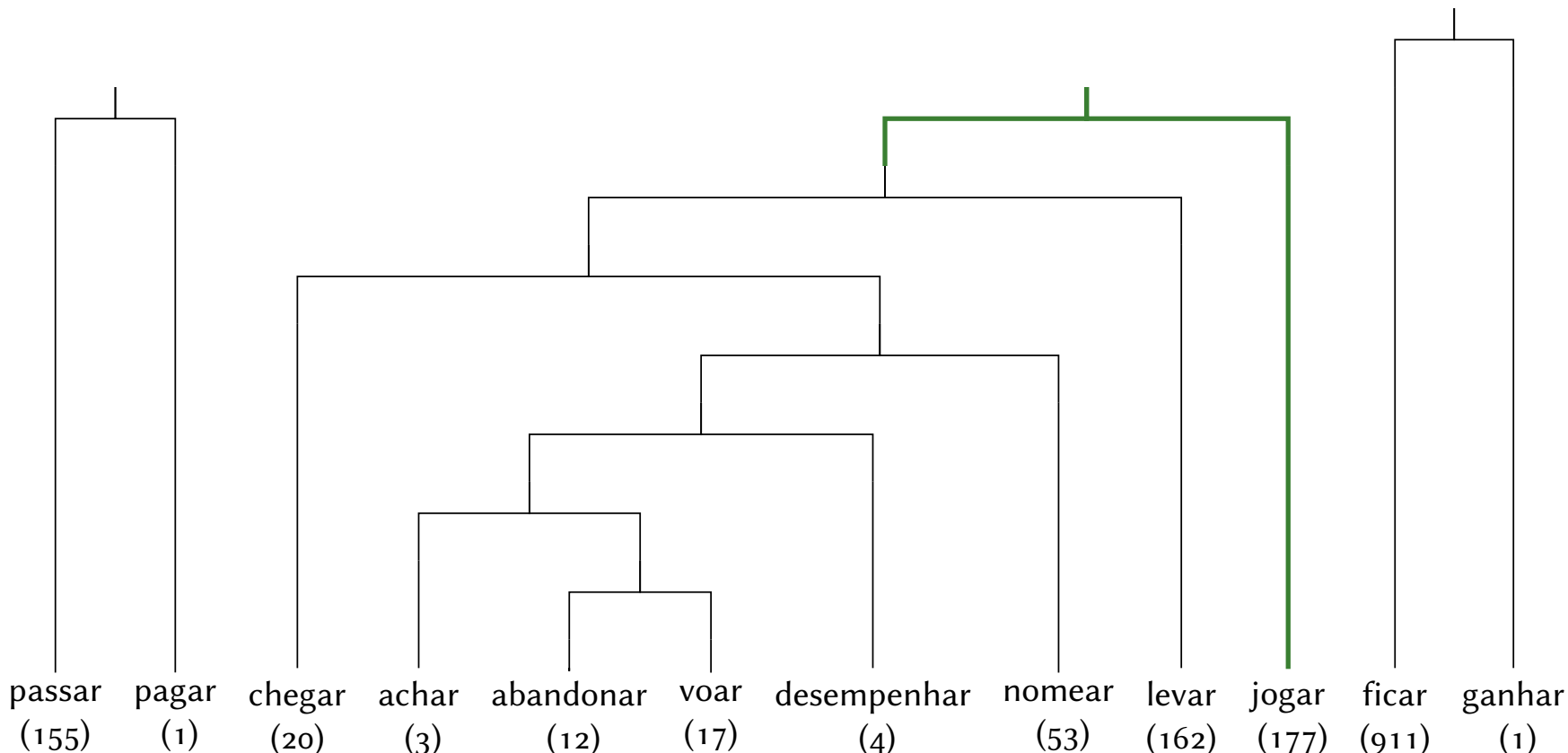
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



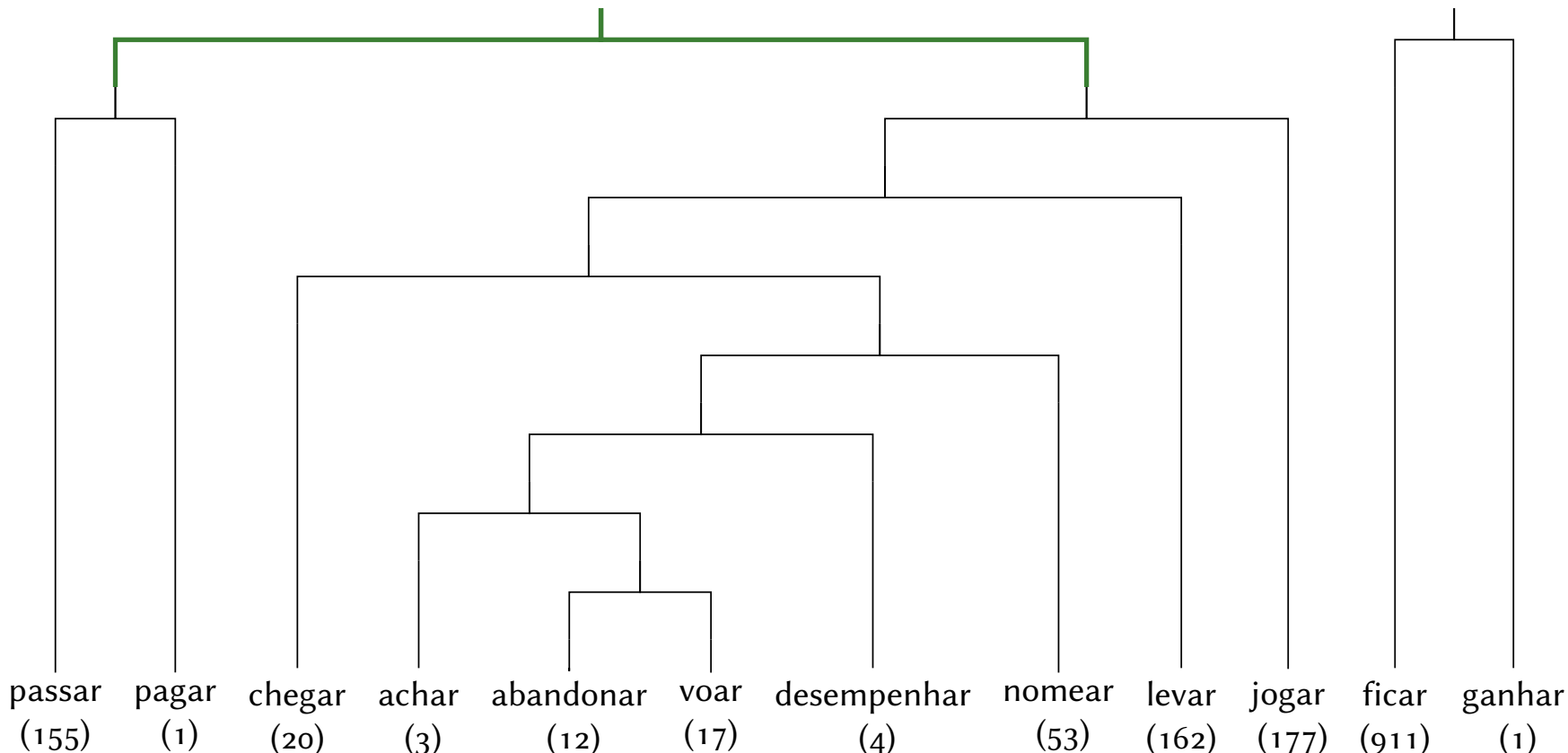
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



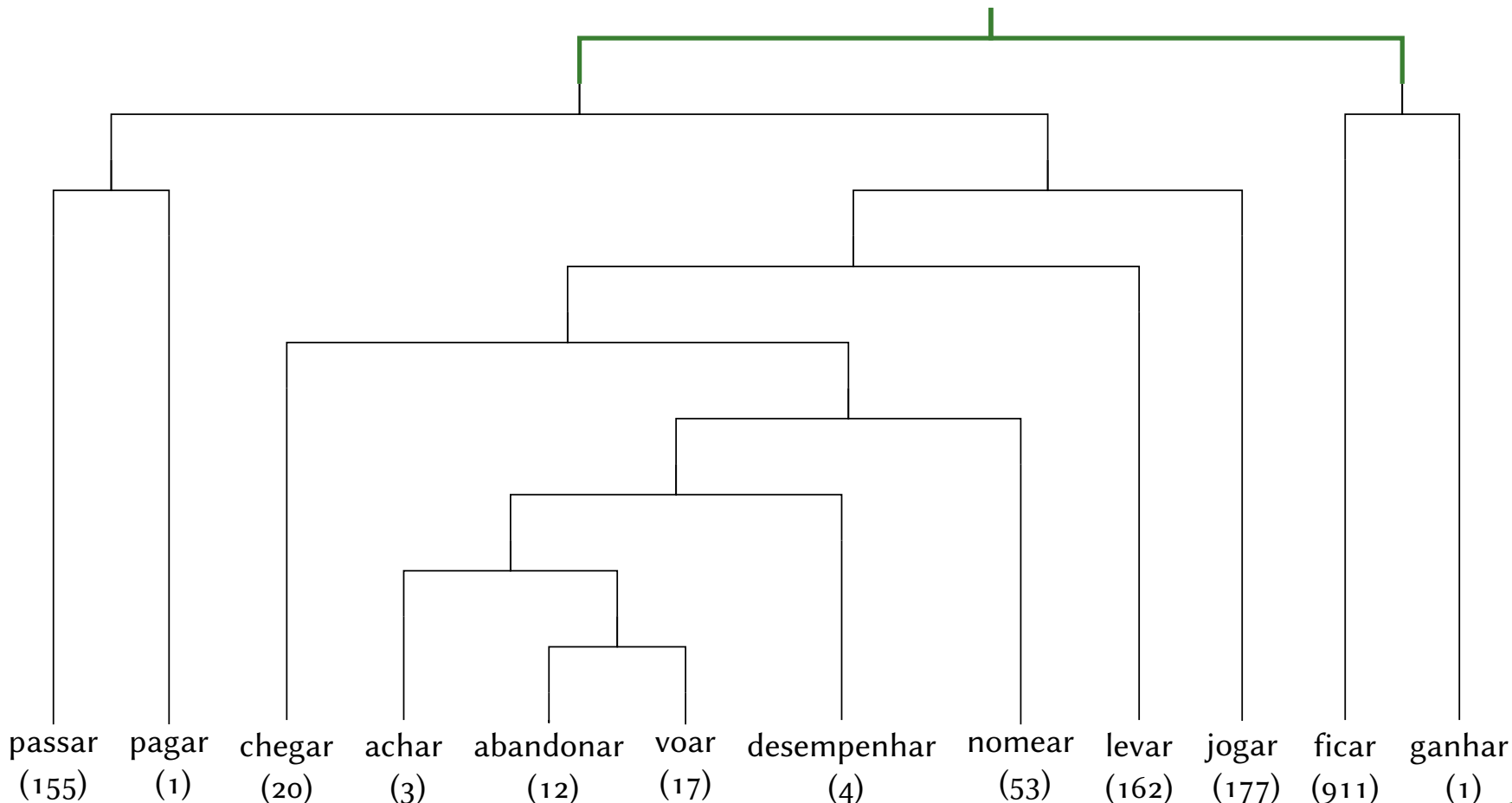
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



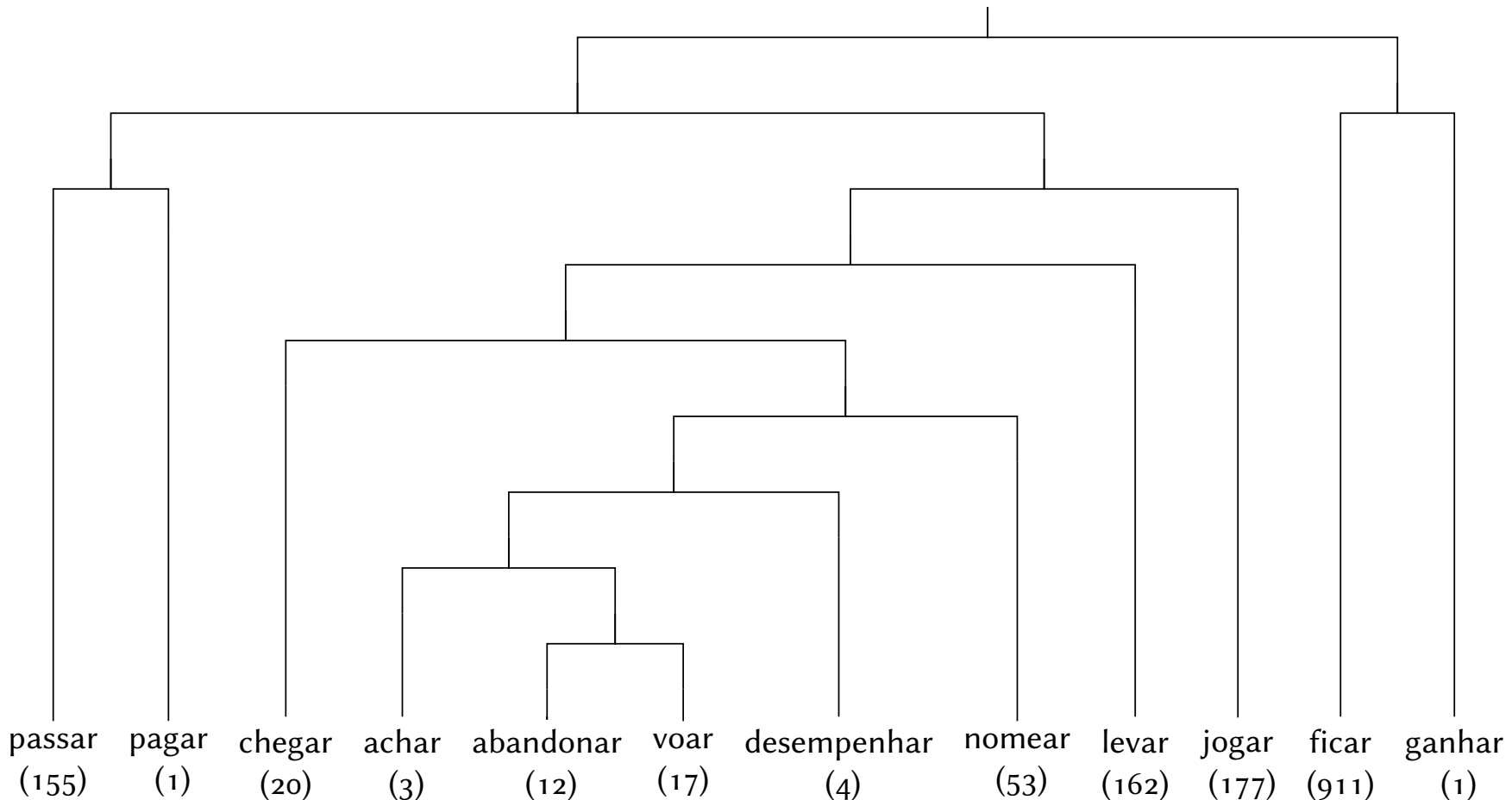
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



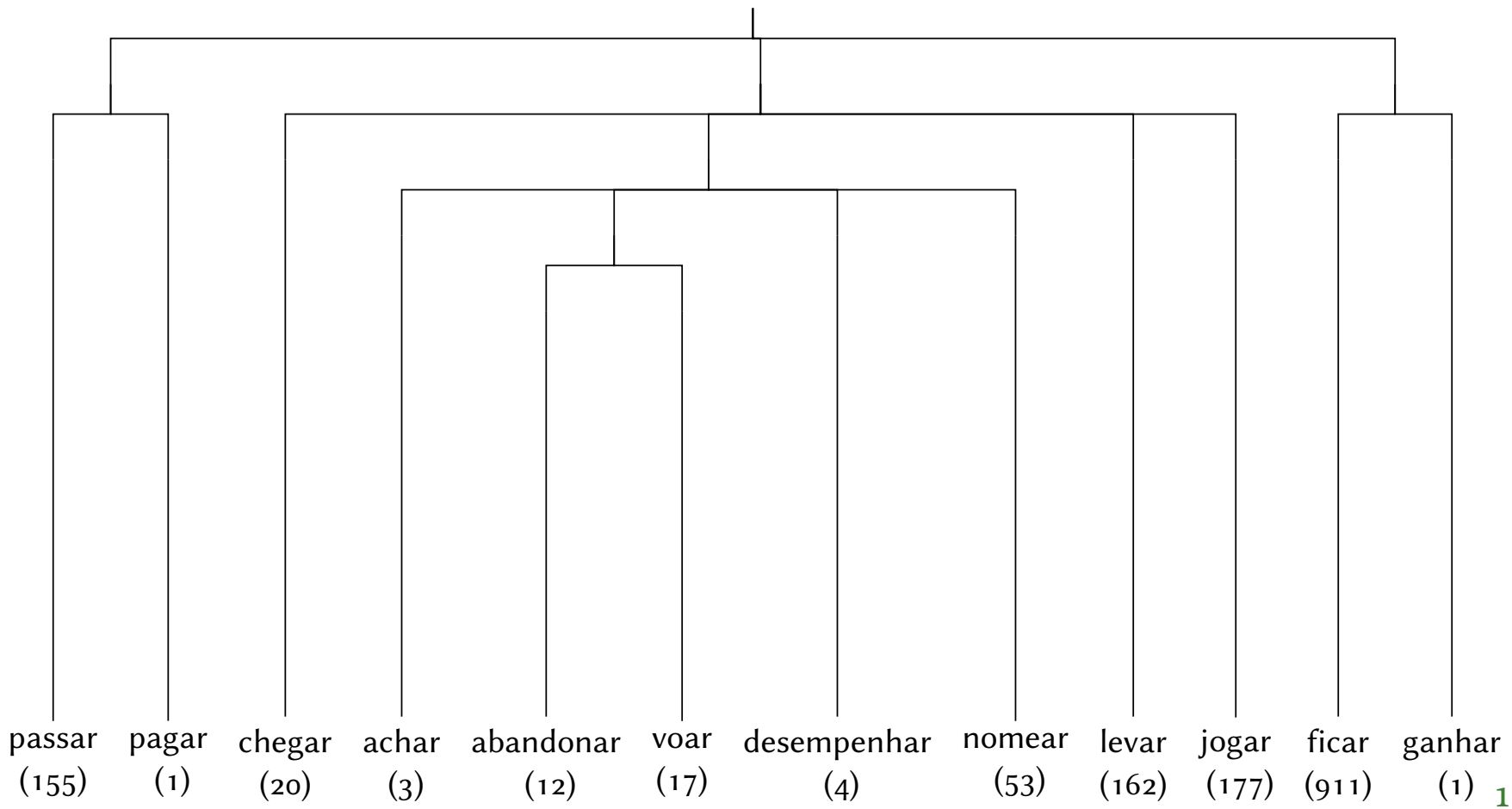
Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.



Clustering algorithm, ex. on European Portuguese conjugation.

(a) Begin with a partition into microclasses.
(b) Merge the pair optimising DL to get a new partition.
(c) Repeat until there is only 1 class.
(d) Run several times, merge variations.



- This allows for an intuitive formal definition of macroclasses
- Macroclasses: The partition that best optimises the description length.
  - As we merge clusters, we first expect the DL to decrease.
  - Macroclasses are reached when DL stops decreasing.
- It is an empirical issue whether a system has macroclasses or not.

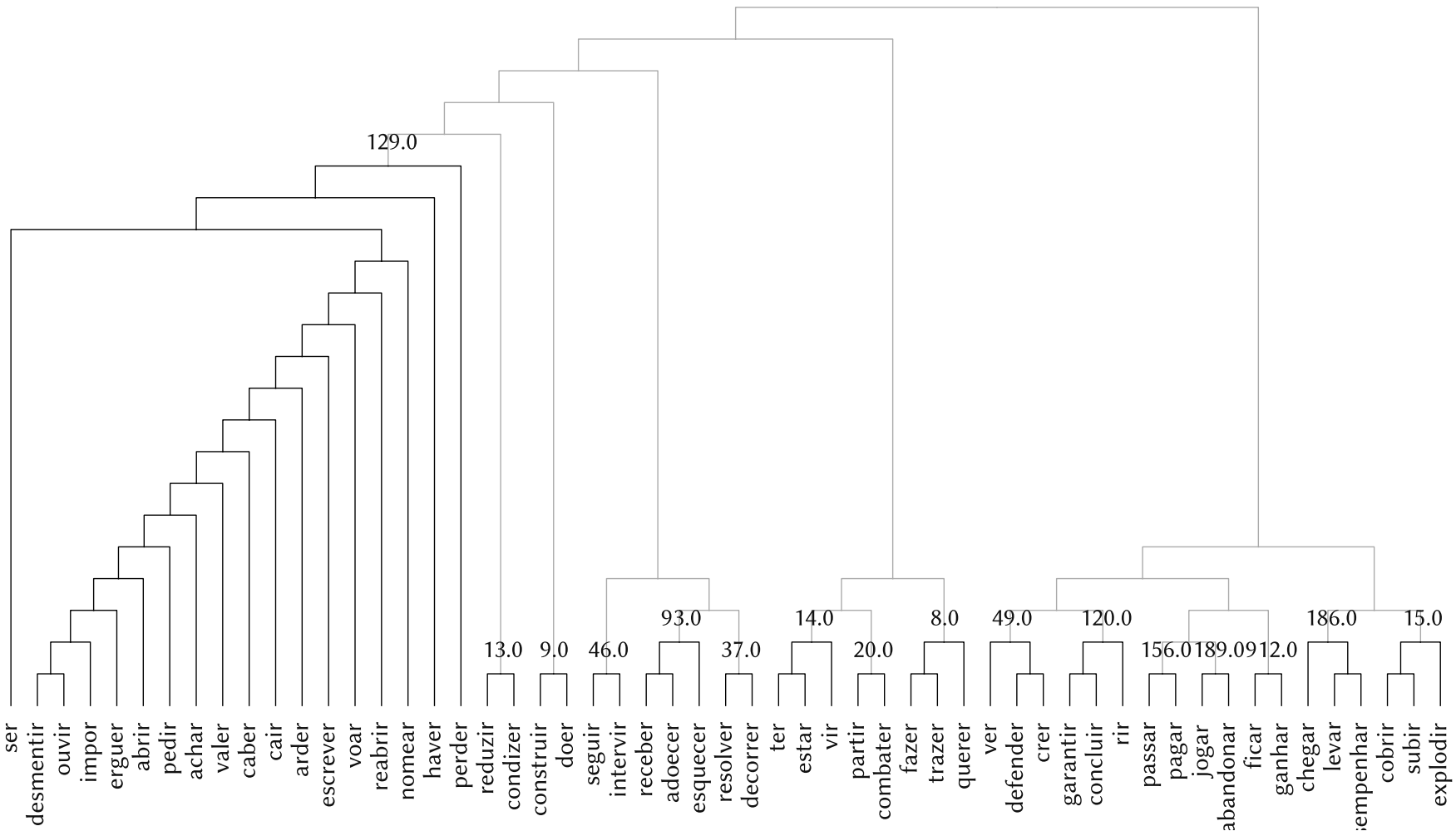
  We demonstrate their existence in French and European Portuguese conjugation systems.

Table of Contents

Datasets

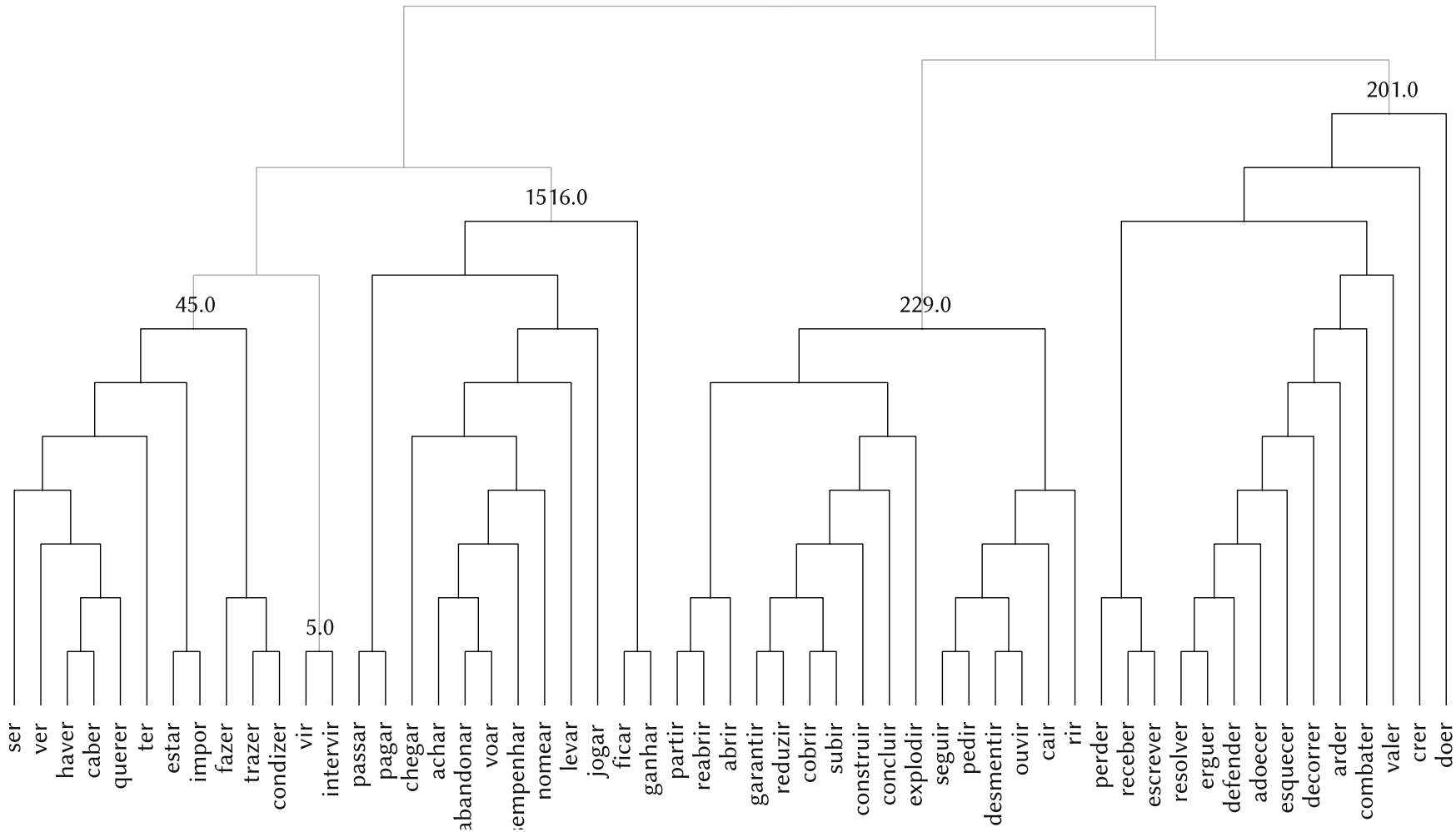
- Paradigm tables contain phonemically transcribed forms.
- **European Portuguese**: Coimbra pronunciation dictionary (Veiga, Candeias, and Perdigão, 2013) (1995 verbal entries).
- **French**: Flexique (Bonami, Caron, and Plancq, 2014) (5406 verbal entries).
- Comparing **local and global** segmentation strategies

- **Global strategy** (stem & exponents): Produces scattered classes with no relationship to conventional knowledge of Portuguese verbal IC.

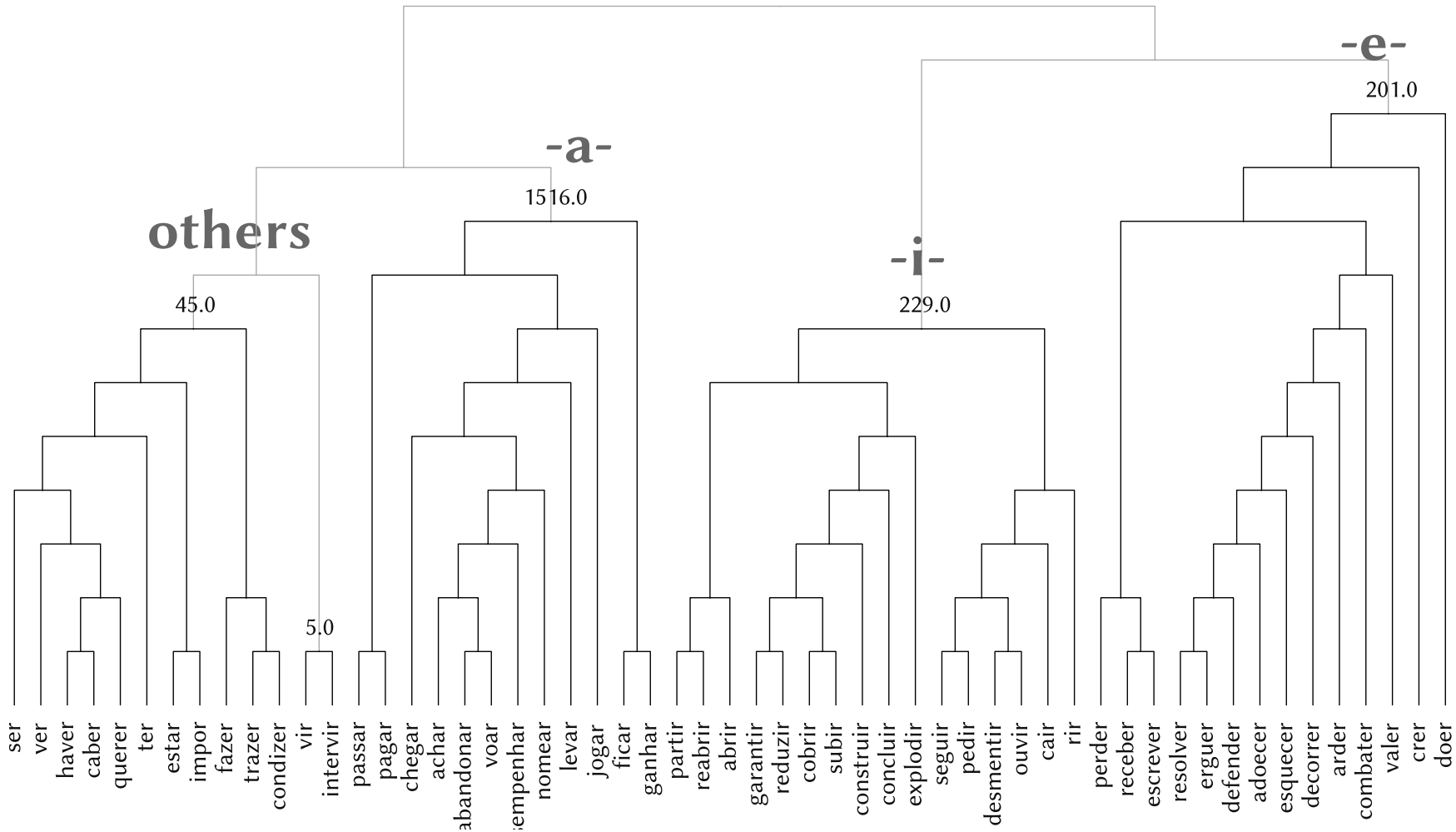


Portuguese classification, local patterns

- **Local strategy** (alternation patterns): finds generalisations that display interesting relationship with traditional accounts.

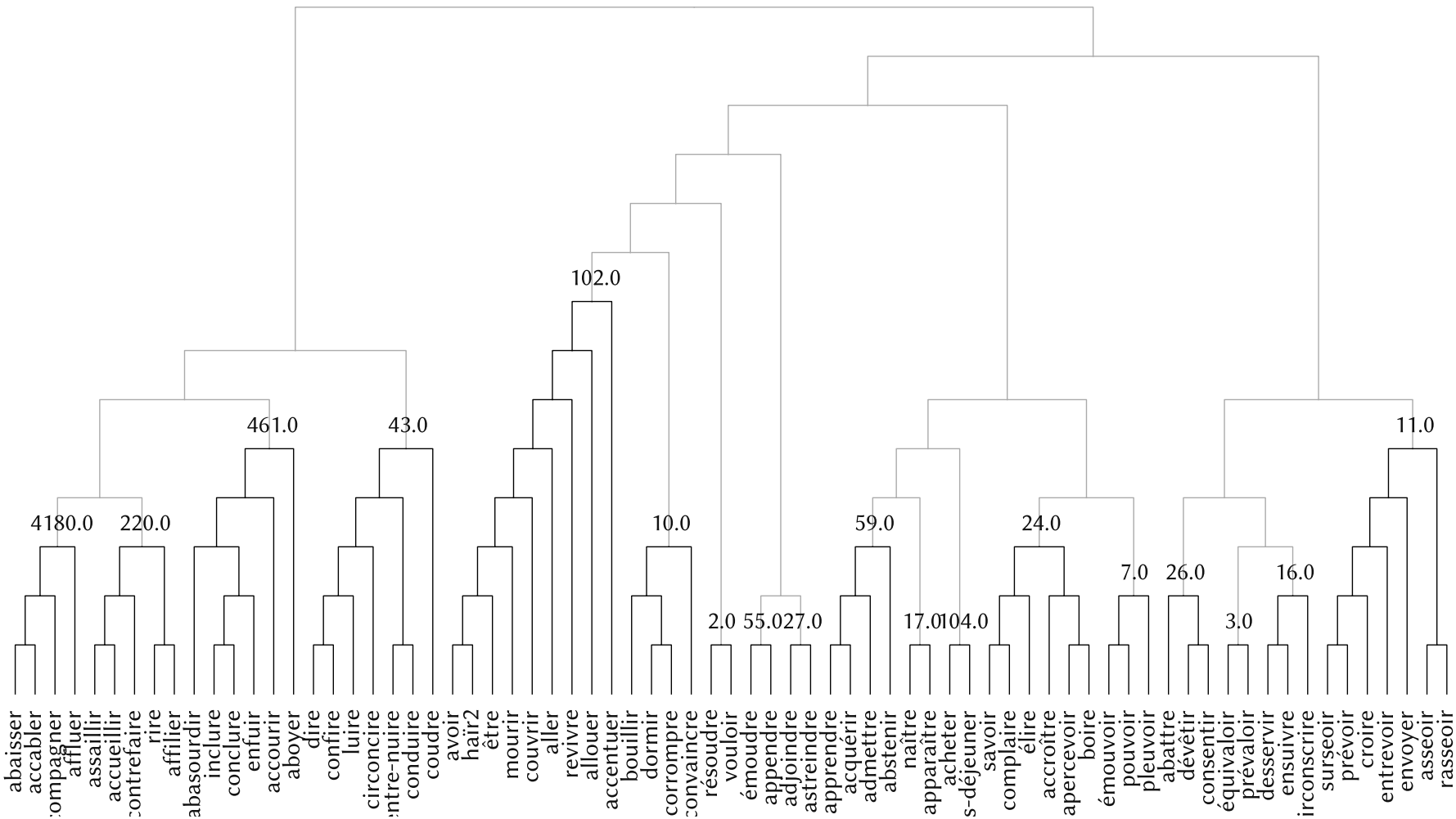


- **Local strategy** (alternation patterns): finds generalisations that display interesting relationship with traditional accounts.



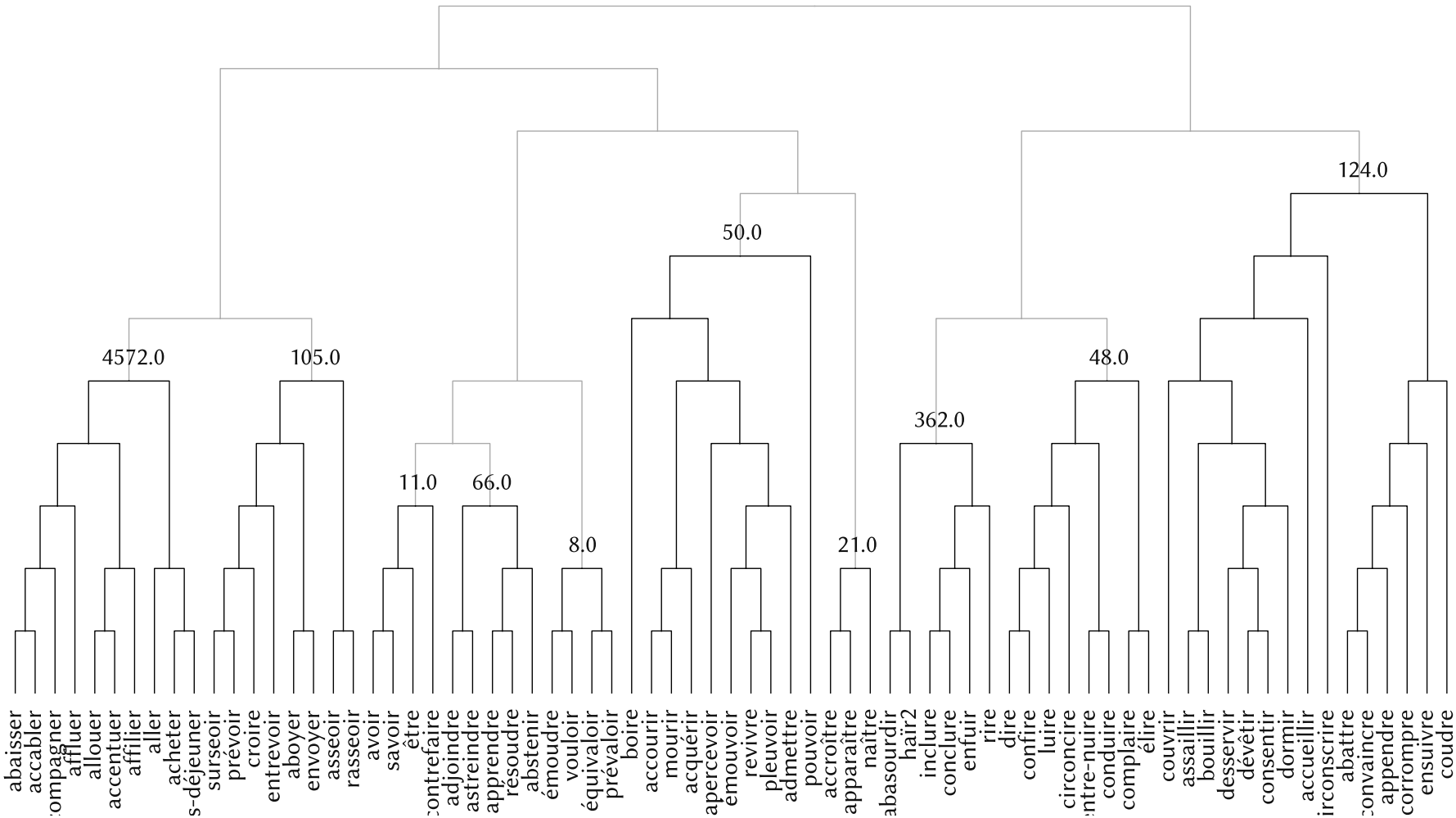
French classification, global patterns

- **Global strategy** (stem & exponents): Produces scattered classes with no relationship to conventional knowledge of French verbal IC.



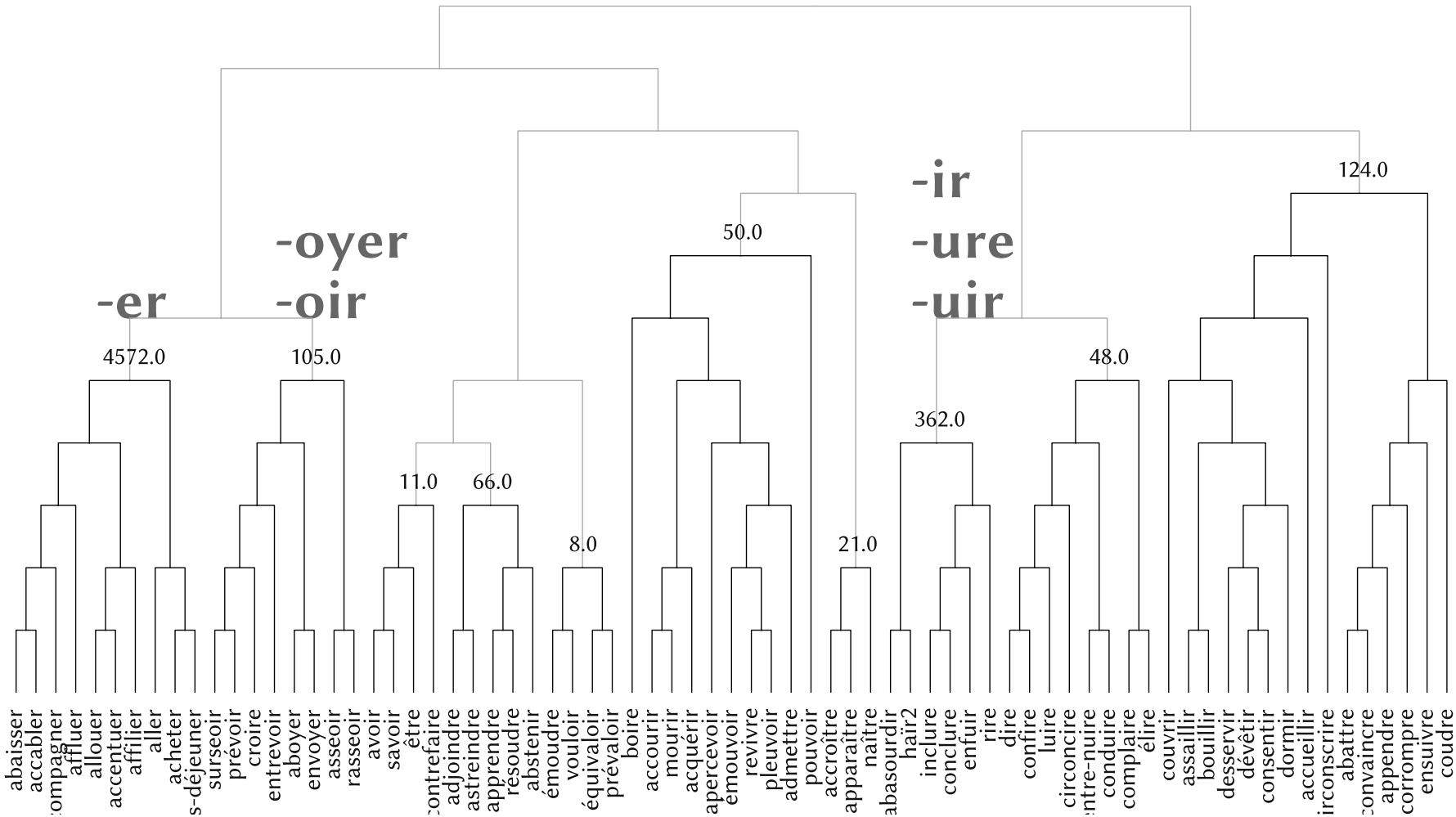
French classification, local patterns

- **Local strategy** (alternation patterns): finds generalisations that display interesting relationship with traditional accounts.



French classification, local patterns

- **Local strategy** (alternation patterns): finds generalisations that display interesting relationship with traditional accounts.



Discussion

- We do find macroclasses

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005
  - The algorithm had no knowledge of previous accounts.

Discussion

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005
  - The algorithm had no knowledge of previous accounts.
- We find groupings that were overlooked:

Discussion

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005
  - The algorithm had no knowledge of previous accounts.
- We find groupings that were overlooked:
  - French: -yer, -oir

Discussion

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005
  - The algorithm had no knowledge of previous accounts.
- We find groupings that were overlooked:
  - French: -yer, -oir
  - French: haïr, finir, -ure, uire

Discussion

- We do find macroclasses
  - Not a bipartition (regular / irregular or productive/unproductive), contra Kilani-Schoch and Dressler, 2005
  - The algorithm had no knowledge of previous accounts.
- We find groupings that were overlooked:
  - French: -yer, -oir
  - French: haïr, finir, -ure, uire
  - Portuguese: two “irregular” groups.

Comparison to other works

	Generalisations	Criterion	Algorithm
Brown and Evans, 2012	raw paradigms	Compression distance	CompLearn
Bonami, 2014	Affixes	Edit distance	UPGMA
Bonami, 2014	Patterns	Hamming distance	UPGMA
Lee and Goldsmith, 2013	Sets of characters	DL variant	greedy bottom-up
This work	Local patterns	DL	greedy bottom-up
This work	Global patterns	DL	greedy bottom-up

Features of our approach:

- Principled notion of Inflectional Realization.
- Using a measure that evaluates the quality of the system allows us to infer macroscopic generalisations.
- No parameters to adjust: **Occam's razor** is the only criterion.

Table of Contents

- **Main properties**:
  - Based on information-theoretic measures.
  - Relies on automatically inferred generalisations.
  - Aims at cross-linguistic applications.
  - Formal definition of macroclasses and microclasses.
- An analysis into macroclasses can be empirically motivated.
- **Local segmentation** better captures the structure in inflection systems than global segmentation.
  - Supports the relevance of local patterns of alternation in abstractive approaches (Blevins, 2006).
  - Complementary to work on information-theoretic modelling of implicative structure (Ackerman, Blevins, and Malouf, 2009; Ackerman and Malouf, 2013; Bonami and Beniamine, 2015)

Code available on my webpage:
`http://www.llf.cnrs.fr/fr/Gens/Beniamine`

Qumín

References I

Ackerman, Farrell, James P Blevins, and Robert Malouf (2009). "Parts and wholes: Patterns of relatedness in complex morphological systems and why they matter". In: Analogy in Grammar: Form and Acquisition, pp. 54–82.

Ackerman, Farrell and Robert Malouf (2013). "Morphological organization: The low conditional entropy conjecture." In: Language 89.3, pp. 429–464.

Blevins, James P. (2006). "Word-based morphology". In: Journal of Linguistics 42 (03), pp. 531–573. ISSN: 1469-7742. DOI: 10.1017/S0022226706004191.

Bonami, Olivier (2014). "La structure fine des paradigmes de flexion". French. Habilitation à diriger des recherches. U. Paris Diderot.

Bonami, Olivier and Sacha Beniamine (2015). "Implicative structure and joint predictiveness". In: ed. by Vito Pirelli, Claudia Marzi, and Marcello Ferro. URL: http://ceur-ws.org/Vol-1347/.

Bonami, Olivier, Gauthier Caron, and Clément Plancq (2014). "Construction d'un lexique flexionnel phonétisé libre du français". In: Actes du quatrième Congrès Mondial de Linguistique Française, pp. 2583–2596.

Brown, Dunstan and Roger Evans (2012). "Morphological complexity and unsupervised learning: validating Russian inflectional classes using high frequency data". In: Current Issues in Morphological Theory: (Ir)regularity, analogy and frequency. Ed. by F. Kiefer, M. Ladányi, and P. Siptár. Amsterdam: John Benjamins, pp. 135–162.

Corbett, Greville G. (2009). "Canonical Inflectional Classes". In: Selected Proceedings of the 6th Décembrettes: Morphology in Bordeaux.

Corbett, Greville G. and Norman M. Fraser (1993). "Network Morphology: a DATR account of Russian nominal inflection". In: Journal of Linguistics 29, pp. 113–142.

References II

Dressler, Wolfgang U. and Anna M. Thornton (1996). "Italian Nominal Inflection". In: *Wiener Linguistische Gazette* 55-57, pp. 1–26.

Kilani-Schoch, Marianne and Wolfgang Dressler (2005). *Morphologie naturelle et flexion du verbe français*. Tübingen: Gunter Narr Verlag.

Lee, Jackson and John A. Goldsmith (2013). "Automatic morphological alignment and clustering". Presented at the 2nd American International Morphology Meeting.

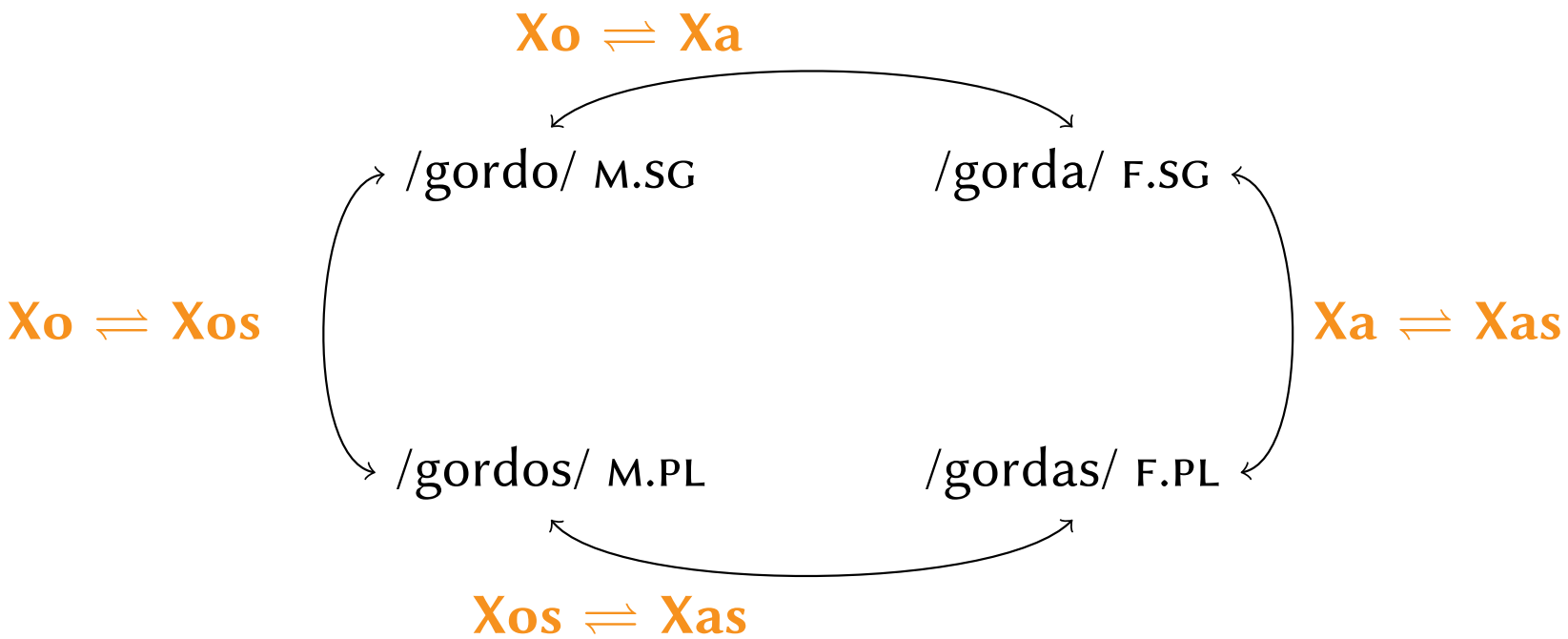
Rissanen, J. (1984). "Universal coding, information, prediction, and estimation". In: *IEEE Tr. on Info. Th.* 30.4, pp. 629–636.

Sagot, Benoît and Géraldine Walther (2011). "Non-canonical inflection: data, formalisation and complexity measures". In: *Systems and Frameworks in Computational Morphology*. Ed. by Cerstin Mahlow and Michael Piotrowski. Vol. 100. Communications in Computer and Information Science. Zurich, Suisse: Springer, pp. 23–45. ISBN: 978-3-642-23137-7.

Veiga, Arlindo, Sara Candeias, and Fernando Perdigão (2013). "Generating a pronunciation dictionary for European Portuguese using a joint-sequence model with embedded stress assignment". English. In: *Journal of the Brazilian Computer Society* 19.2, pp. 127–134. ISSN: 0104-6500. DOI: 10.1007/s13173-012-0088-0.

Walther, Géraldine (2013). "On canonicity in morphology:an empirical, formal and computational approach". PhD thesis. Université Paris Diderot, École doctorale de sciences du langage 132, U.F.R. de linguistique.

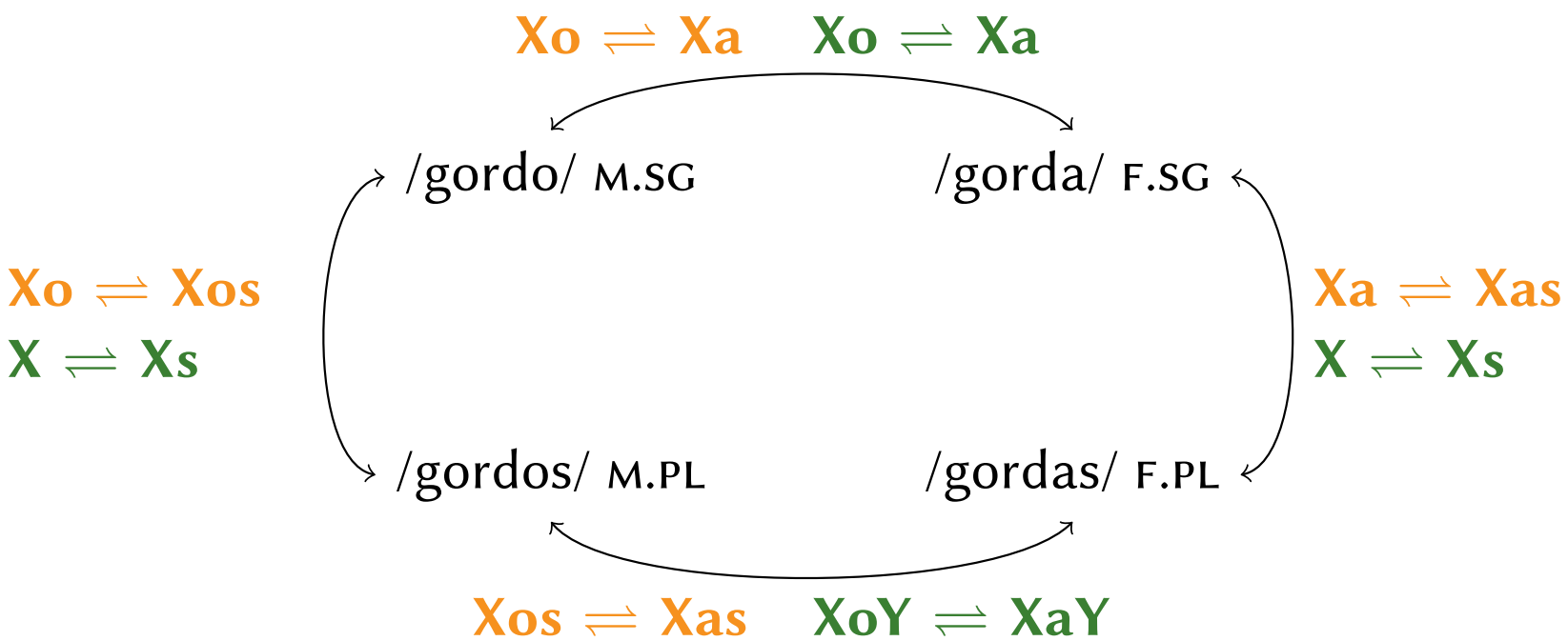
Both can be used in an abstractive approach:



global segmentation

Spanish adjective GORDO ‘fat’.

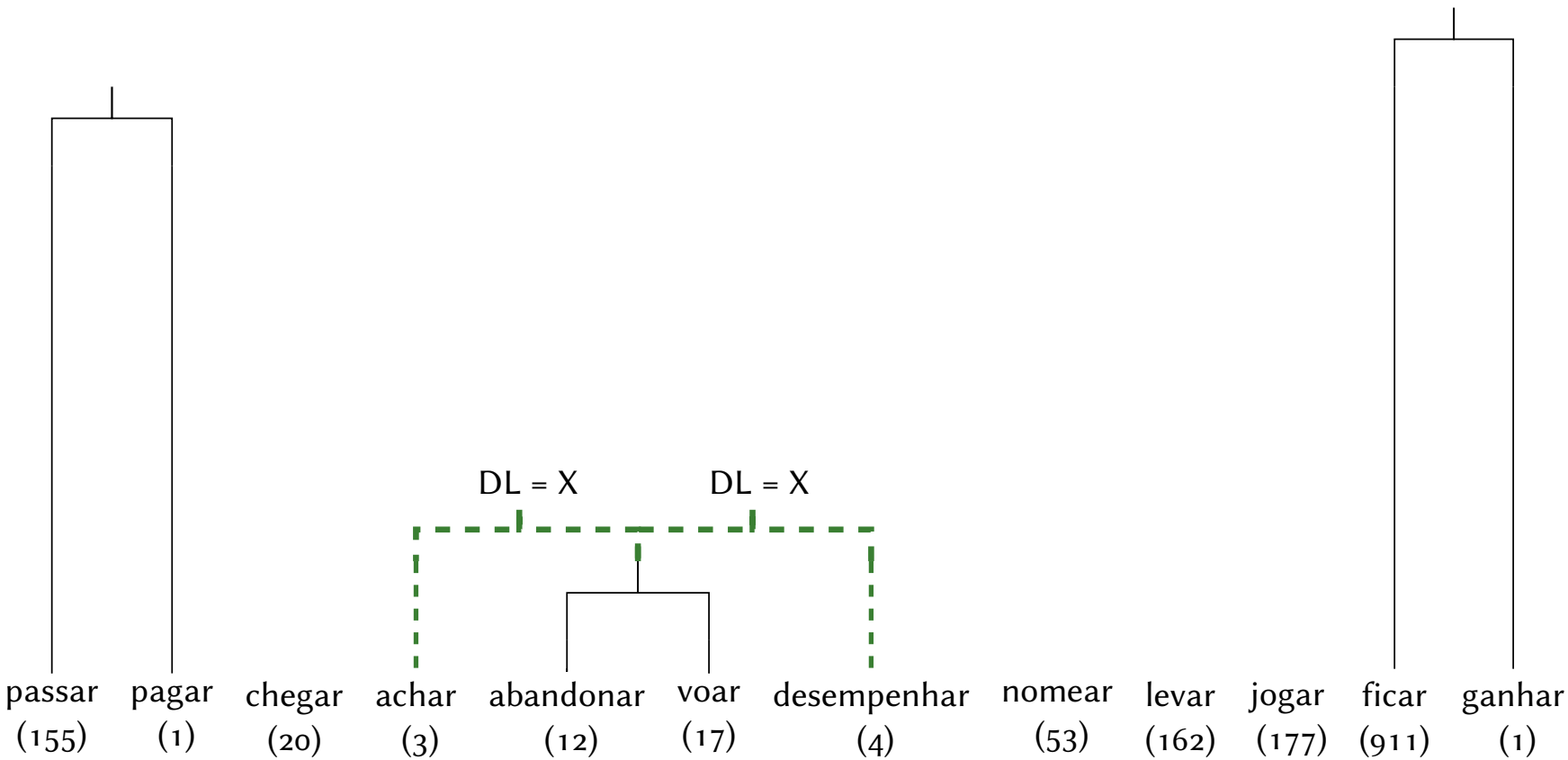
Both can be used in an abstractive approach:



global segmentation vs **local segmentation**

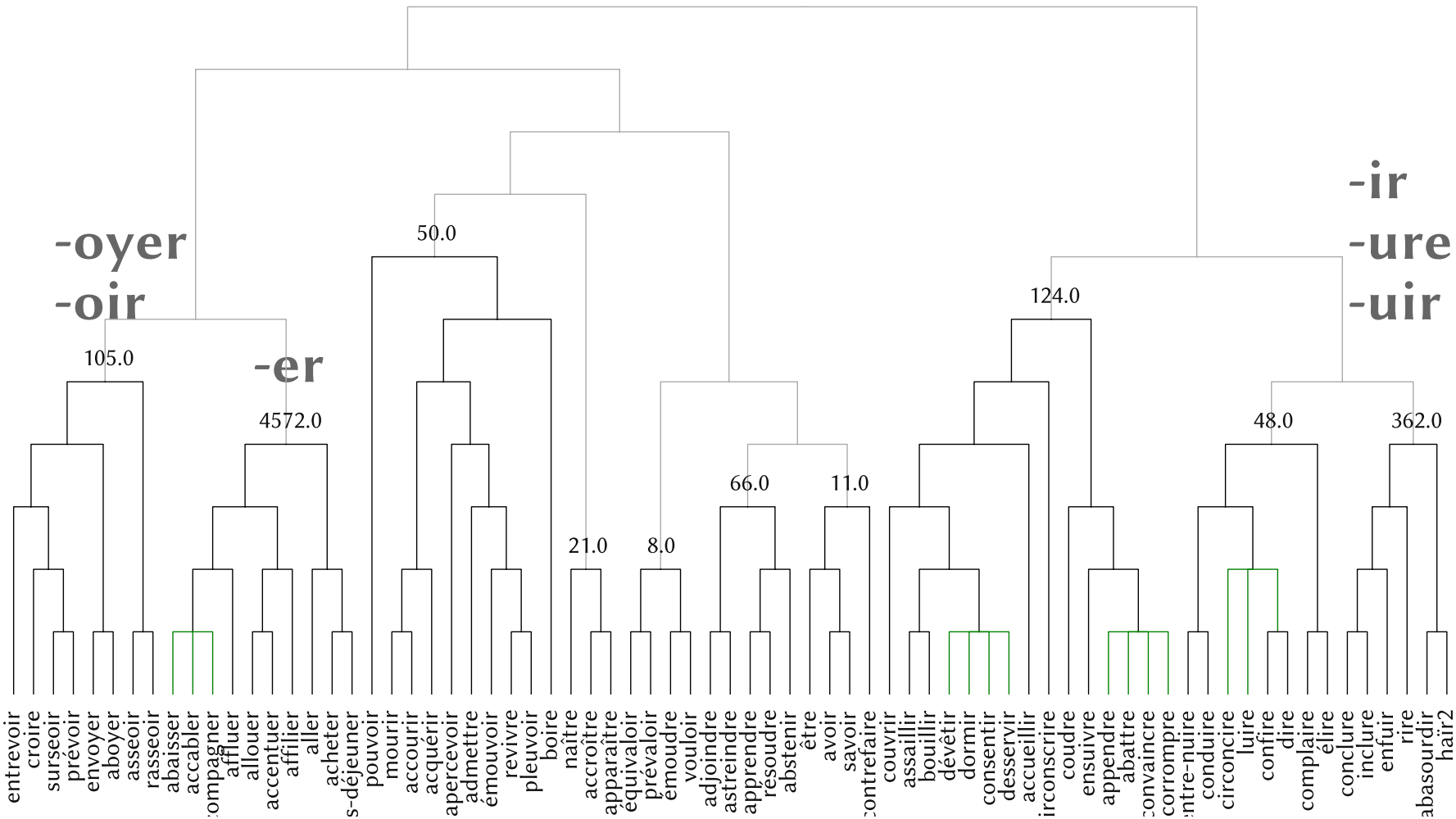
Spanish adjective GORDO ‘fat’.

Non Determinism



French classification, local patterns, merged results

- **Local strategy** (alternation patterns): finds generalisations that are in line with traditional accounts.



Portuguese classification, local patterns, merged results

- **Local strategy** (alternation patterns): finds generalisations that are in line with traditional accounts.