

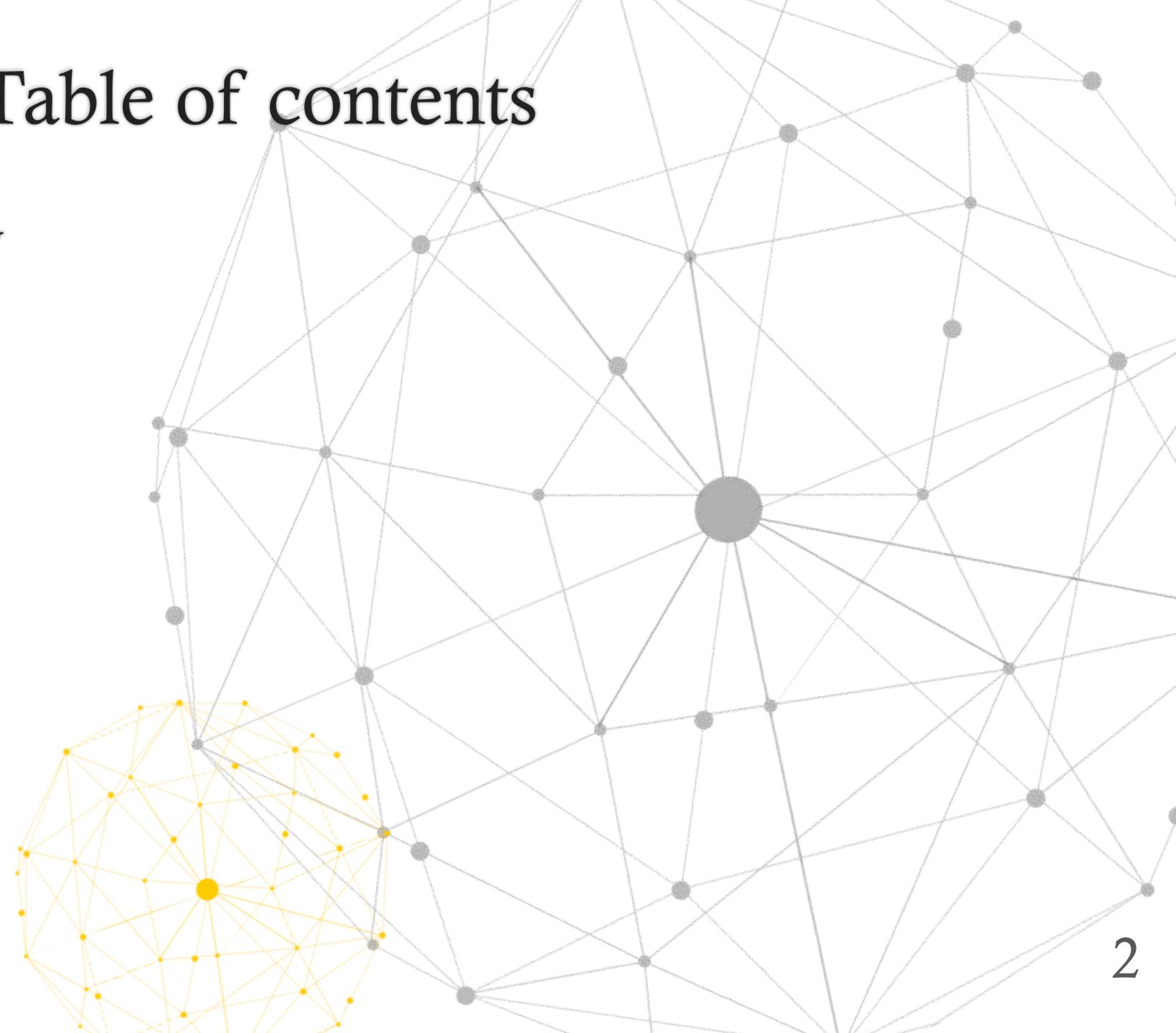
# Anthroponyms and paradigmatic derivation in French

Mathilde Huguin

*First Workshop on Paradigmatic Word Formation Modeling*

# Table of contents

1. Subject overview
2. Observations
3. Data
4. Interpretations
5. Conclusion



# Subject overview

- Our study focuses on morphologically complex words based on proper names of French politicians (henceforth PPN 'Politician Proper Name')
  1. PATRICK BALKANY<sub>NPr</sub> > BALKANYESQUE<sub>Adj</sub>

« Curieux comment le fisc s'est montré si peu curieux avec le patrimoine **Balkanyesque** pendant toutes ces années ! »

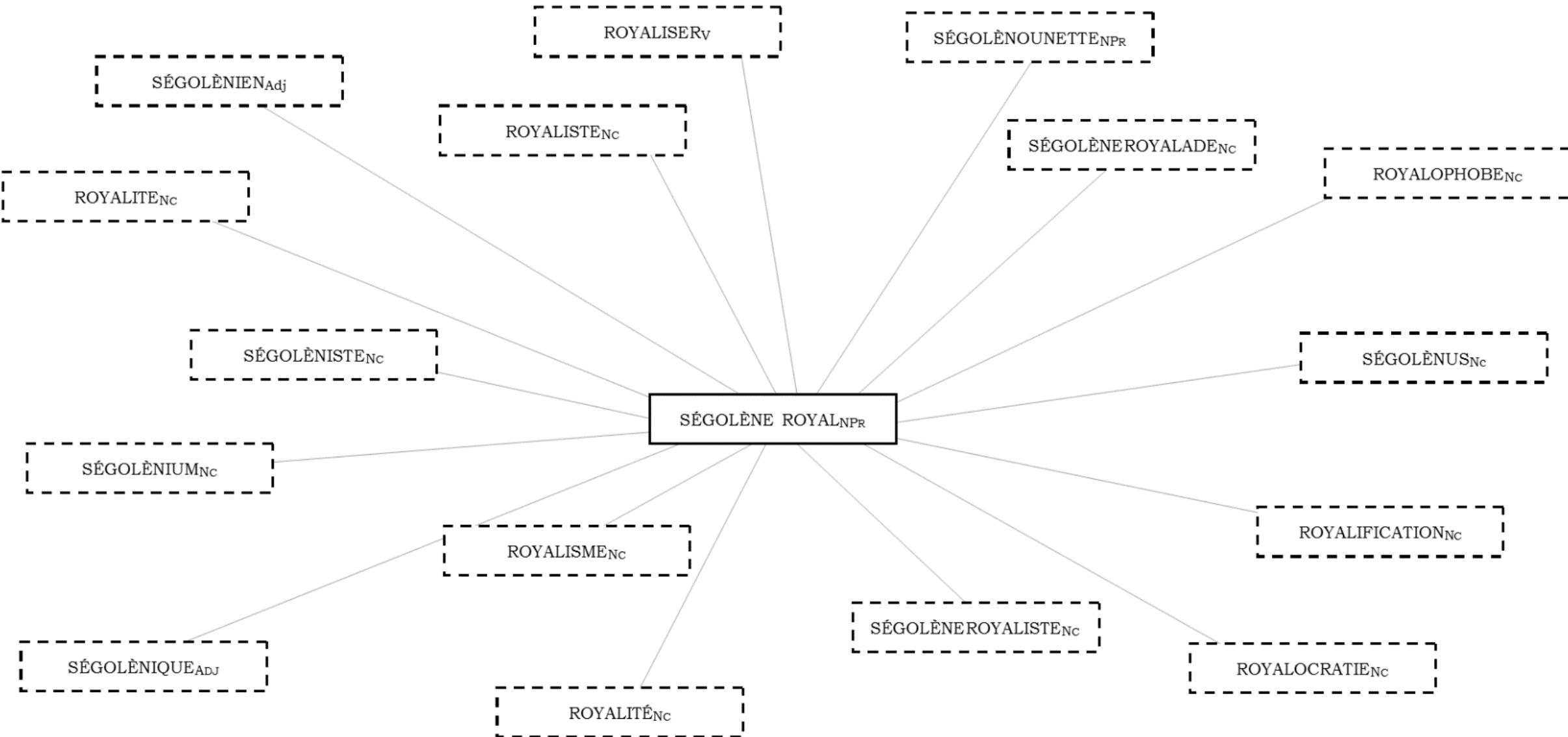
*'Curious how the fisc has shown so little curiosity with the (Balkany -esque) heritage during all these years!'*
  2. NADINE MORANO<sub>NPr</sub> > MORANOLOGUE<sub>Nc</sub>

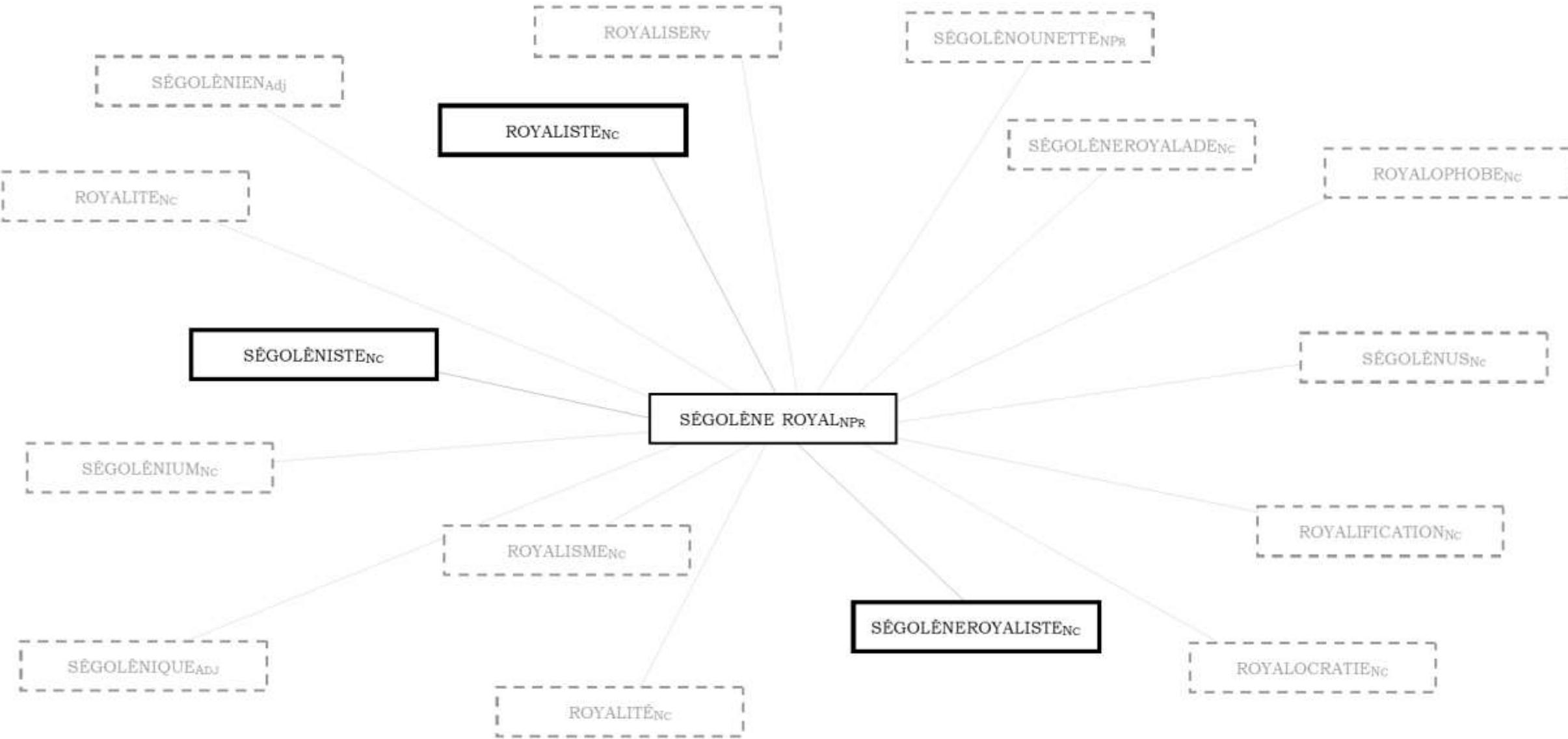
« Je ne comprends pas... Qu'est-ce que j'ai raté ? Je ne suis pas un **moranologue**. »

*'I don't understand... What did I miss? I'm not a (Morano -logist).'*

# Observations

SÉGOLÈNE ROYAL\_NPR







« " Il n'y a pas deux Partis socialistes ", a assuré le **ségoléniste** Manuel Valls. »

*"There are not two socialist parties": assured the (Ségolène –ist) Manuel Valls.'*

→ 'follower of Ségolène Royal'



« Sans vouloir "ségoléniser", je présente mes excuses à mon lectorat très nombreux en cette période pré-électorale... »

*'Without wanting "(to Ségolène –ize)", I apologize to my many readers in this pre-election period... '*

→ '~ to apologize for something that you did not do'



## **Ségolène Royal demande "pardon" pour le "discours de Dakar" de Nicolas Sarkozy**

*"Quelqu'un est venu ici vous dire que 'l'Homme africain n'est pas entré dans l'histoire'. Pardon, pardon pour ces paroles humiliantes et qui n'auraient jamais dû être prononcées et – je vous le dis en confidence – qui n'engagent ni la France, ni les Français", a déclaré, lundi 6 avril, Ségolène Royal devant plusieurs centaines de personnes réunies au siège du Parti socialiste sénégalais, à Dakar.*

**Le Monde**

## ***Ségolène Royal apologizes for Nicolas Sarkozy's speech in Dakar***

*"Someone came here to say that 'Africans have not fully entered into history'. Sorry, sorry for these humiliating words that should never have been uttered and – I tell you in confidence – that do not commit France or French people", declared, Ségolène Royal, on Monday 6<sup>th</sup> April, before several hundreds of people at the Socialist Party headquarters, in Dakar.*

« Sans vouloir "ségoléniser", je présente mes excuses à mon lectorat très nombreux en cette période pré-électorale ... »

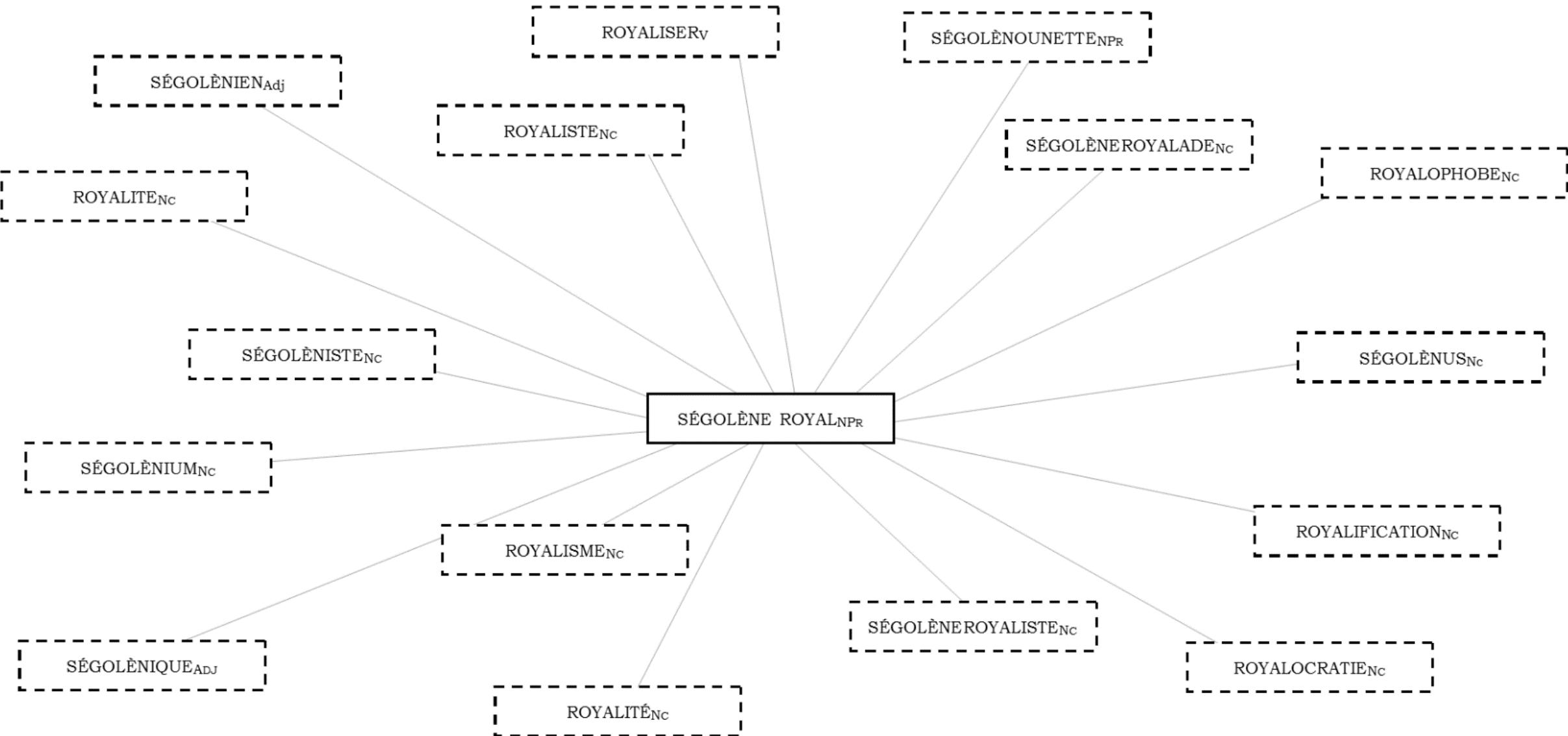
*'Without wanting "(to Ségalène -ize) ", I apologize to my many readers in this pre-election period...'*

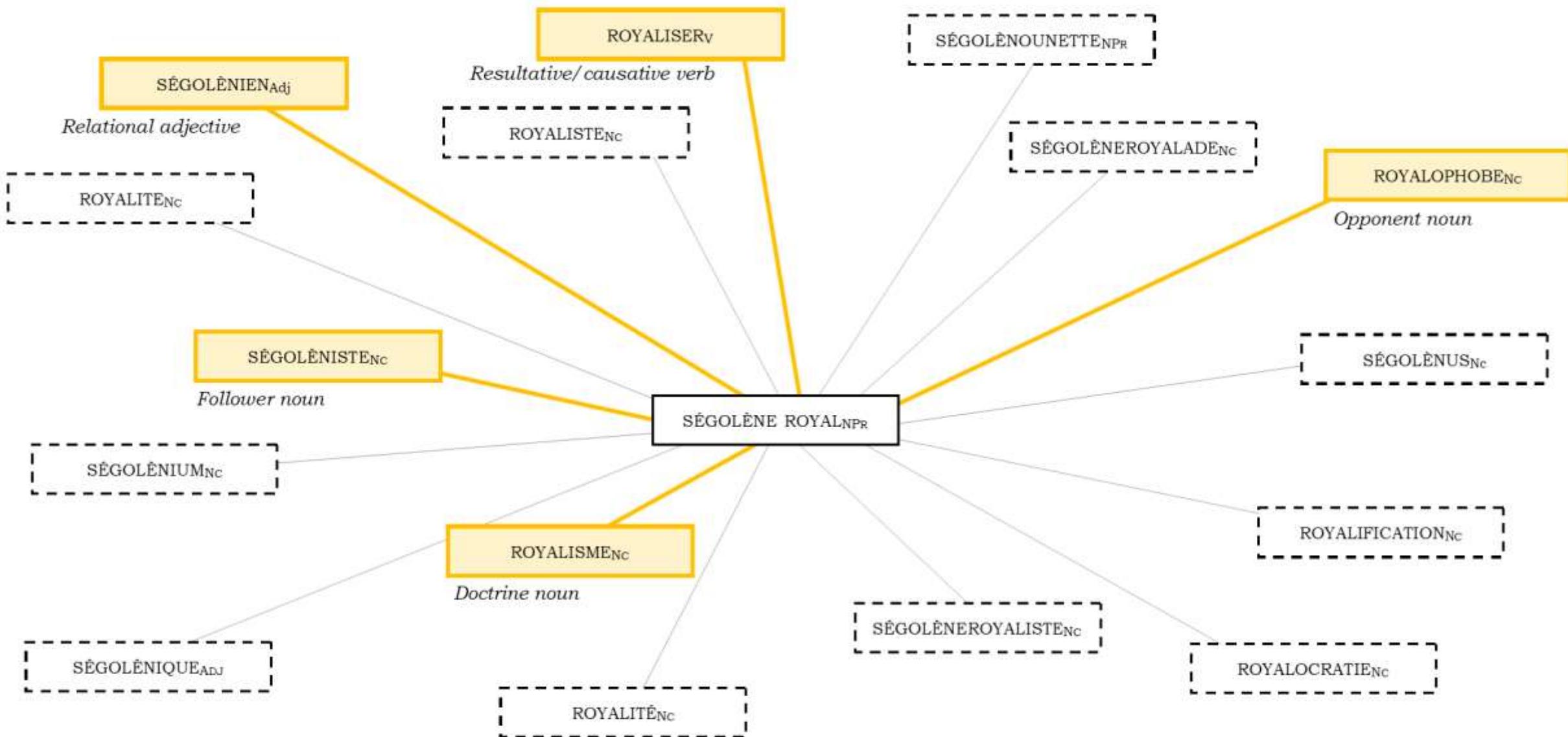
→ ‘~ to apologize for something that you did not do’

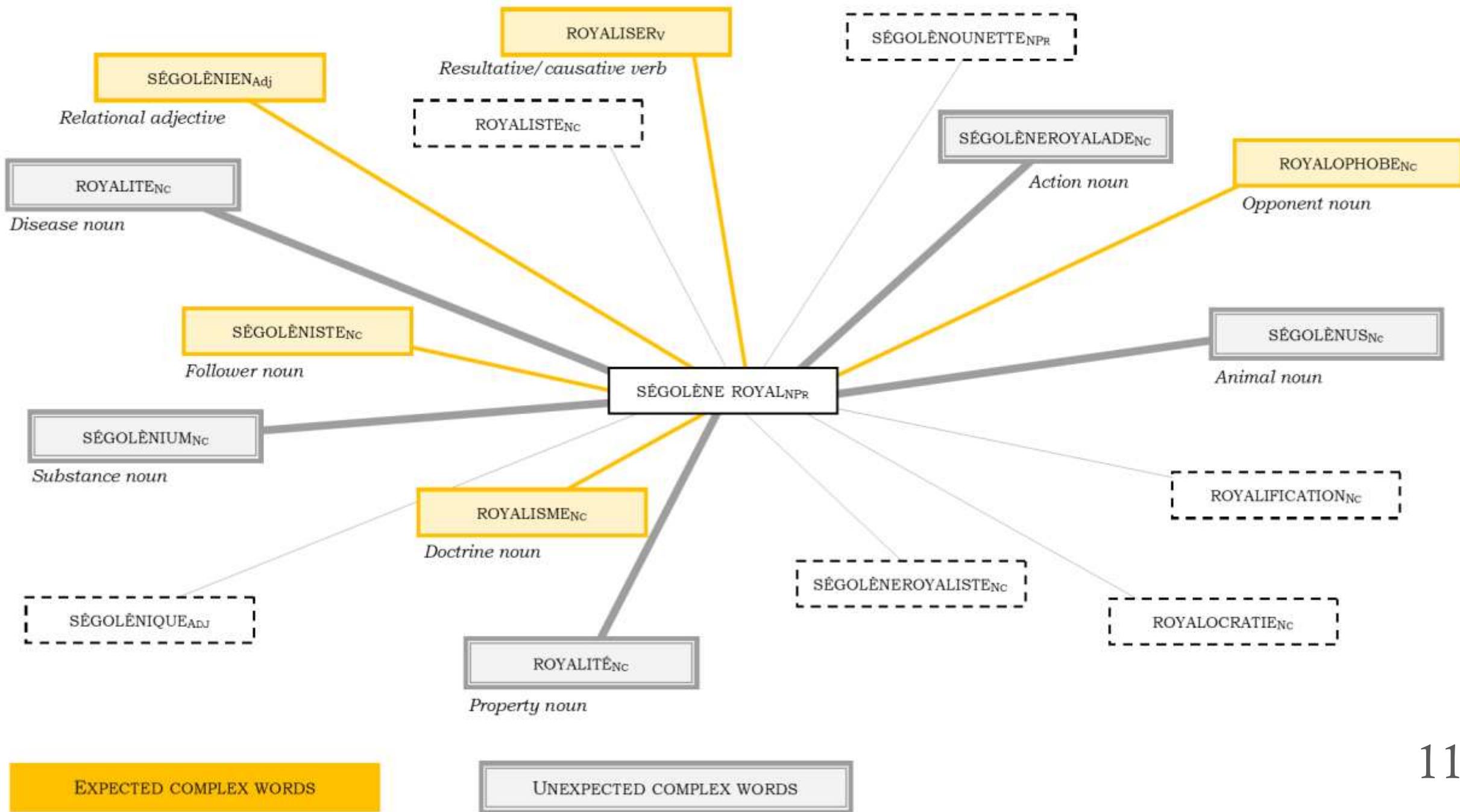
## **Les pardons de Ségolène Royal ou le comique de répétition**

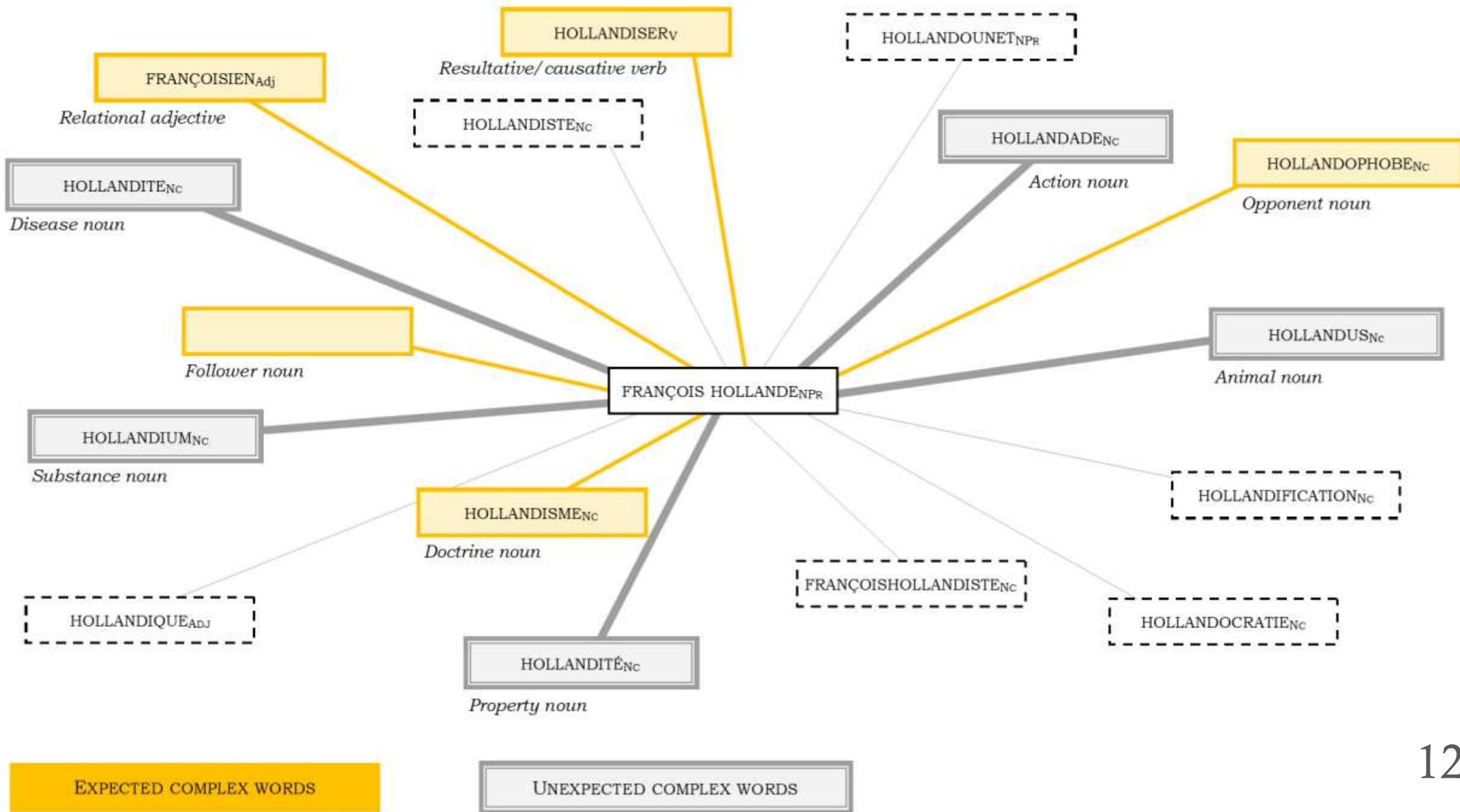
*Le nouvel  
Observateur*

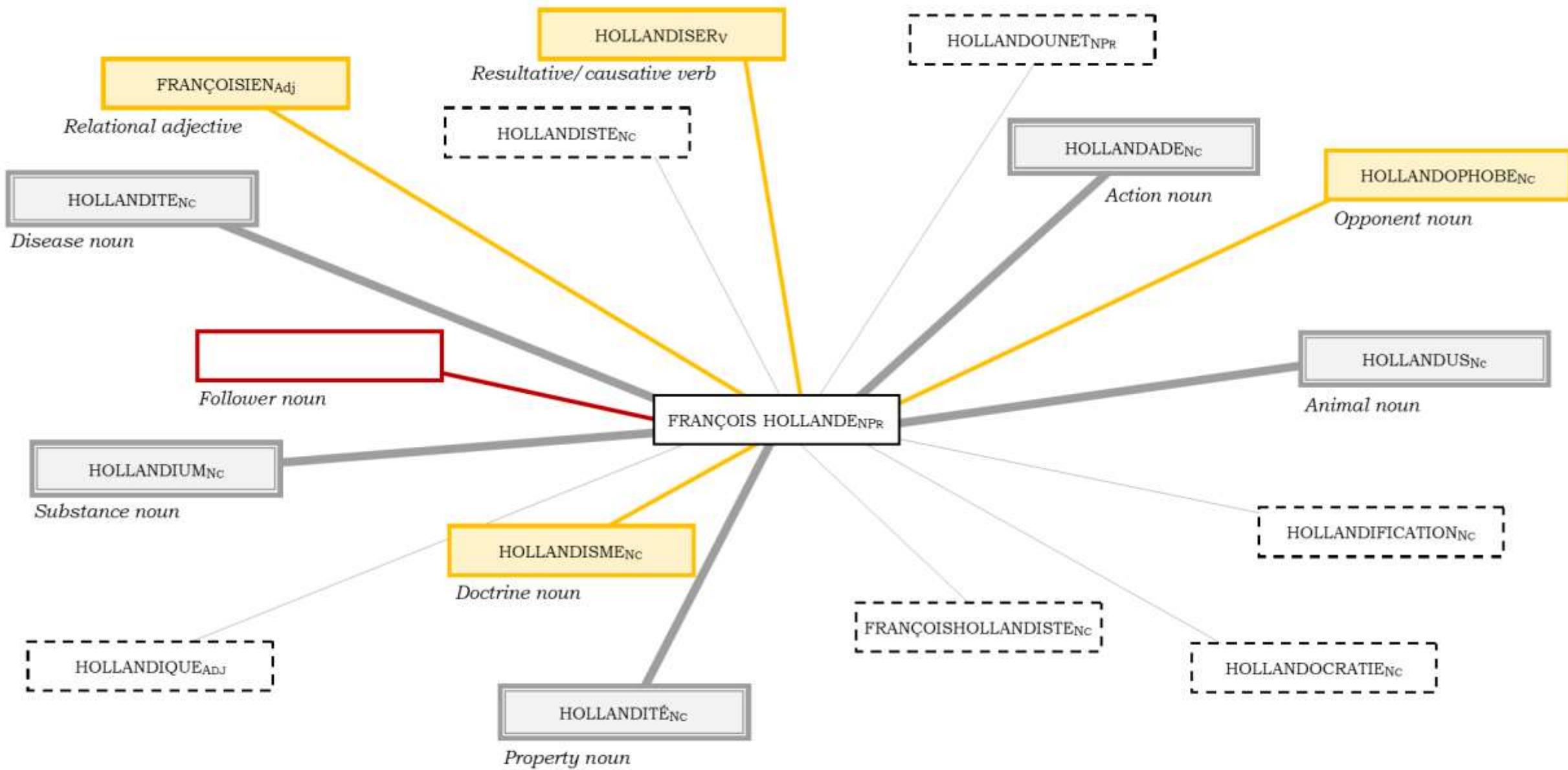
## ***Ségolène Royal's requests for forgiveness or the running gag***

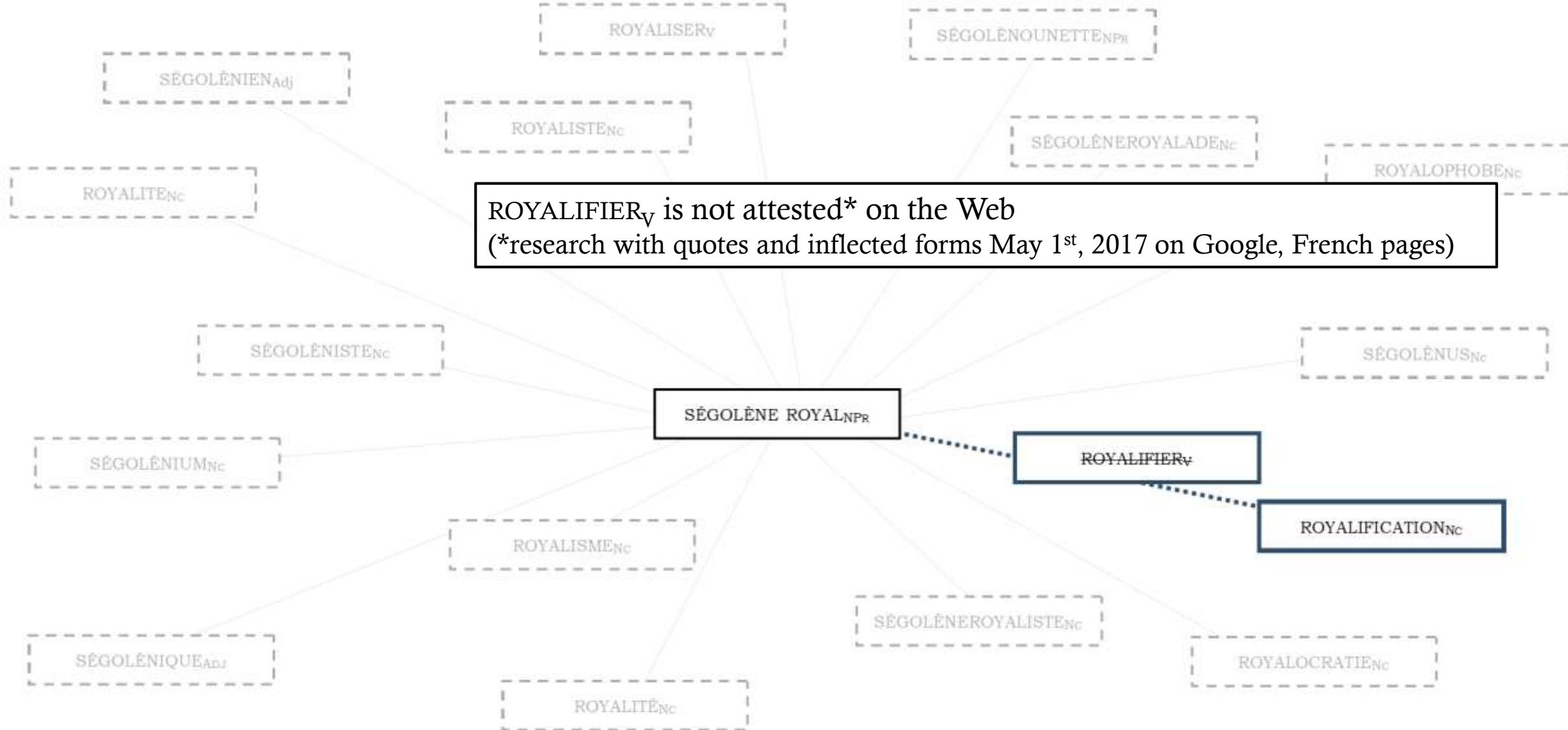








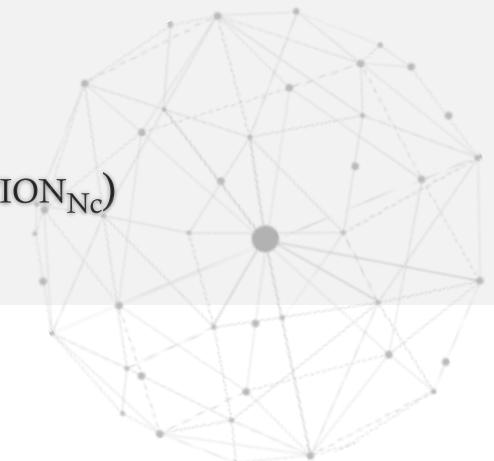




# Observations

## Summary of observations

- i. PPNs appear in different forms in the complex words ( $\text{SÉGOLÈNISTE}_{\text{Nc}}$  /segolɛn/,  $\text{ROYALISTE}_{\text{Nc}}$  /r w a jal/,  $\text{SÉGOLÈNEROYALISTE}_{\text{Nc}}$  /segolɛnr w a jal/)
- ii. Complex words can be interpreted in two ways
- iii. PPNs are the bases of expected (follower:  $\text{ROYALISTE}_{\text{Nc}}$ ) and unexpected complex words (substance:  $\text{SÉGOLÈNIUM}_{\text{Nc}}$ )
- iv. Derivational families have structural similarities
- v. Some words do not have any attested bases ( $\text{ROYALIFIER}_V > \text{ROYALIFICATION}_{\text{Nc}}$ )



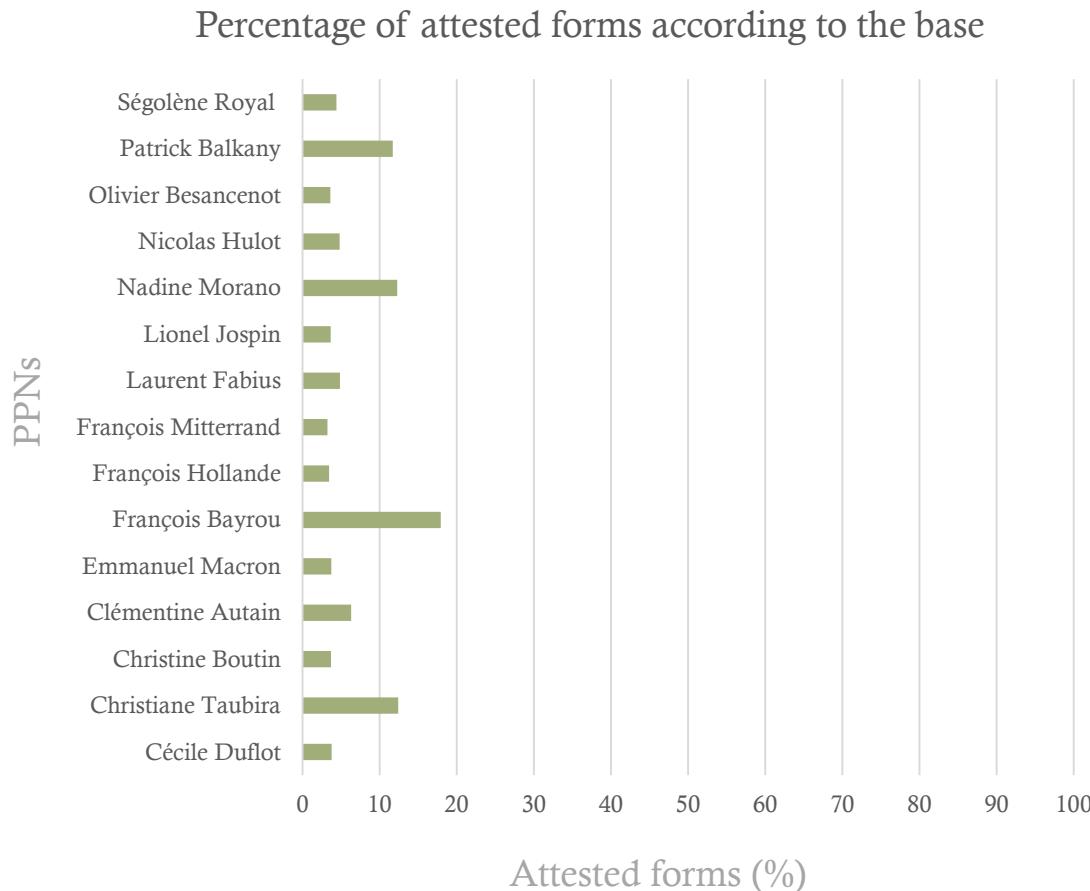
# Data



- Generation of ~130,000 candidate forms from:
  - A set of 90 PPNs
  - A set of morpho-phonological and graphic constraints (based on *Optimality Theory*: Price & Smolensky, 1993; Plénat, 1997; Lignon & Plénat, 2009; Roché & Plénat, 2014)
  - A set of 100 exponents, i.e. suffixes and ‘suffixoid’ elements (e.g. *-issime*, *-phobe*; cf. Lasserre & Montermini, 2014; Huguin, 2015)
- For a given PPN and one exponent there are several candidate forms

CHRISTIANE TAUBIRA<sub>NPr</sub> > *taubiratiste, taubirassiste, taubiraiste, taubiriste, christianetaubiraiste etc.*
- Candidate forms were collected from the Web (collaboration with Data-Observer, using Bing Search API)

# Data | sub-corpus



- 15 PPNs and their derivational families

CÉCILE DUFLOT, CHRISTIANE TAUBIRA, CHRISTINE BOUTIN, CLÉMENTINE AUTAIN, EMMANUEL MACRON, FRANÇOIS BAYROU, FRANÇOIS HOLLANDE, FRANÇOIS MITTERRAND, LAURENT FABIUS, LIONEL JOSPIN, NADINE MORANO, NICOLAS HULOT, OLIVIER BESANCENOT, PATRICK BALKANY, SÉGOLÈNE ROYAL

- 15,963 candidate forms (i.e. web requests) and 1,272 attested forms

# Interpretations | stem collection

- i. PPNs appear in different forms in the complex words ( $\text{SÉGOLÈNISTE}_{\text{Nc}}$  /segolɛn/,  $\text{ROYALISTE}_{\text{Nc}}$  /r wajal/,  $\text{SÉGOLÈNEROYALISTE}_{\text{Nc}}$  /segolɛnr wajal/)

« *L'information phonologique associée à un lexème prend donc la forme d'une série de cases contenant chacune un radical. Nous appelons espace thématique l'ensemble de ces cases.* »  
(Bonami, Boyé & Kerleroux, 2009: 107)

*'The phonological information associated with a lexeme thus takes the form of a series of cells each containing a stem. This set of cells is called stem space.'*

→ **PPNs have a stem network that differs from the *stem space* definition**

# Interpretations | stem collection

## Differences between PPN stems and *stem space* stems

- They do not have the same dependency links (cf. Bonami & Boyé, 2003)
- They are not selected by a rule of word-formation (cf. Bonami, Boyé & Kerleroux, 2009; Roché & Plénat, 2014)
- They can be deduced from the anthroponym graphic form
- They are all syntactically autonomous

# Interpretations | stem collection

---

EMMANUEL MACRON<sub>NPr</sub>

Last Name	First Name	First Name Last Name
<i>Macron</i>	<i>Emmanuel</i>	<i>Emmanuel Macron</i>
/makrɔ̃/	/emanyɛl/	/emanyɛlmakrɔ̃/

---

# Interpretations | stem collection

**NAJAT VALLAUD-BELKACEM<sub>NPr</sub>**

Last Name	First Name	First Name Last Name	Last Name <sub>1</sub>	Last Name <sub>2</sub>	Acronym
<i>Vallaud-Belkacem</i>	<i>Najat</i>	<i>Najat Vallaud-Belkacem</i>	<i>Vallaud</i>	<i>Belkacem</i>	<i>NVB</i>
/valo bεlkasɛm/	/naʒat/	/naʒatvalo bεlkasɛm/	/valo/	/bεlkasɛm/	/ɛnvebe/

1. Difficile d'épingler correctement le « taubiratus » et le « **vallaubelkacemus** ».
2. On peut aussi sortir du système **najatesque** en faisant l'école à la maison.
3. Merci pour cette réplique d'un ton parfaitement **najat-vallaud-belkacemien**.
4. Il est clair qu'en Taubiristan et en **Vallaudistan**, ça paraît un abus pur et simple.
5. Le rayonnement de la France pour les prochaines générations en cours de **belkacemisation** c'est au pire la trace de l'activité nucléaire, au mieux une allusion aux années les plus sombres de la colonisation et du commerce triangulaire.
6. Ne serait-ce que parce qu'ils énervent la bien-pensance « progressiste » **NVBiste**.

# Interpretations | semantic network

ii. Complex words can be interpreted in two ways

- **A denominative meaning**

1. FRANÇOIS BAYROU<sub>NPr</sub> > BAYROUISTE<sub>Nc</sub> ‘follower of/François Bayrou/’

« Voilà, aucun **bayrouiste** ne répond à ma question : pourquoi ce soutien ? »

‘So, no (Bayrou –ist) answers my question: why this support?’

- **A set of stereotipic features**

2. NADINE MORANO<sub>NPr</sub> > MORANISTE<sub>Adj</sub> ‘property to be conversative’

« Quand on défend la « pureté » de la langue française contre les importations de mots étrangers, mais aussi contre les tournures de phrases à la mode considérées agrammaticales ou incorrectes (quand on critique l’usage intensif de « trop » dans « elle est trop belle »), on est, en un sens, **moraniste**. »

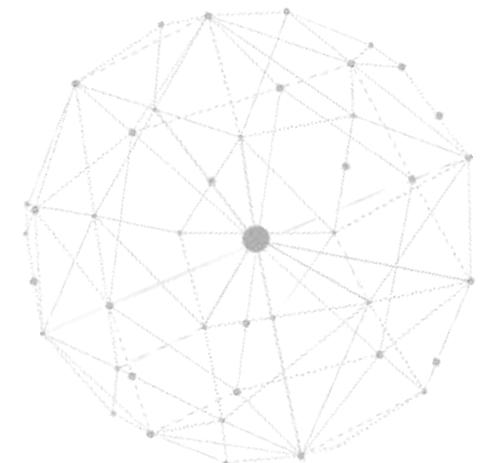
‘When we defend the "purity" of the French language against the foreign imports, but also against fashionable turns of phrase considered ungrammatical or incorrect (when criticizing the intensive use of "trop" in "elle est trop belle", we are, in a sense, (Morano -ist).’

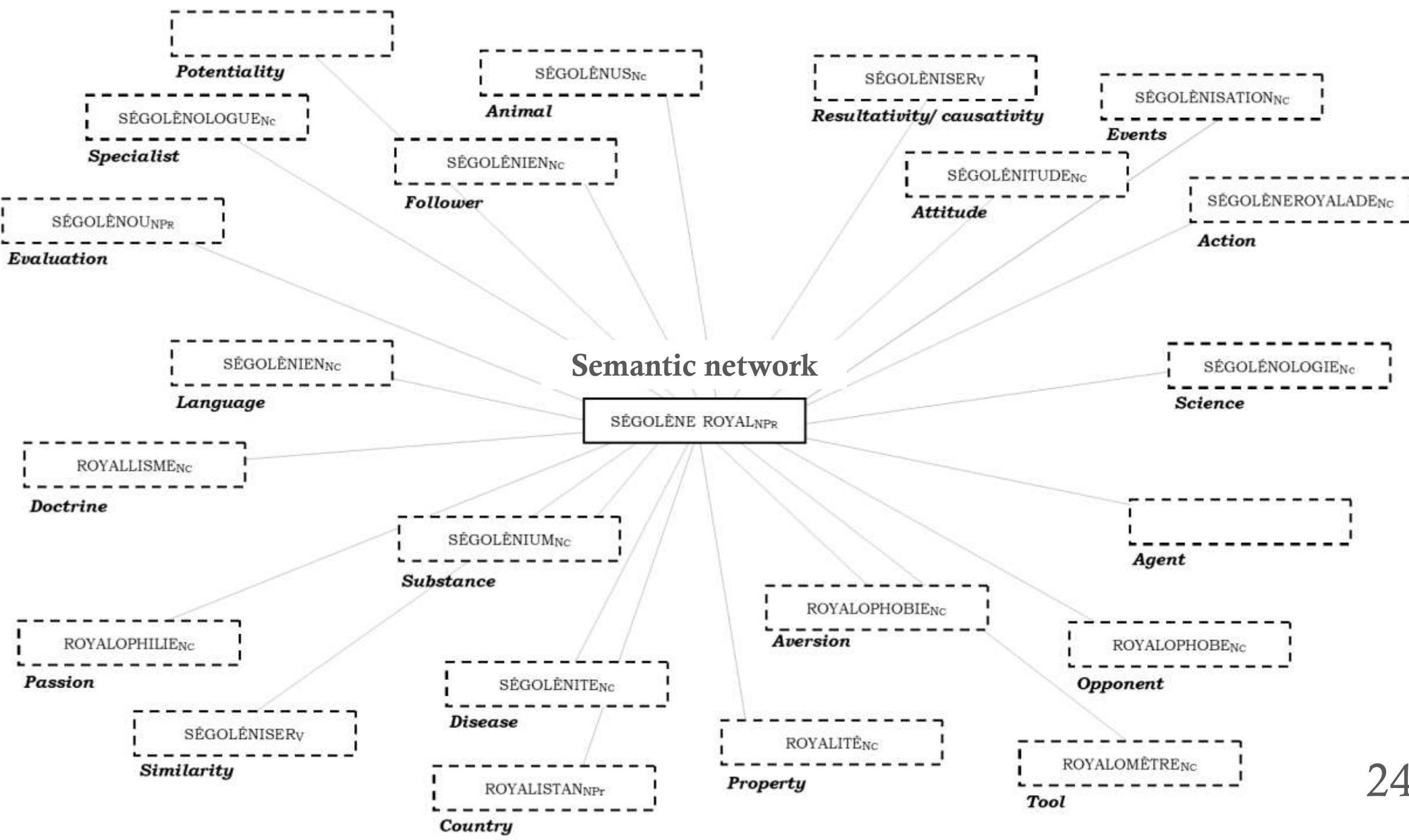
→ PPNs have a twofold meaning

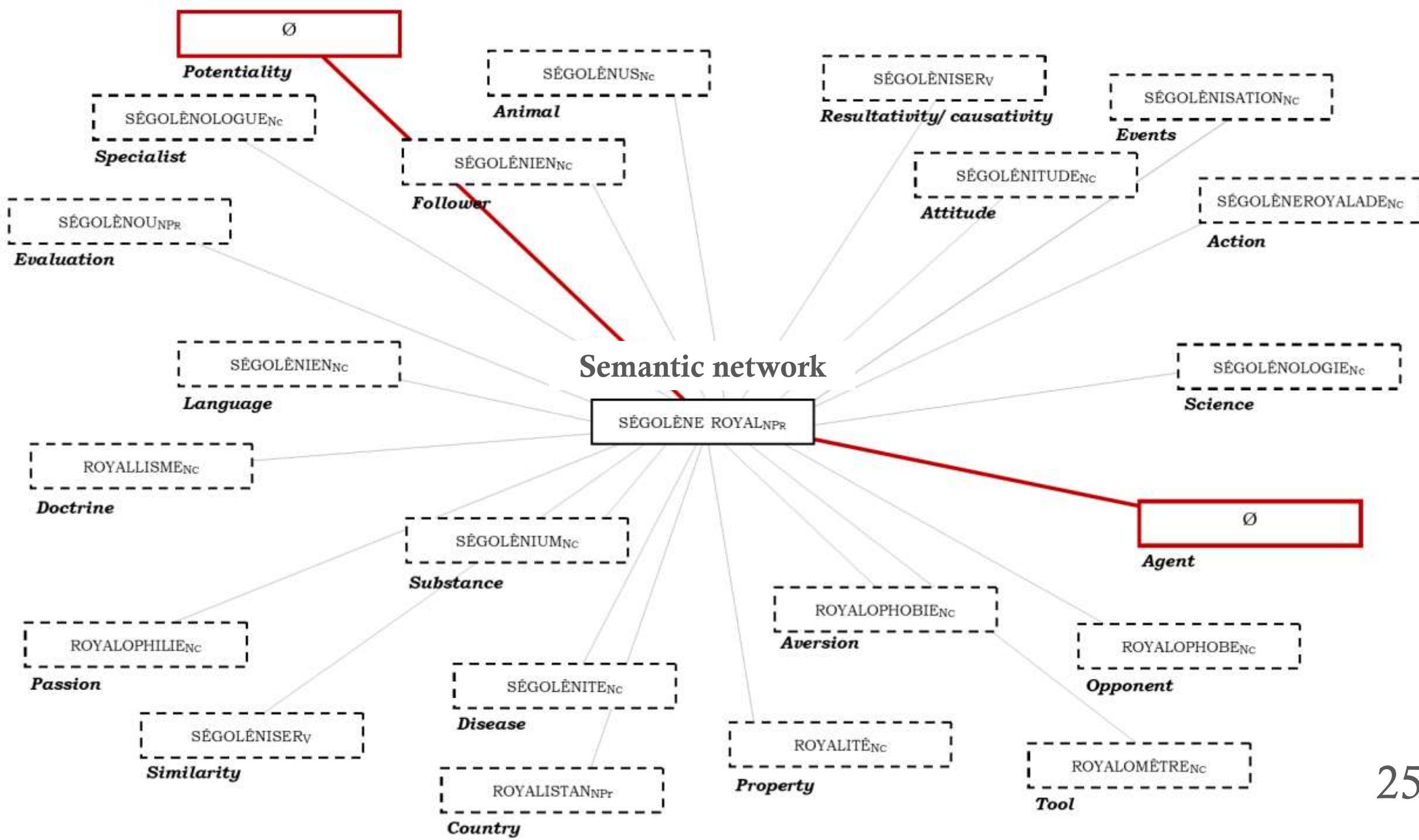
# Interpretations | derivational families

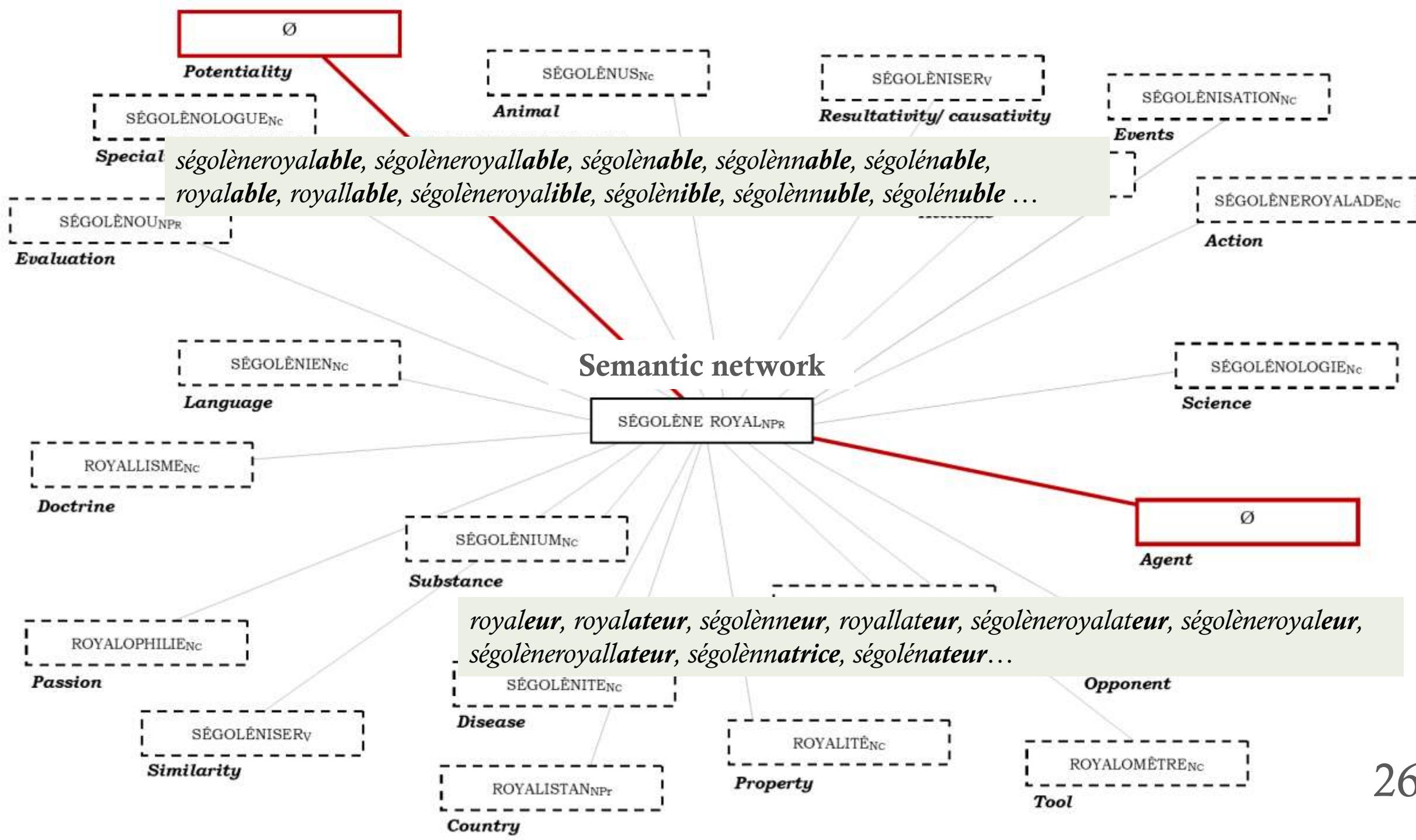
- iii. PPNs are the bases of expected (follower: ROYALISTE<sub>Nc</sub>) and unexpected complex words (substance: SÉGOLÈNIUM<sub>Nc</sub>)
- iv. Derivational families have structural similarities

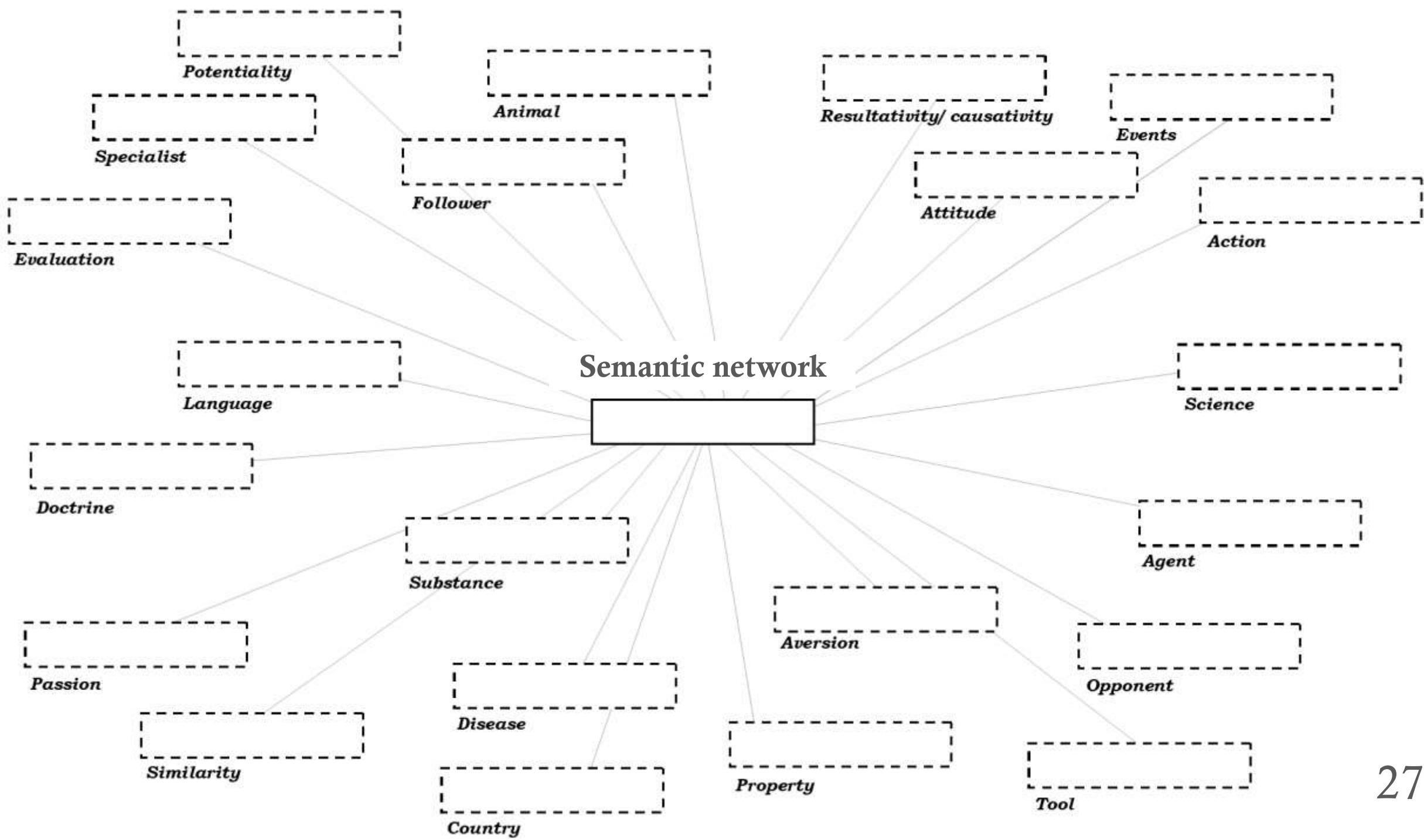
- Derivational families follow a number of invariants
  - Convert in terms of observability (attested or non-attested forms)
  - Visible if we separate the formal, syntactic and semantic dimensions

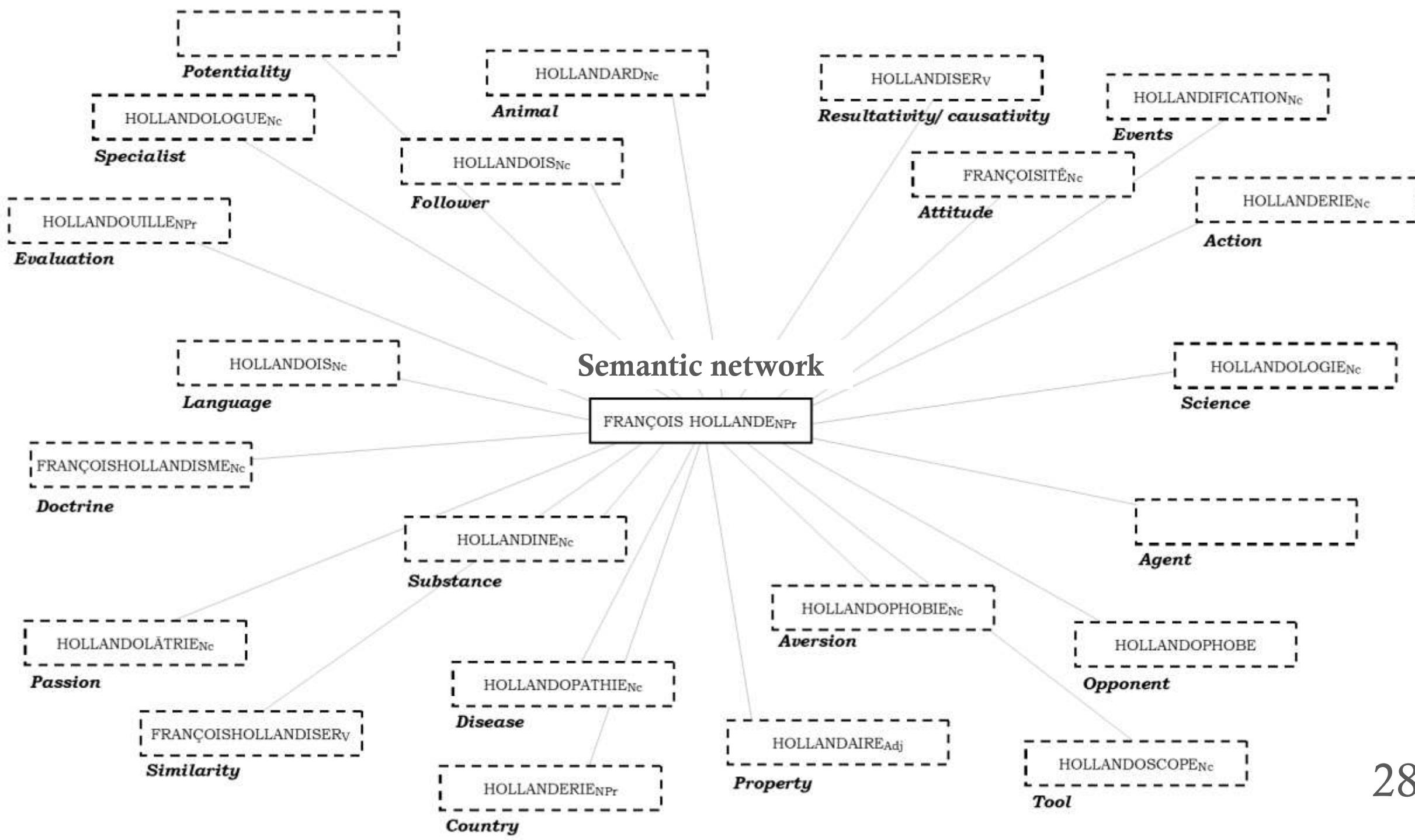


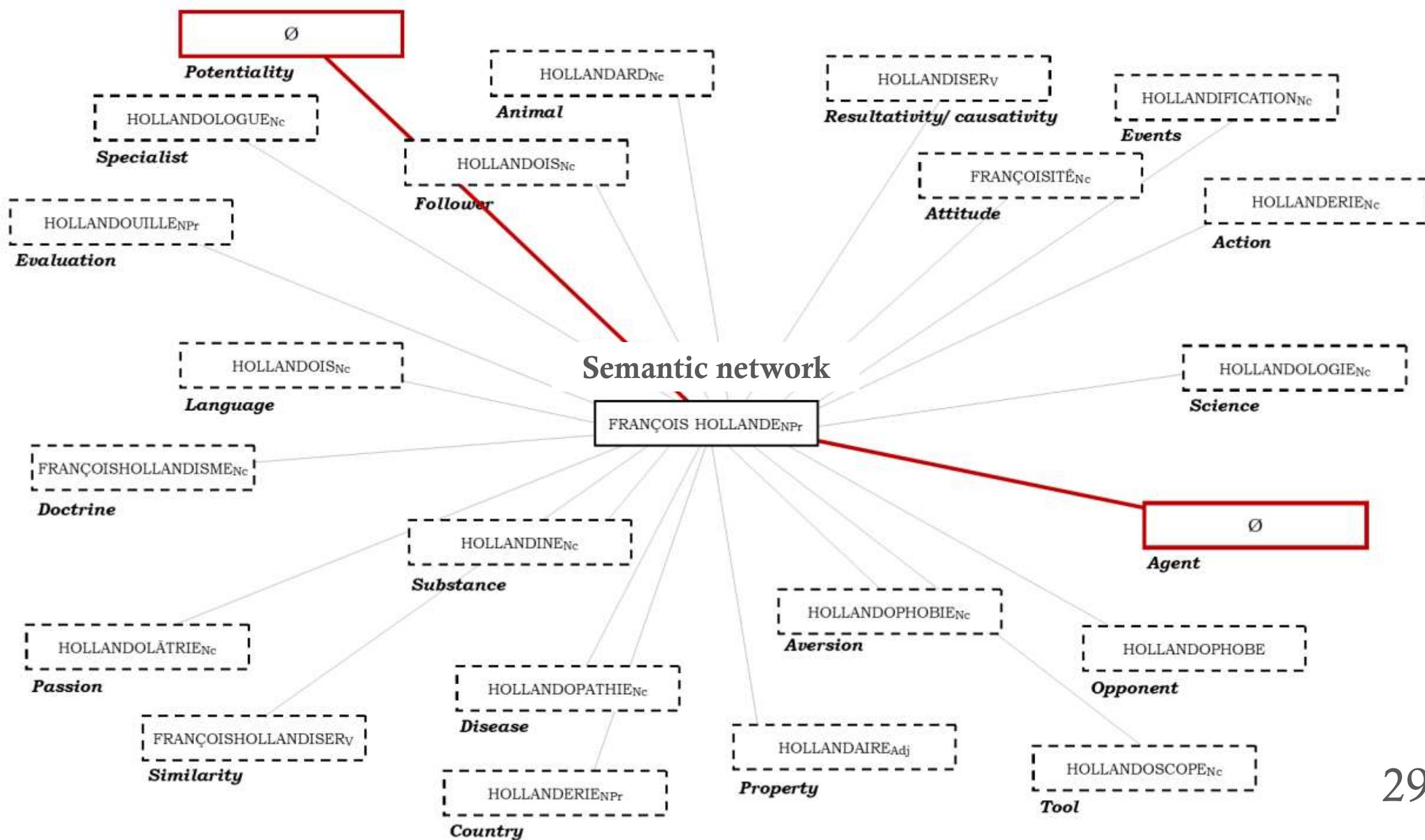


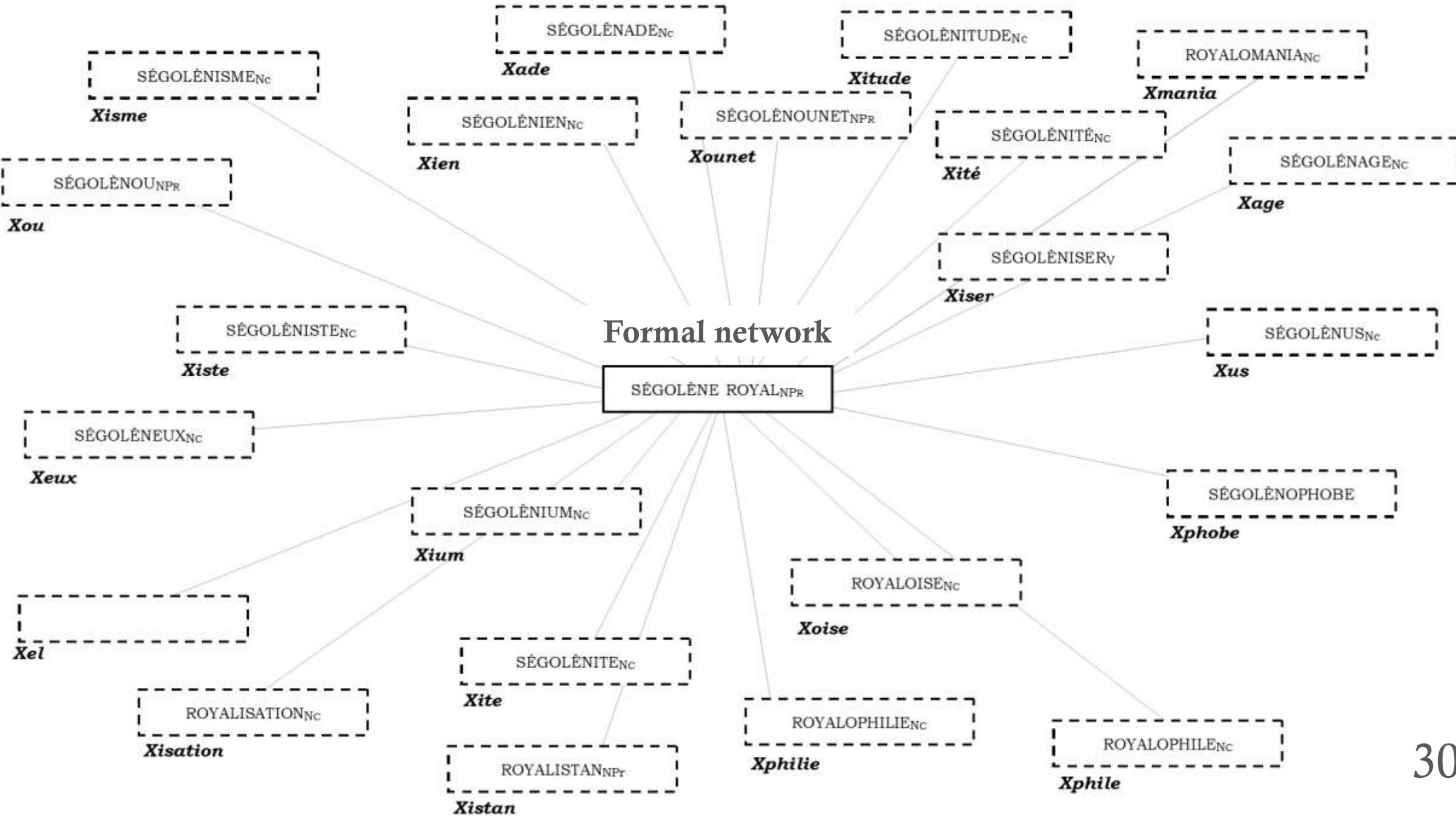


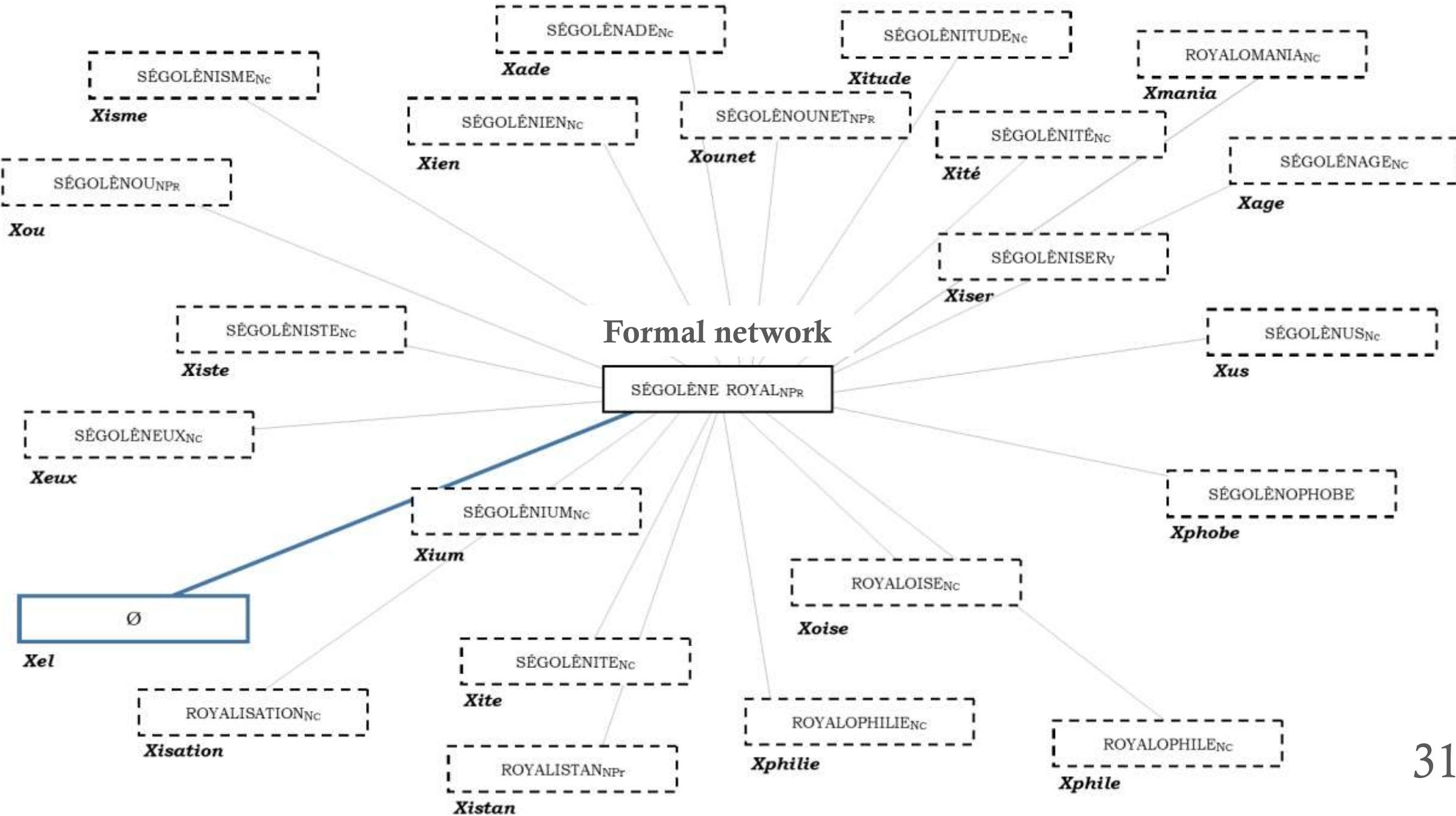


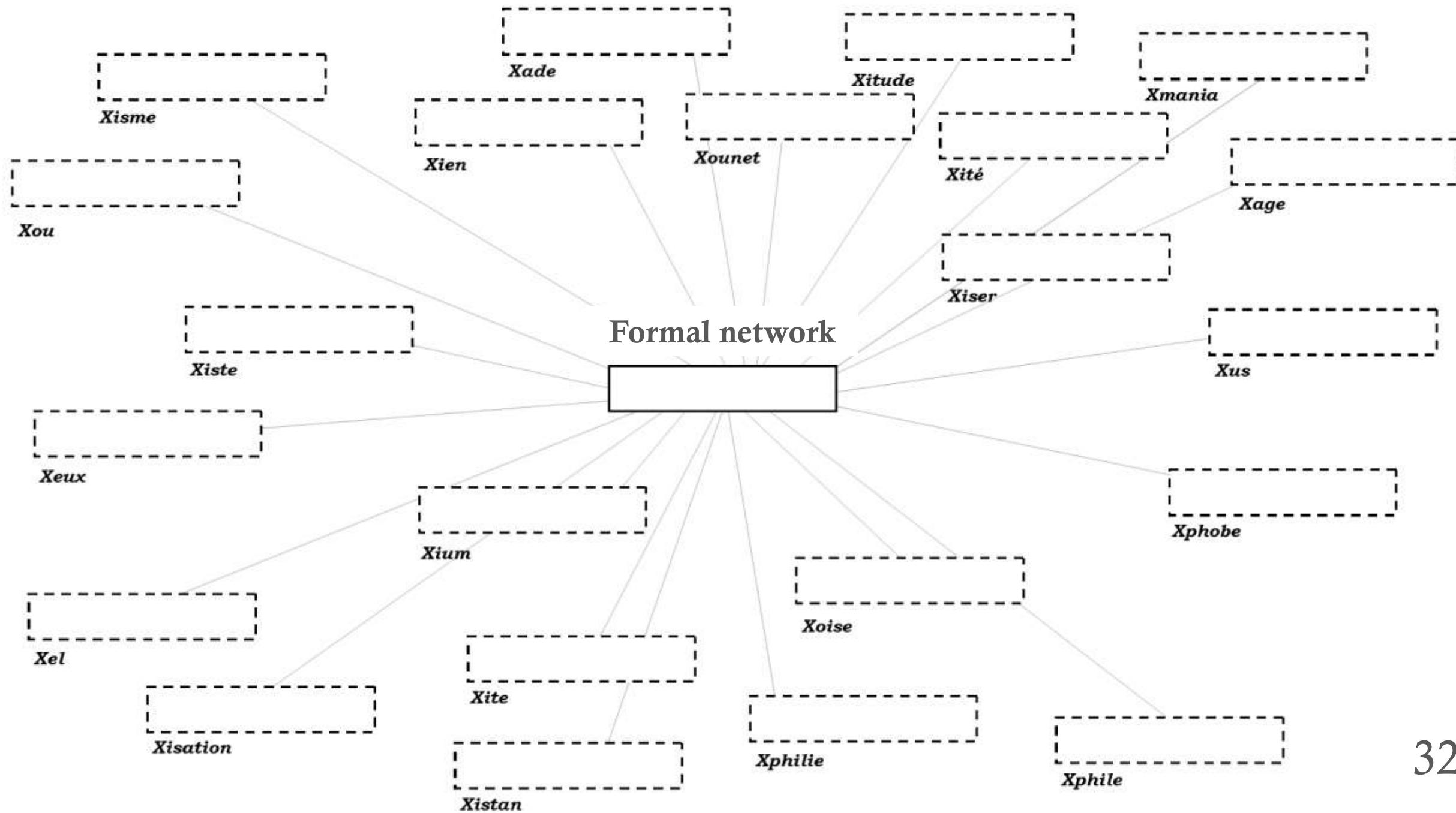


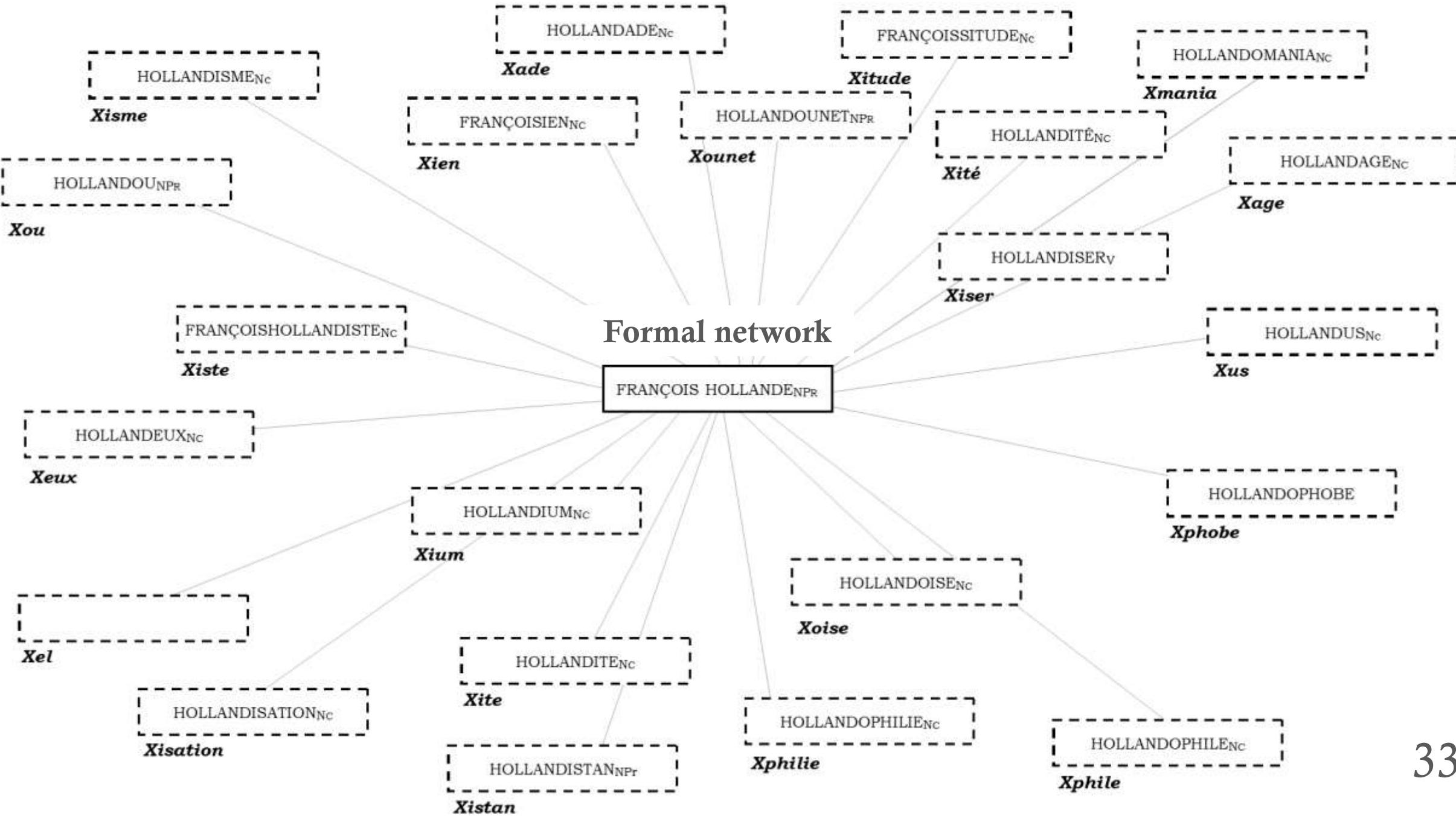


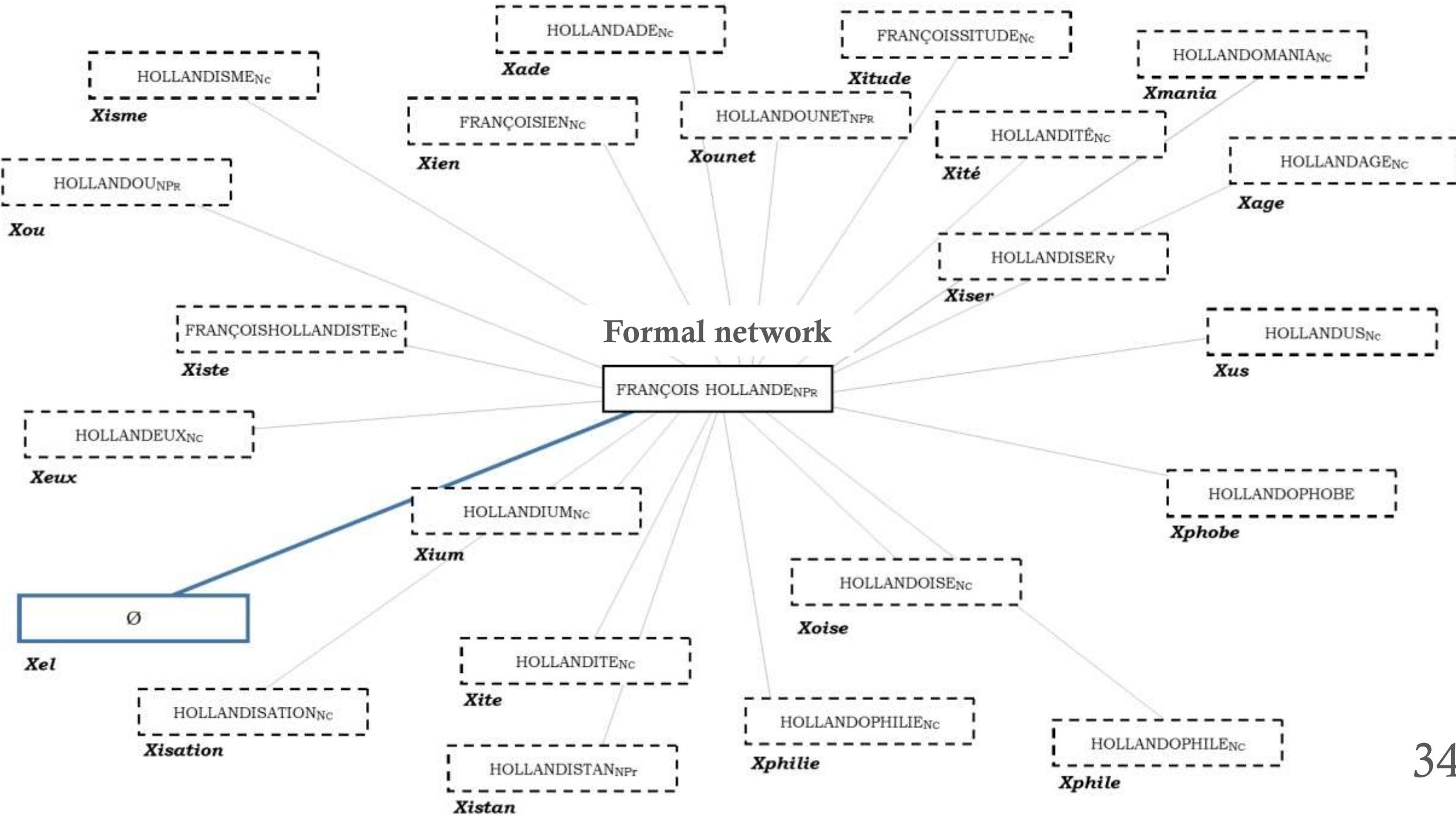






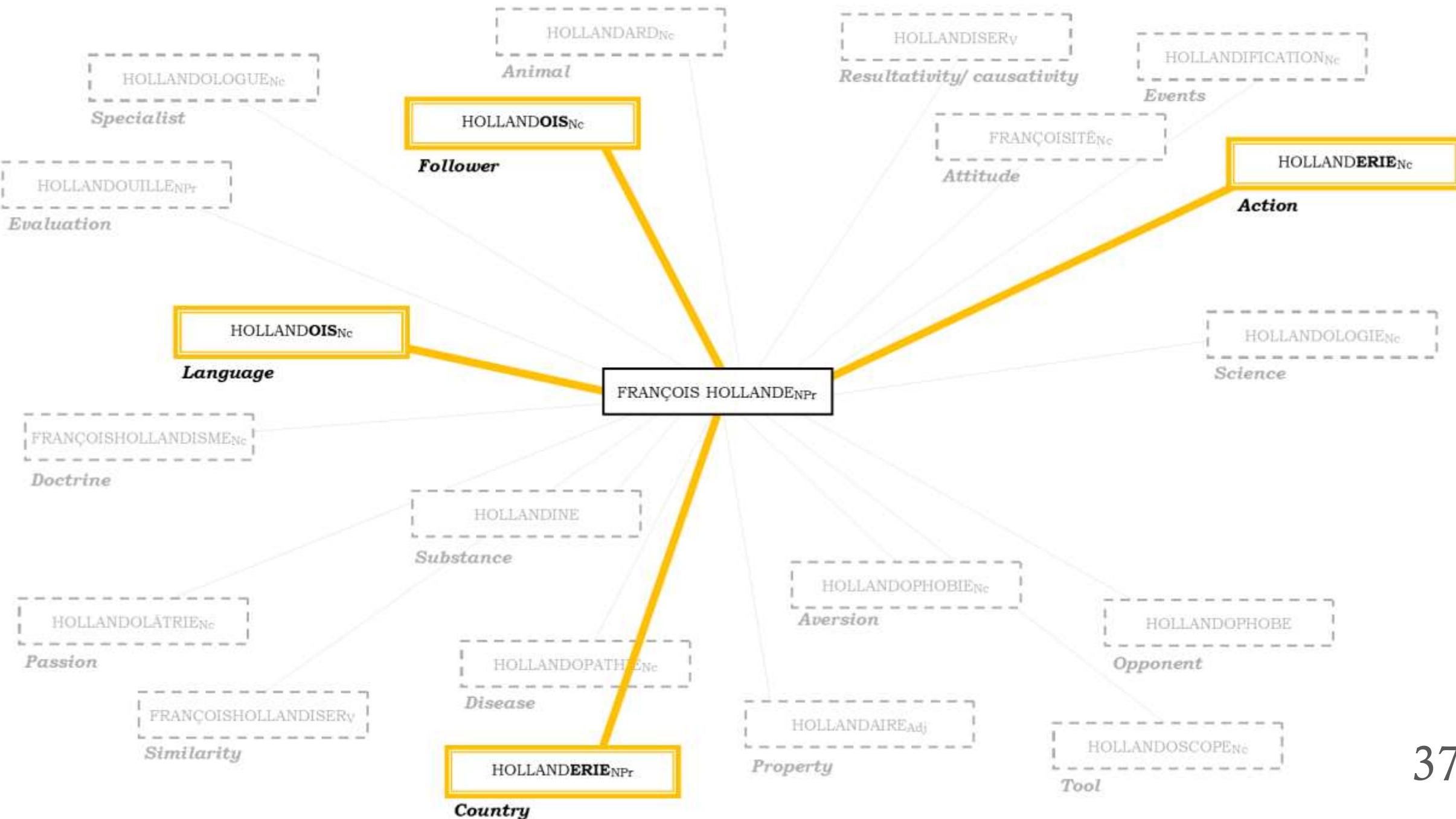


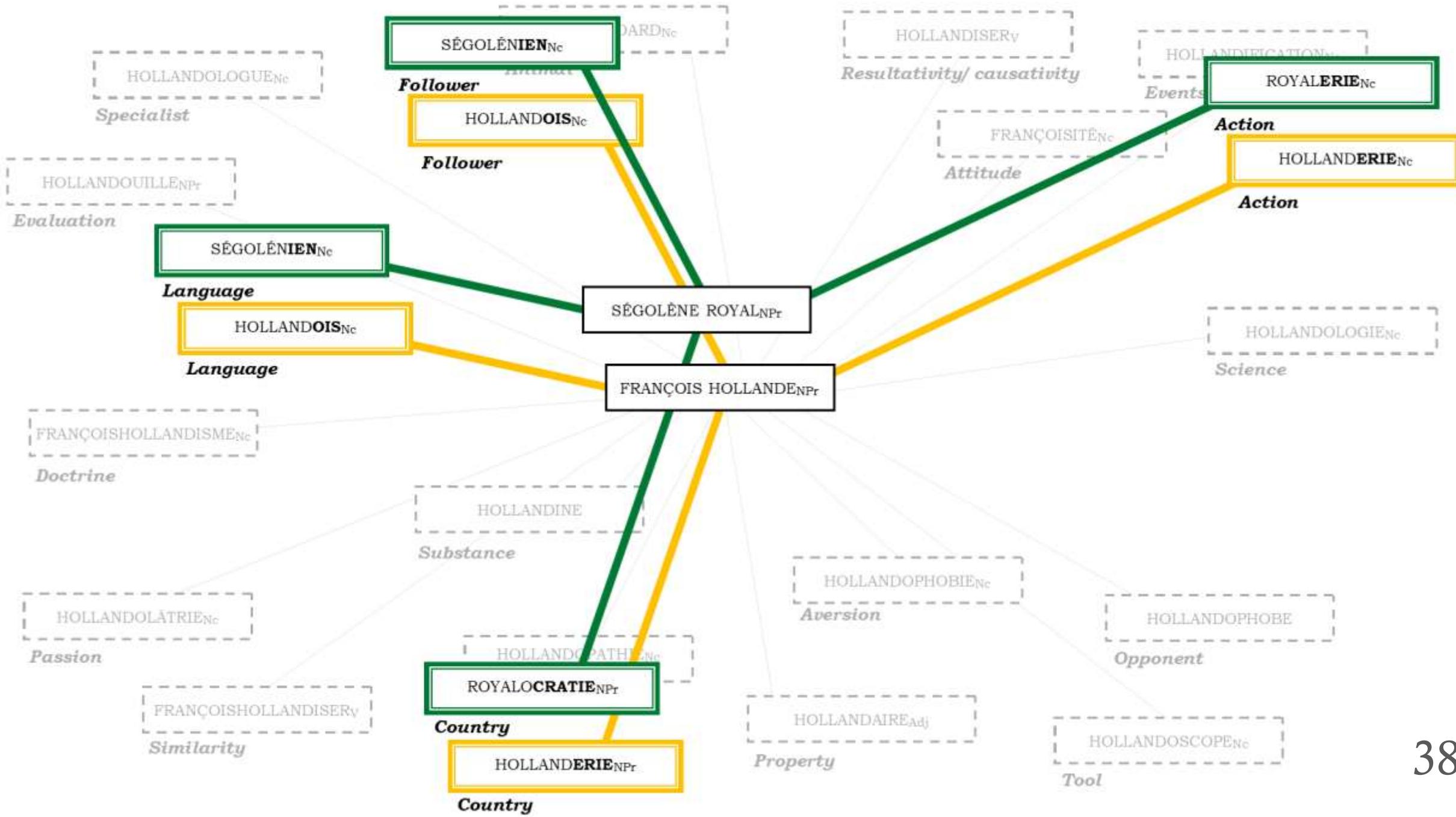


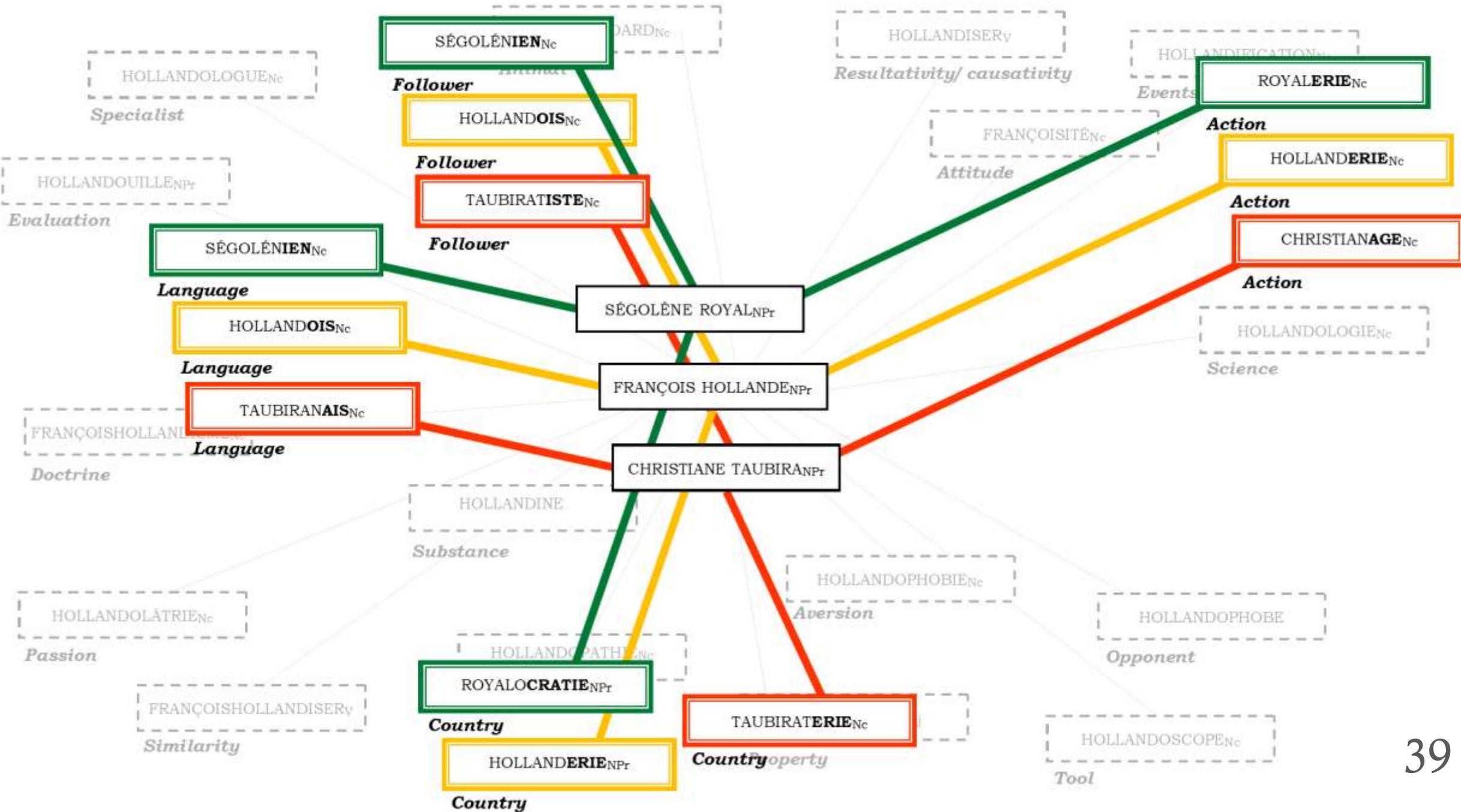


Why study the networks independently?









# Interpretations | paradigmatic links

## Characteristics of networks

- Ability to abstract derivational families in two predictable patterns
- Some parts of the networks are generalizable to all PPNs

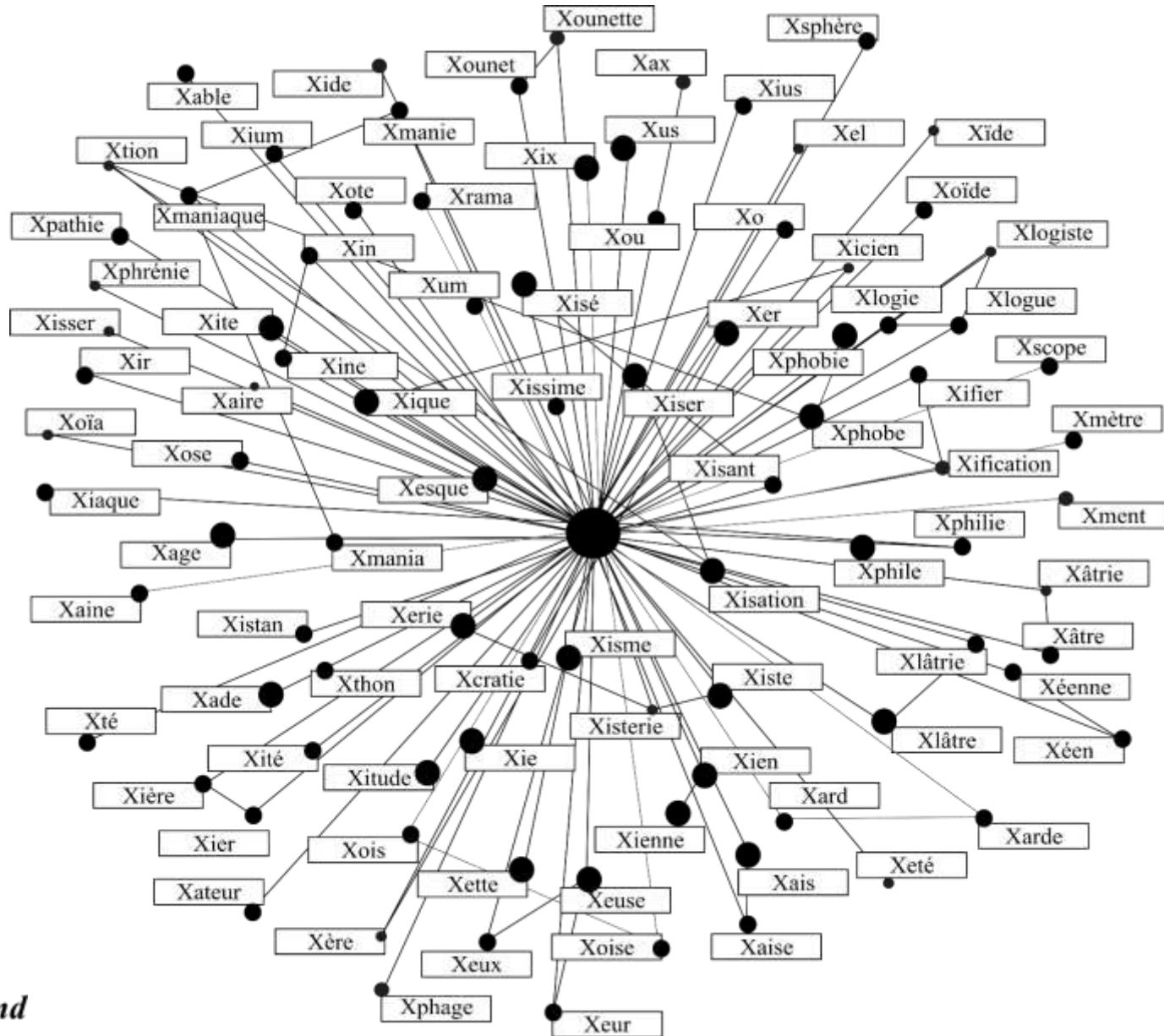
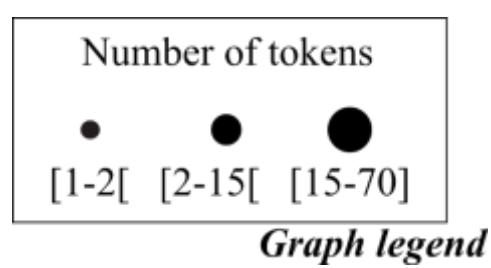
*'a paradigm is a series of morphologically related forms which share a base or base-type'*

*'a paradigm provides a typical set of realisations that are generalizable (to some extent) beyond the individual base word chosen as an illustration. This is the core of what a paradigm is, and is central to the notion of a paradigm as a pattern.'*

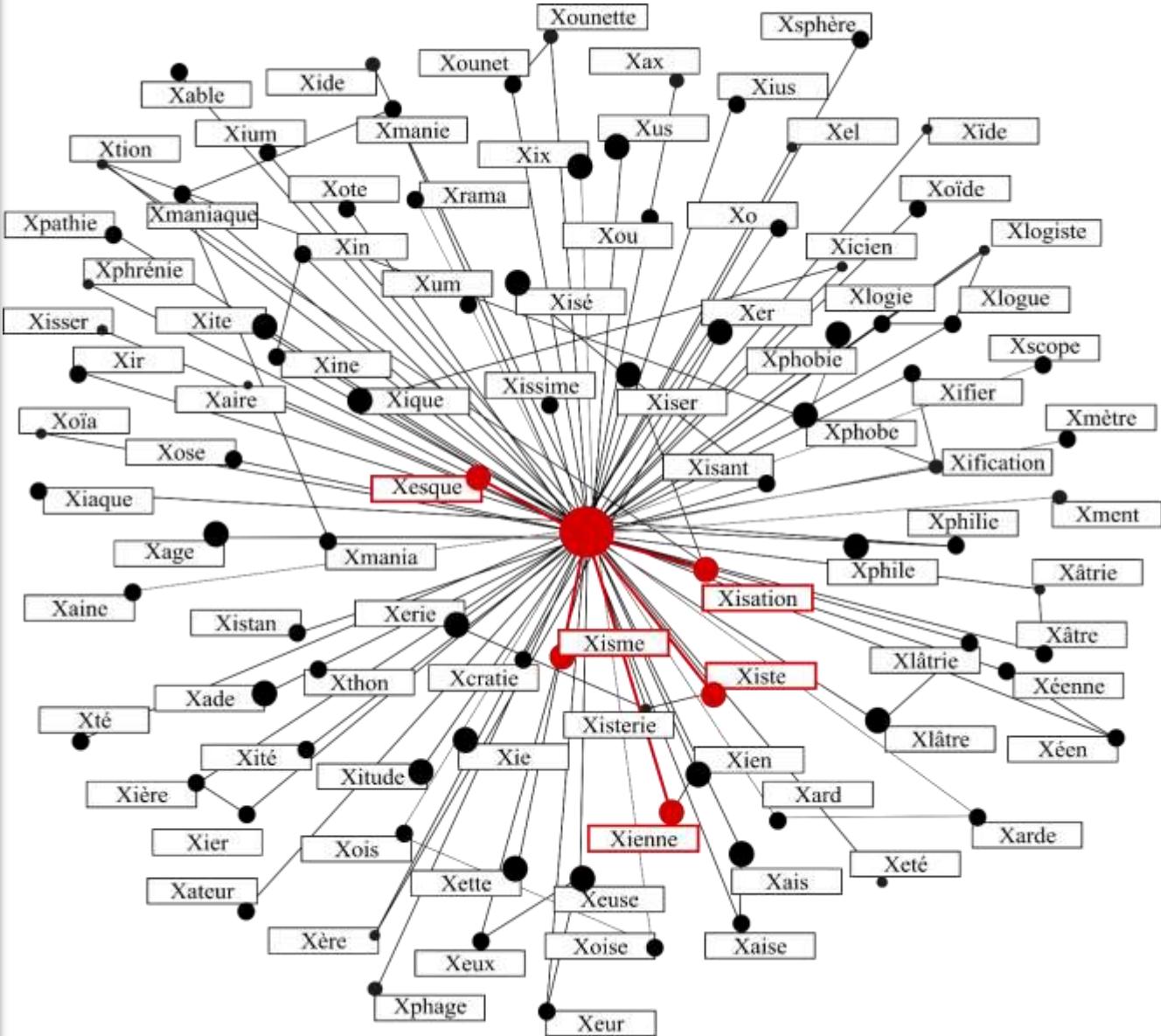
(Bauer, 1997: 244)

→ The formal and semantic networks **contain paradigmatic links**

# Interpretations | formal abstract network

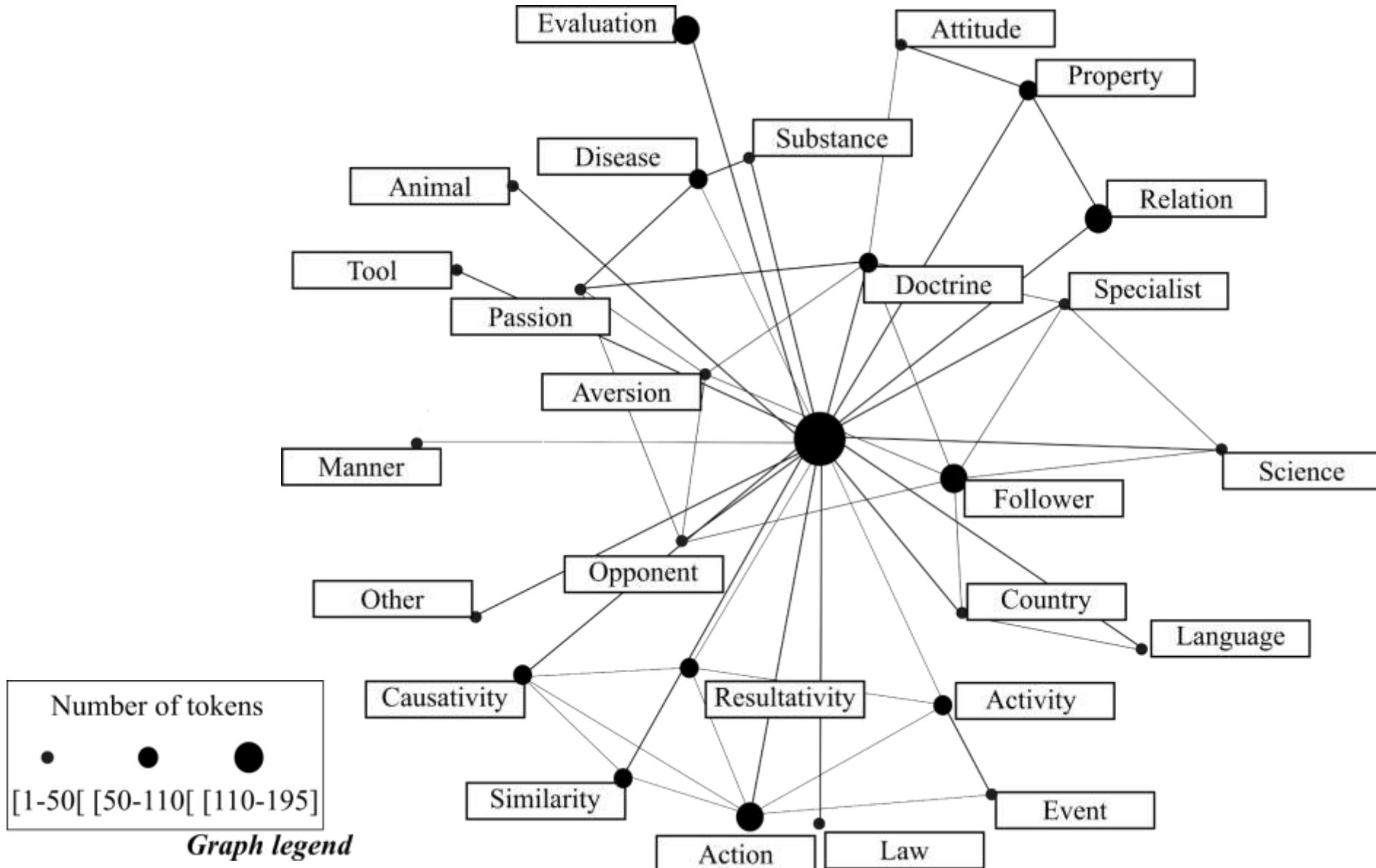


# Interpretations | formal abstract network



- Contains a set of construction models
  - Contains **paradigmatic links**
  - Makes it possible to explain the existence of lexemes without any attested bases
- v. Some words do not have any attested bases ( $\text{ROYALIFIER}_V > \text{ROYALIFICATION}_{N_c}$ )

# Interpretations | semantic abstract network



# Interpretations | semantic abstract network

- Contains a set of semantic categories
- Contains **paradigmatic links**

*'the forms in a paradigm are semantically related by more than the meaning of the base.'*  
(Bauer, 1997: 244)

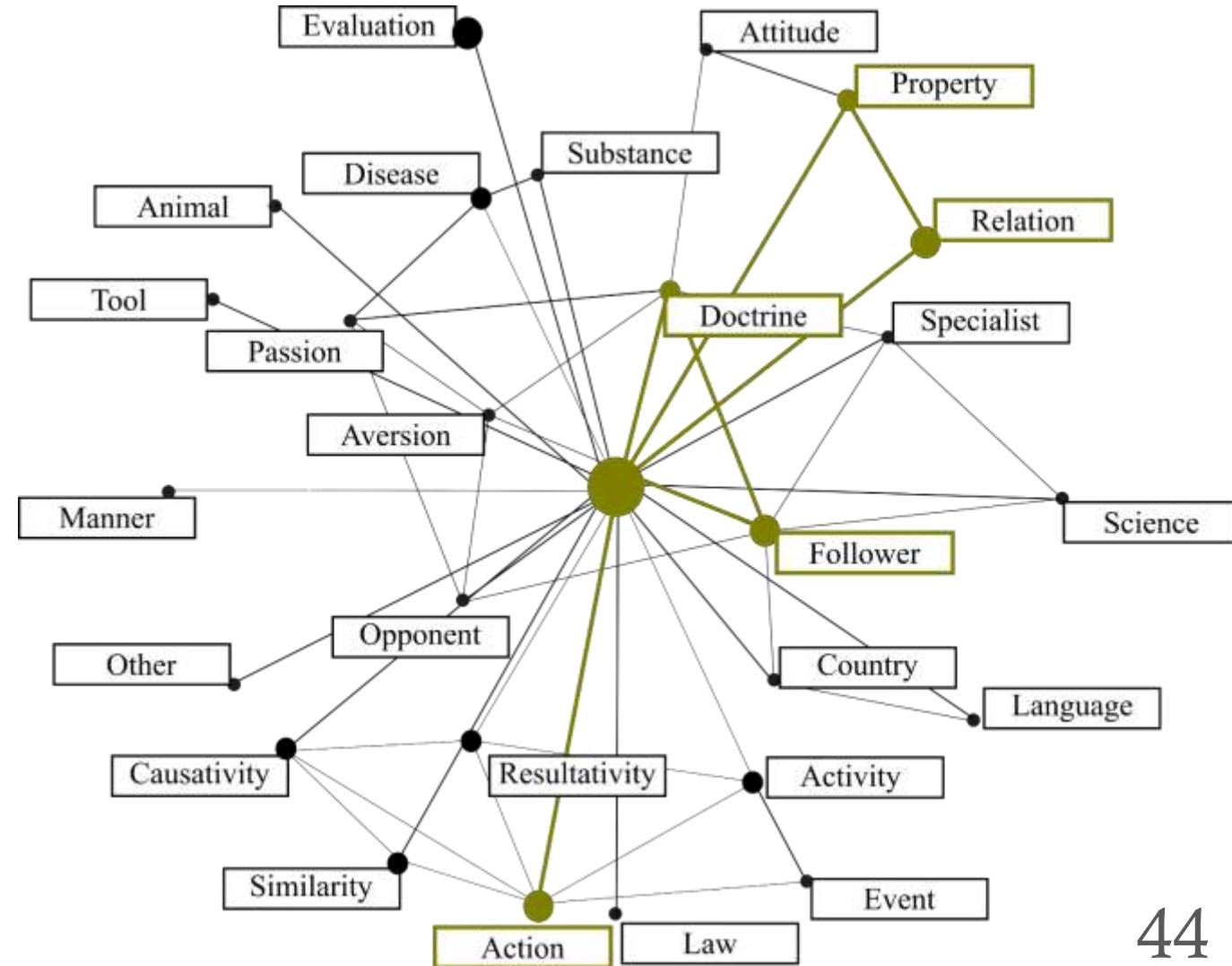
HOLLANDIEN<sub>Nc</sub>

'follower of François Hollande/hollandisme'

HOLLANDISME<sub>Nc</sub>

'ideology of François Hollande/ hollandiens '

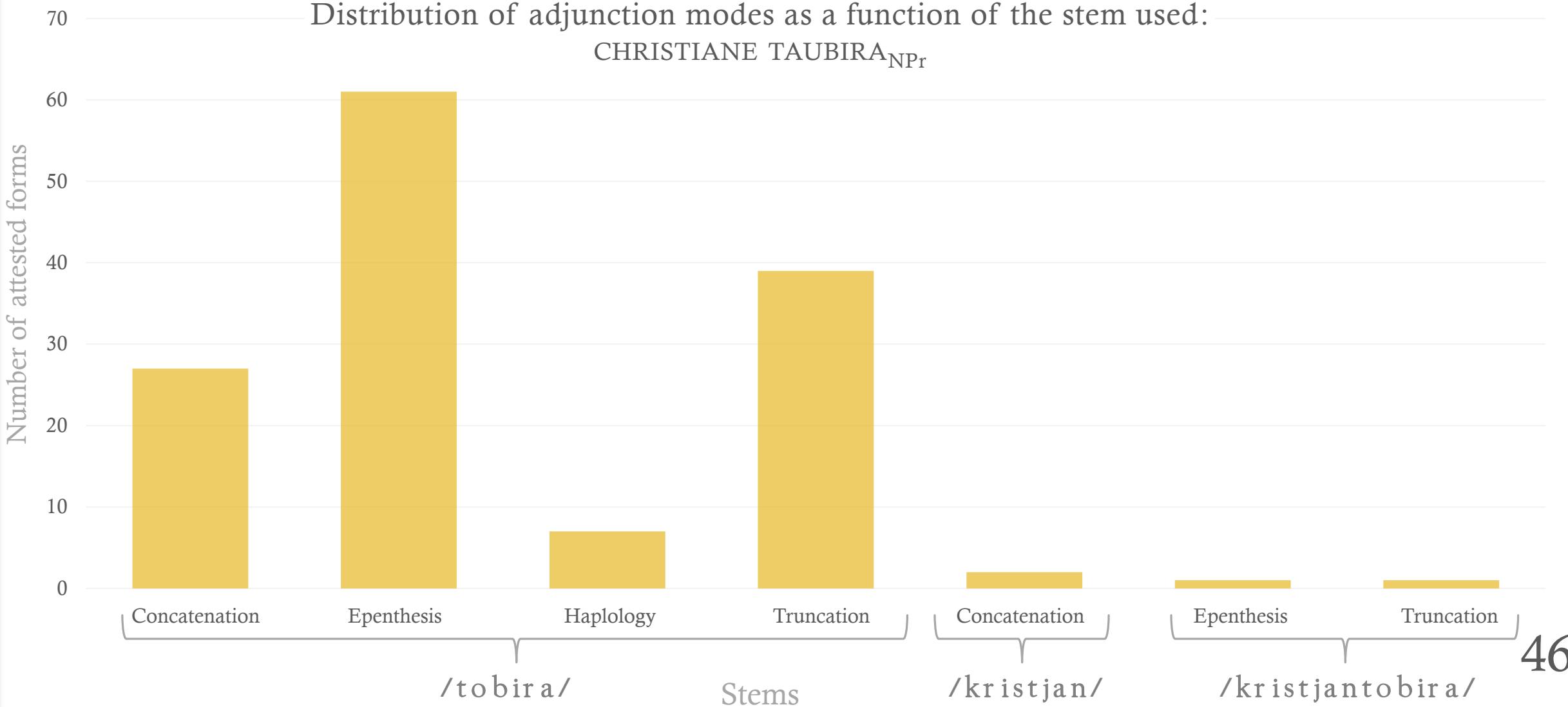
(cf. Marxism/Marxist: Bauer, 1997 ; Booij, 2008: 33; Roché, 2011; Štekauer, 2014: 363; among others)



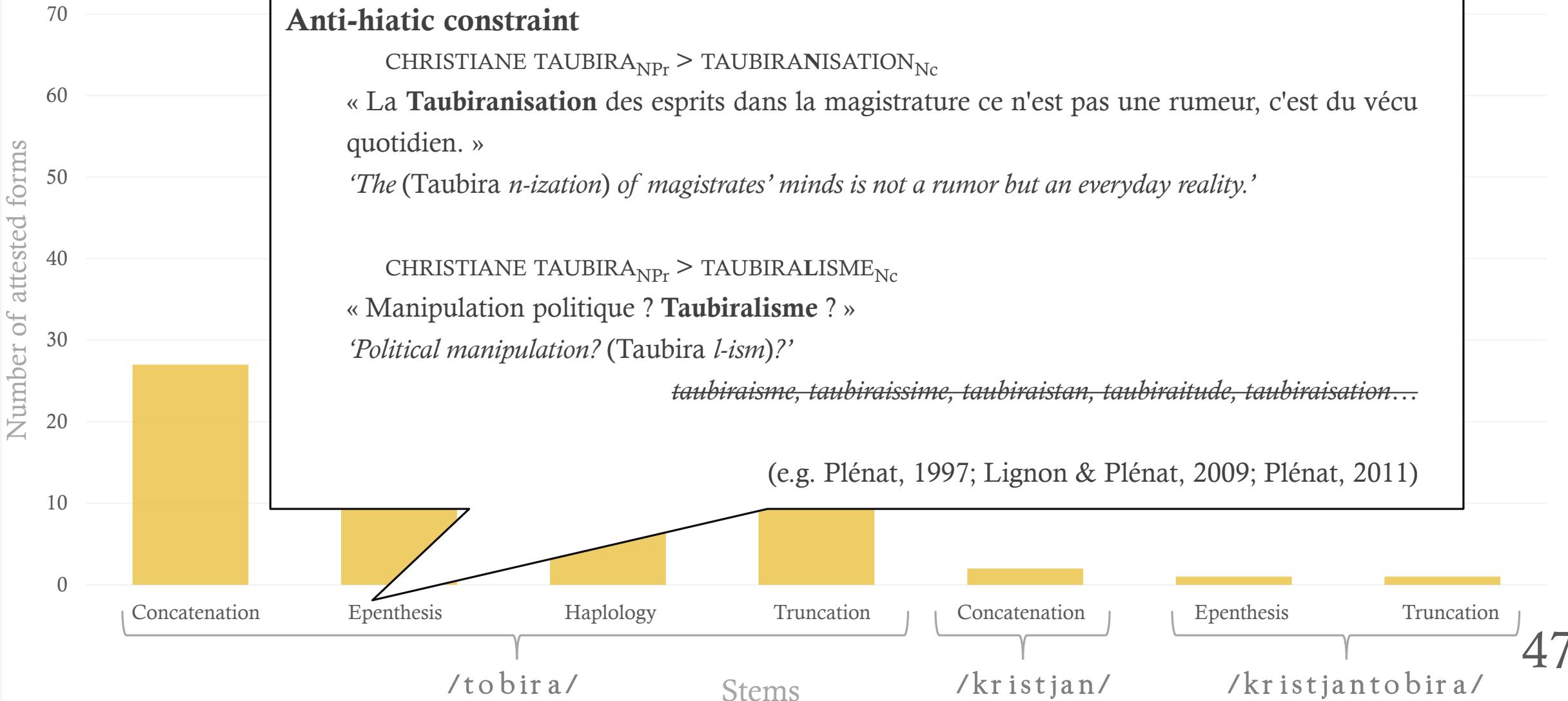
# Interpretations | network realization

- The network realization obeys **morpho-phonological constraints** which determine:
  - The choice of stem in the stem collection (e.g. GEORGES TRON<sub>NPr</sub>; fidelity and size constraints cf. Villoing, 2012; Roché & Plénat, 2014)
  - The stem form in the complex word (i.e. adjunction mode)

# Interpretations | network realization



# Interpretations | network realization

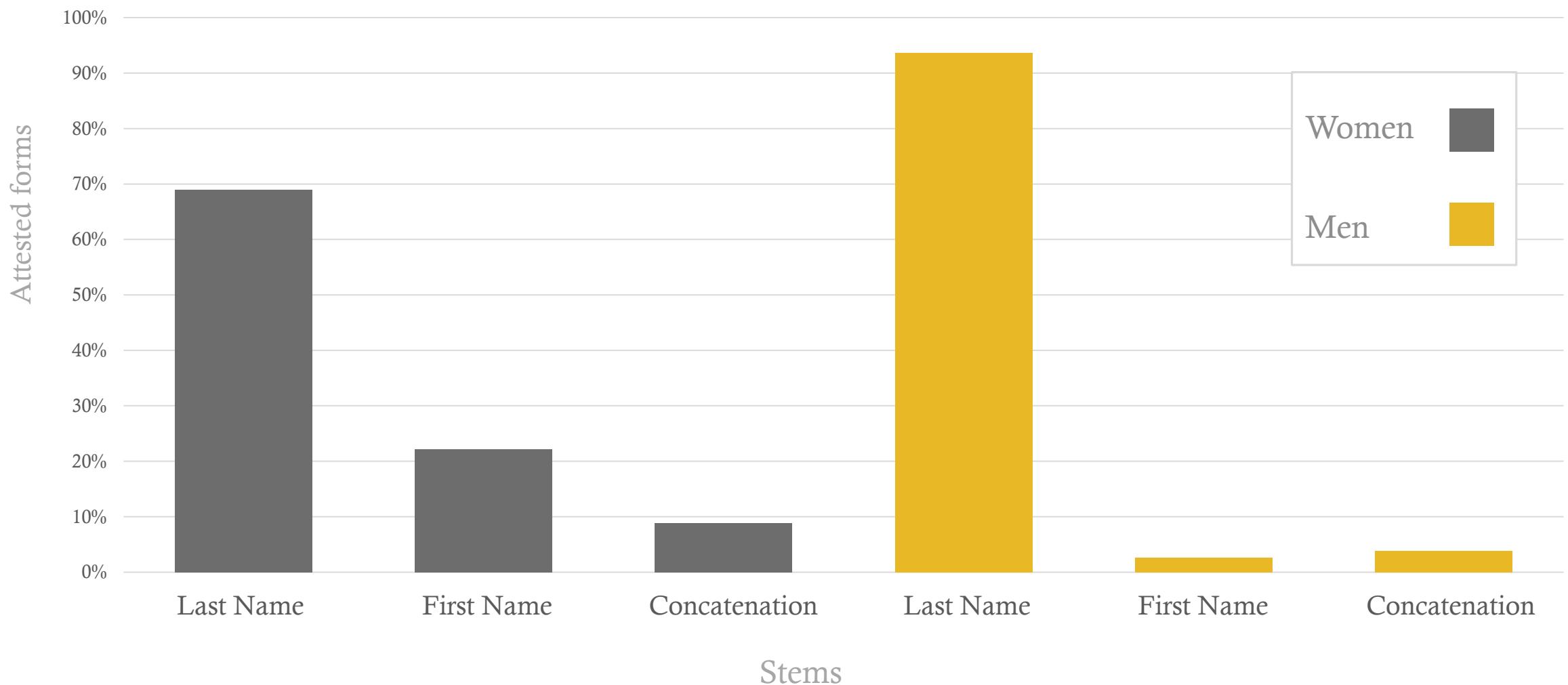


# Interpretations | network realization

- The network realization obeys **morpho-phonological constraints** which determine:
  - The choice of stem in the stem collection (e.g. GEORGES TRON<sub>NPr</sub>; fidelity and size constraints cf. Villoing, 2012; Roché & Plénat, 2014)
  - The stem form in the complex word (i.e. adjunction mode)
- The network realization obeys **extra-linguistic constraints** which determine:
  - The choice of stem in the stem collection
  - The number and the semantic category of complex words

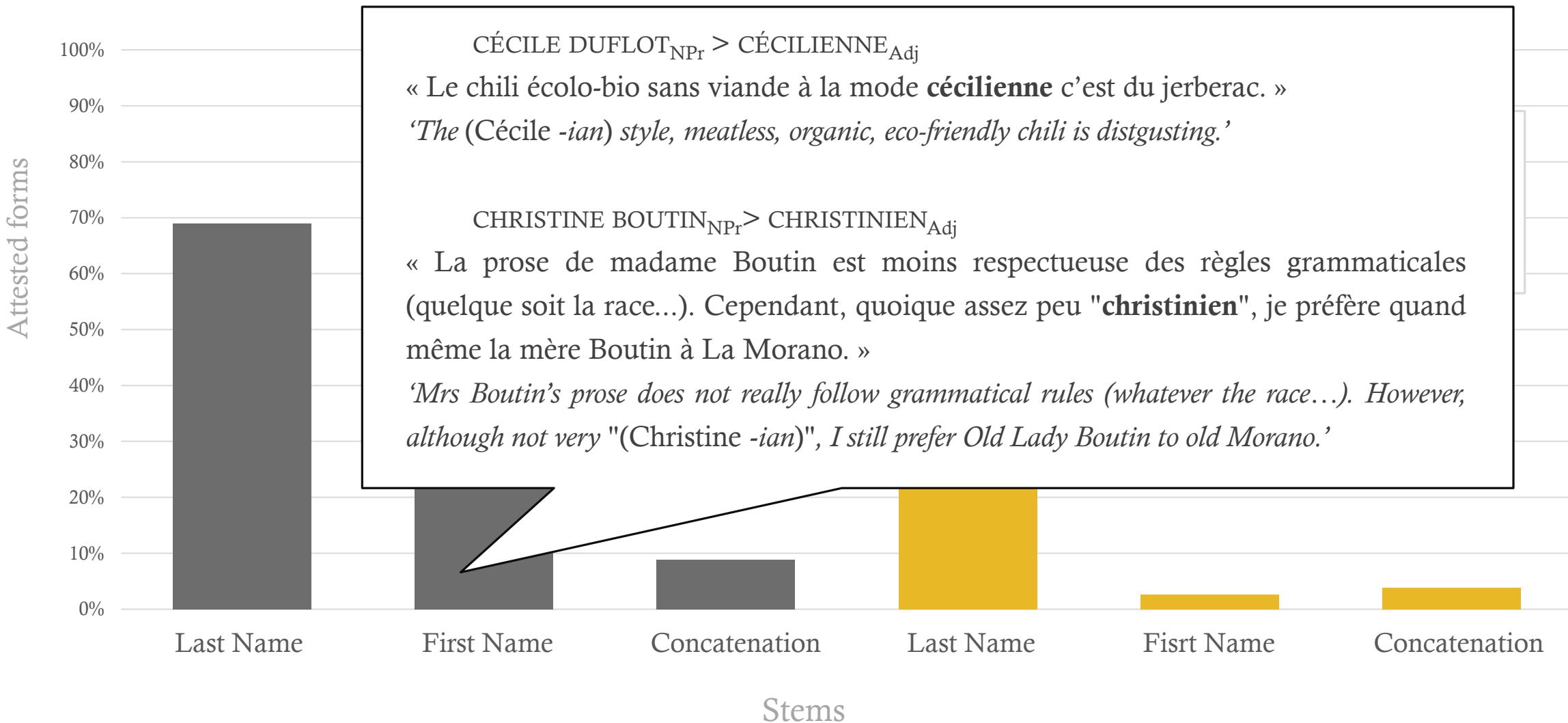
# Interpretations | network realization

Percentage of stem use as a function of the gender of the public figure

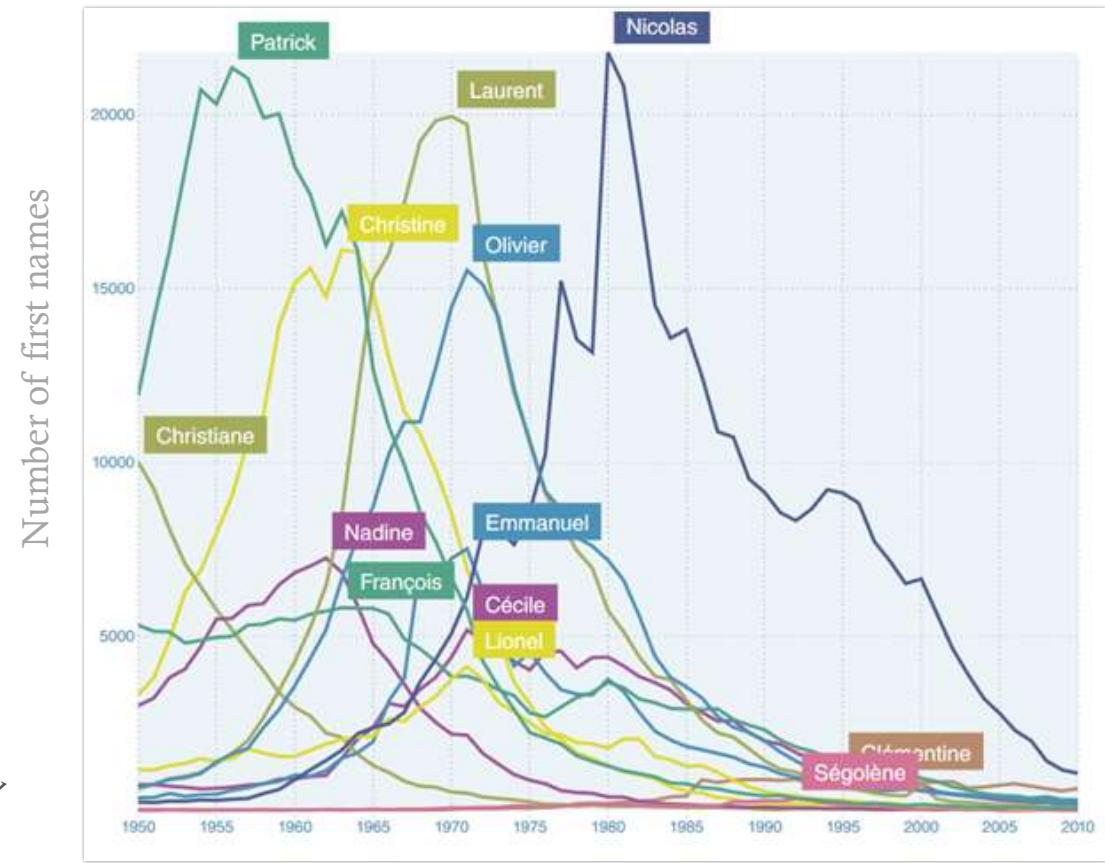
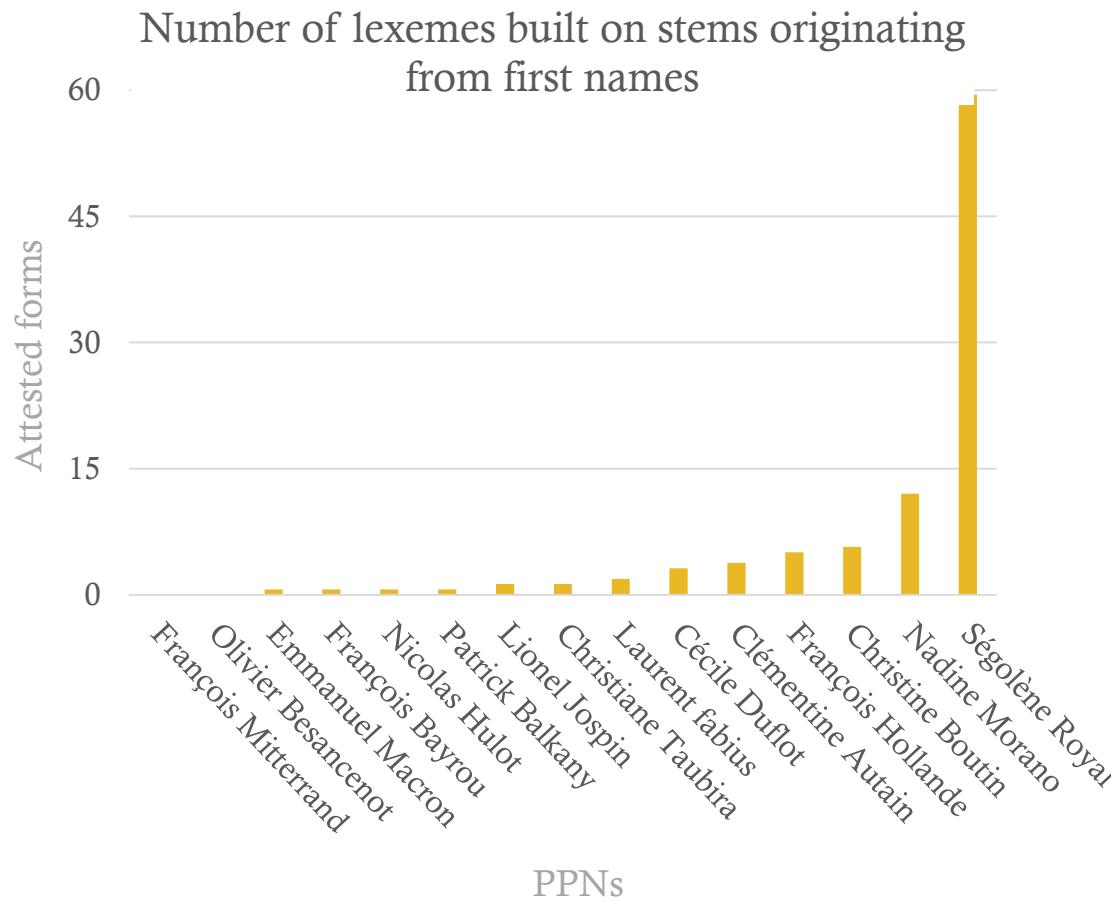


# Interpretations | network realization

Percentage of stem use as a function of the gender of the public figure



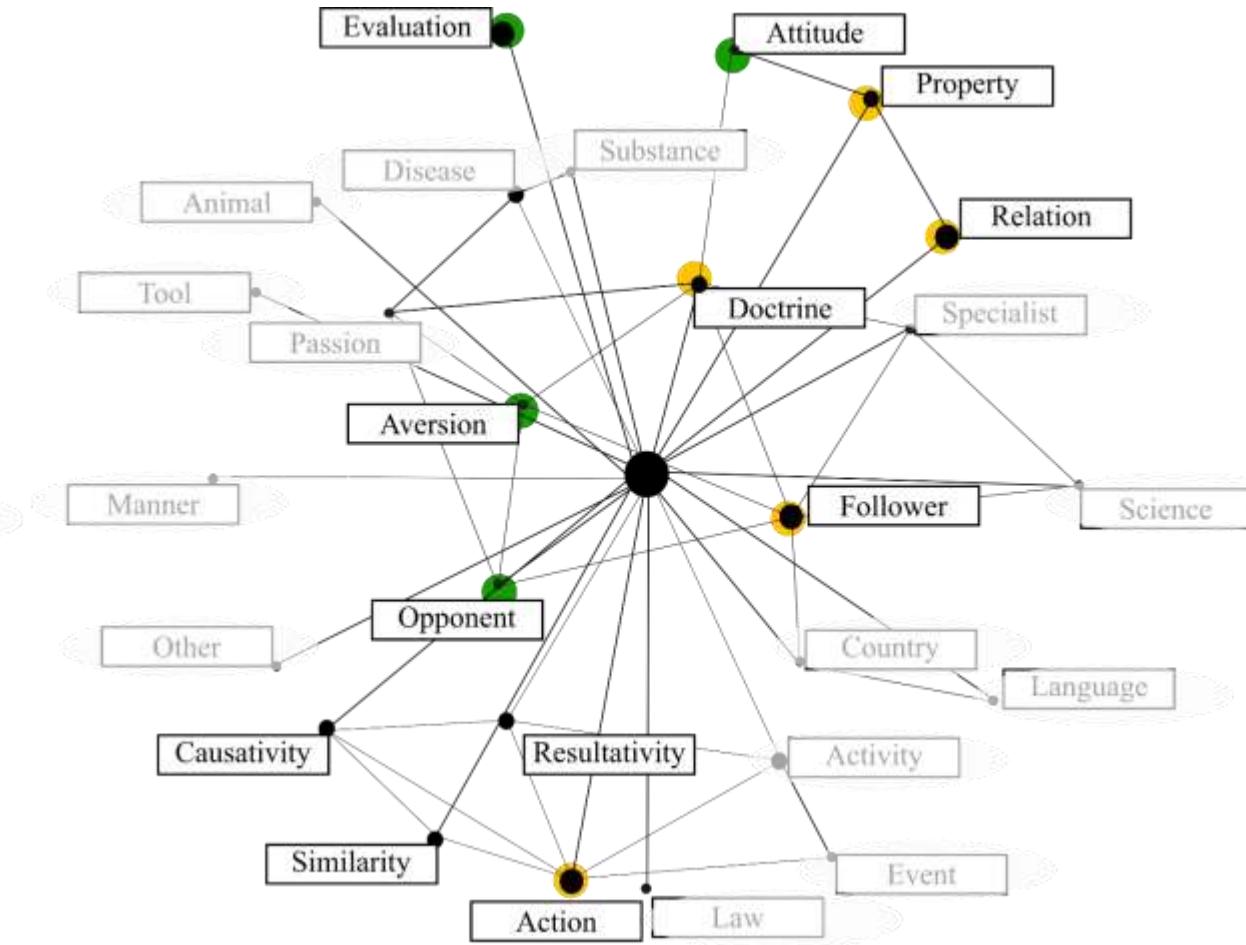
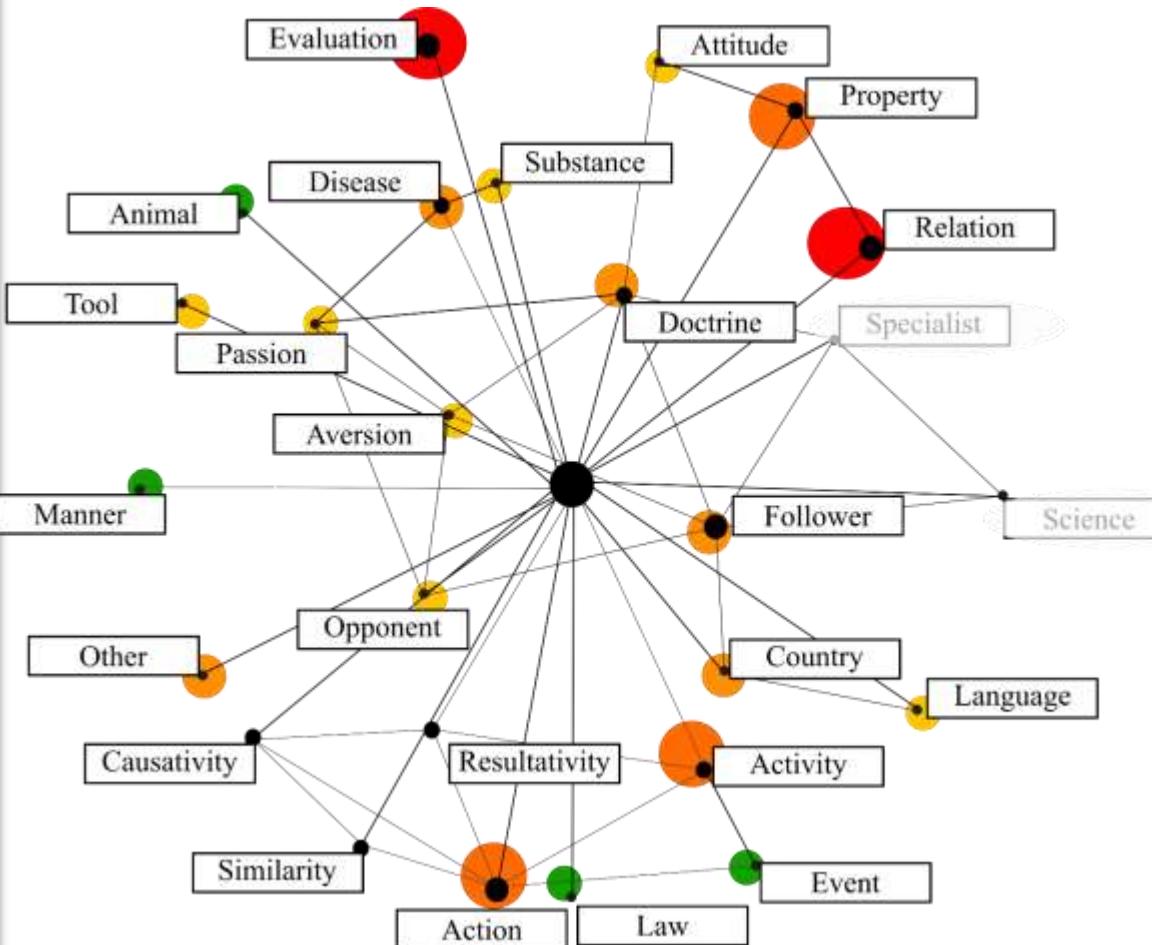
# Interpretations | network realization



First names given to newborns in France between 1950 and 2010

cf. Insee, 2010

# Interpretations | network realization



*Number of tokens*

All PPNs



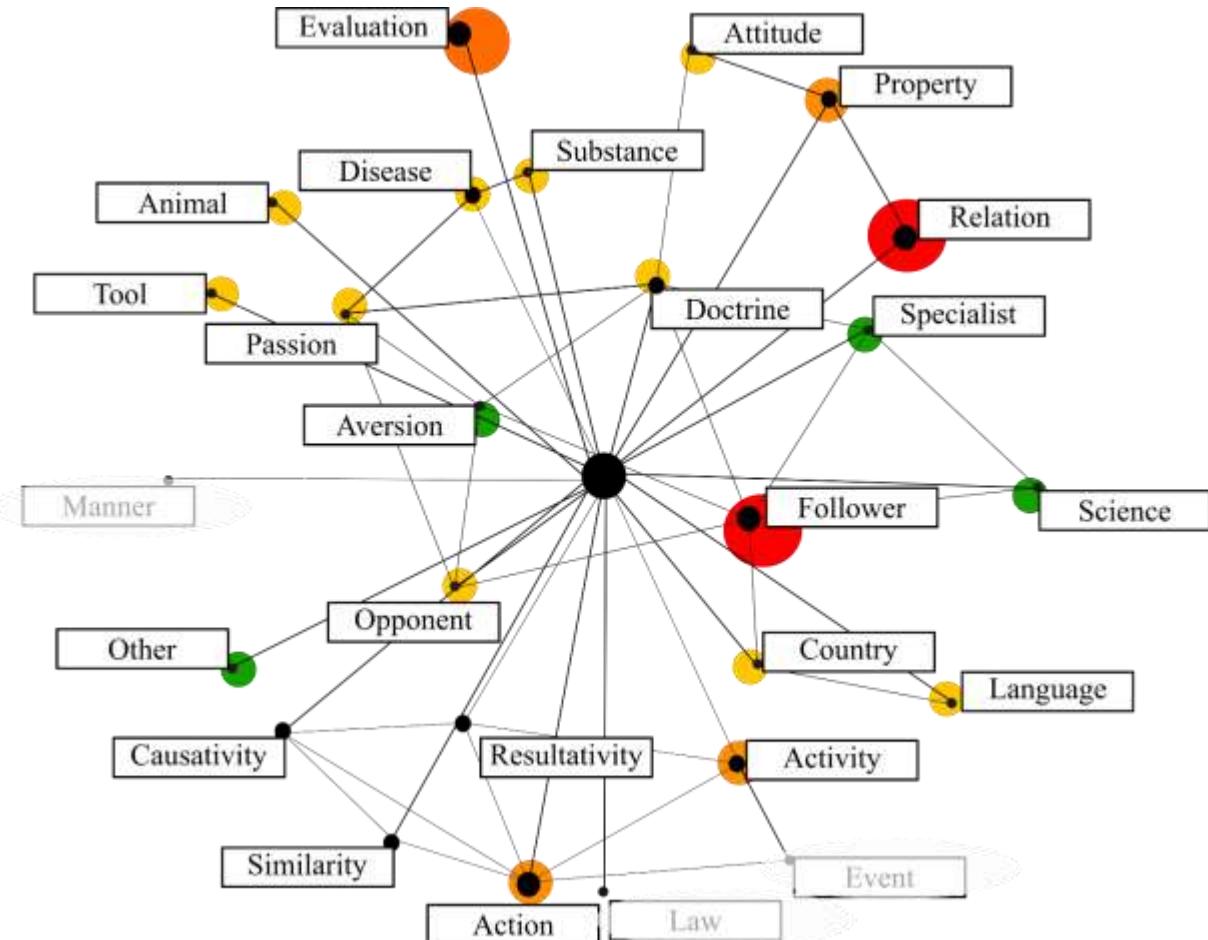
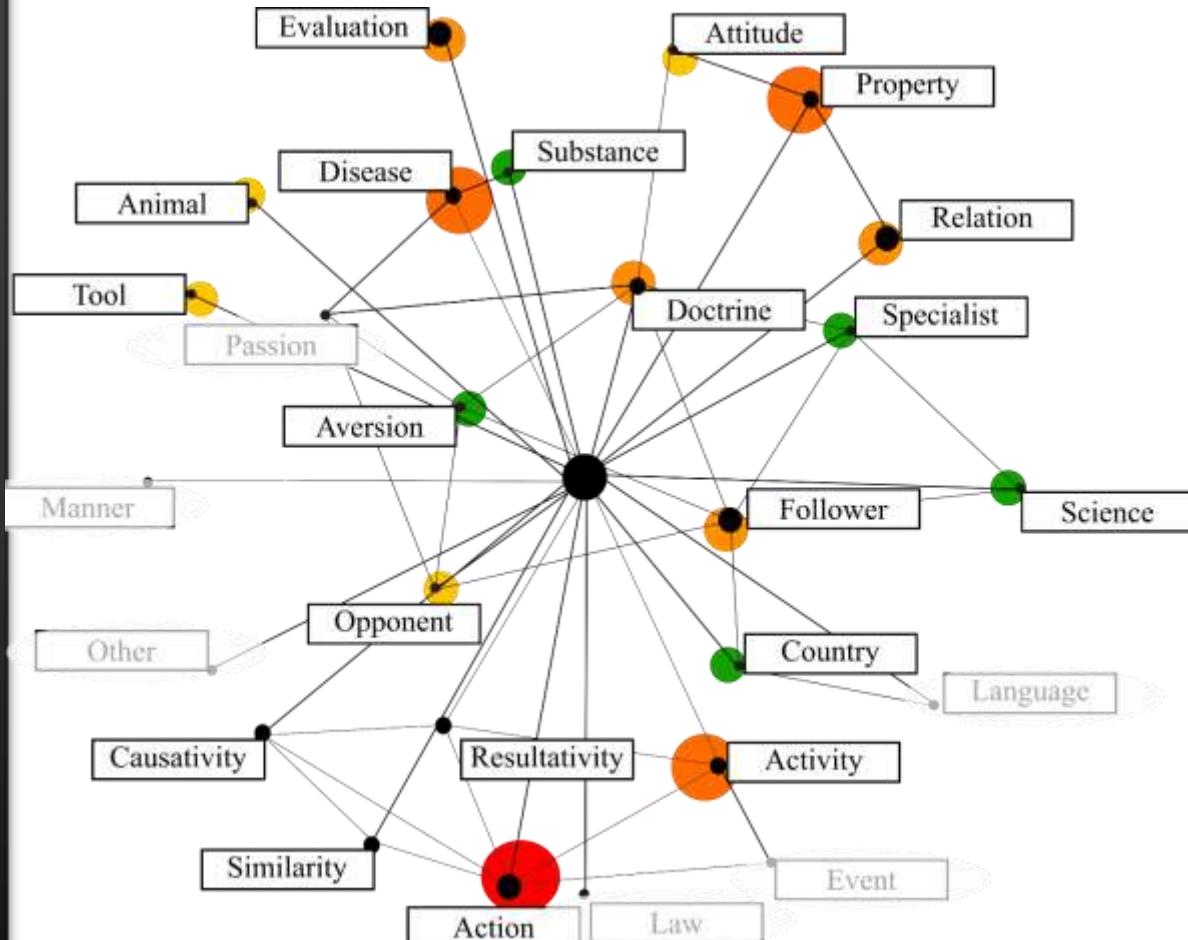
[1-50[ [50-110[ [110-195]

PPN



[1] ]1-5] ]5-10] ]10-15] ]15-25]

# Interpretations | network realization



NADINE MORANO<sub>NPr</sub>

*Number of tokens*

All PPNs



PPN



FRANÇOIS HOLLANDE<sub>NPr</sub>

# Conclusion

**PPNs have a stem collection and a twofold meaning**

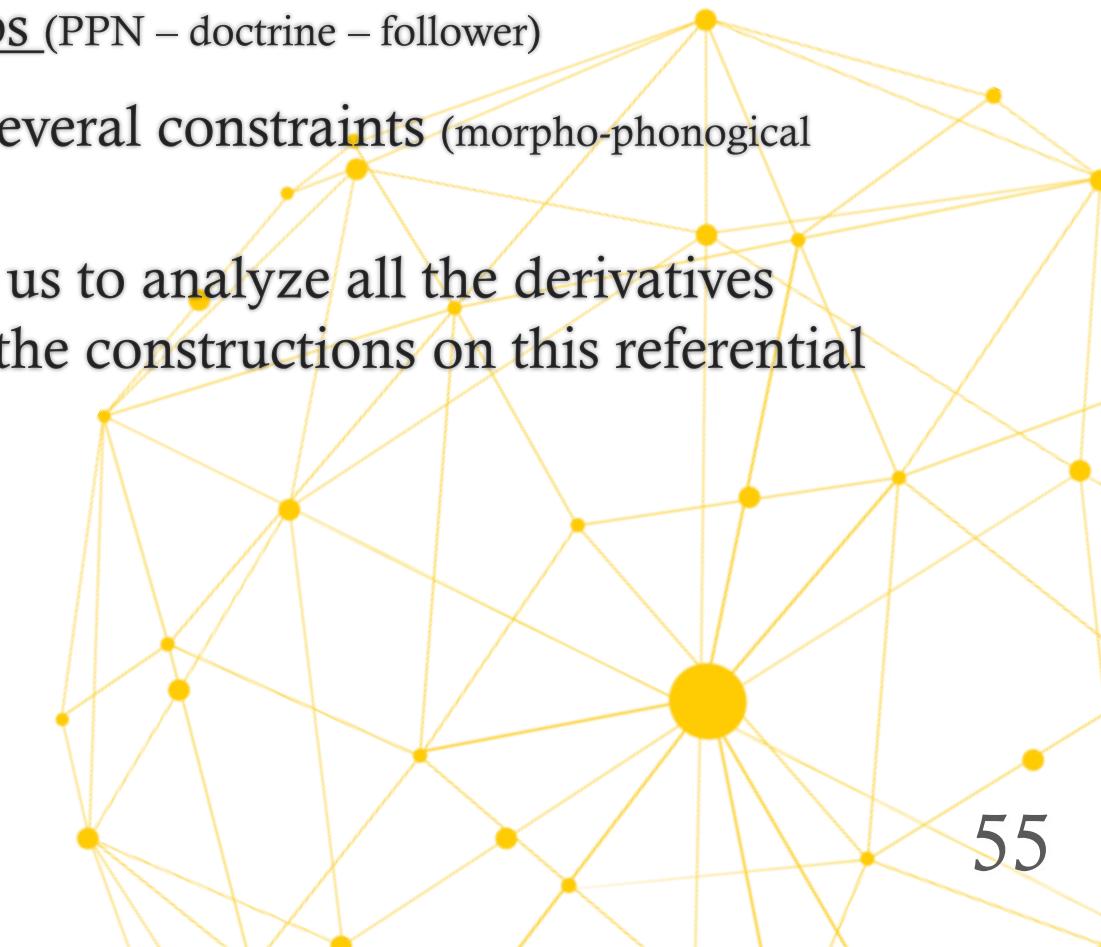
- Such a proper name subclass is therefore justified not only from a referential point of view but also from a linguistic point of view

PPN						
(F)	First Name	Last Name	First Name Last Name	Last Name <sub>1</sub>	Last Name <sub>2</sub>	Acronym
(Sx)				NPr		
(S)	'denominative meaning'			'set of stereotypic properties'		

# Conclusion

## Derivational families have a network organisation

- The formal, syntactic and semantic networks are not superposable
- The networks include paradigmatic relationships (PPN – doctrine – follower)
- The realization of the networks was driven by several constraints (morpho-phonological and extra-linguistic)
- Abandoning a binary-oriented approach allows us to analyze all the derivatives ( $\text{ROYALIFIER}_V > \text{ROYALIFICATION}_{NC}$ ) and helps us predict the constructions on this referential subclass of proper names



Thank you !

# References

- Bauer L. (1997). Derivational paradigms. In: G. Booij & J. van Marle (eds), *Yearbook of Morphology 1996*. Dordrecht: Kluwer, 243-256.
- Booij G. (2008). Paradigmatic morphology. In: B. Fradin (éd), *La raison morphologique. Hommage à la mémoire de Danielle Corbin*. Amsterdam/Philadelphia: John Benjamins, 29-37.
- Bonami O. & Boyé G. (2003). Supplétion et classes flexionales dans la conjugaison du français. *Langages* 152, 102-126.
- Bonami O., Boyé G. & Kerleroux F. (2009) L'allomorphie radicale et la relation flexion-construction. In : B. Fradin, F. Kerleroux & M. Plénat (eds), *Aperçus de morphologie du français*. Saint-Denis : Presses Universitaires de Vincennes, 103-125
- Hathout N., Montermini F. & Tanguy L. (2008). Extensive data for morphology: using the World Wide Web. *Journal of French Language Studies* 18, 67-85.
- Huguin M. (2015). *Lexèmes morphologiquement construits sur base de noms propres de personnalités politiques*. Mémoire de Master non publié, Université de Lorraine, France.
- Langendonck W.V. (2007). *Theory and Typology of Proper Names*. Berlin/New York: Mouton de Gruyter.
- Lasserre M. & Montermini F. (2014). Une approche distributionnelle pour l'analyse des dérivés néoclassiques. *Verbum (Presses Universitaires de Nancy)*.

# References

- Lignon S. & Plénat M. (2009). Échangisme suffixal et contraintes phonologiques. In : B. Fradin, F. Montermini & M. Plénat (eds), *Aperçu de morphologie du français*. Saint Denis : Presses Universitaires de Vincennes, 65-81.
- Plénat M. (1997). Analyse morpho-phonologique d'un corpus d'adjectif en *-esque*. *Journal of French Language Studies* 7, 163-179.
- Plénat, M. (2011). Enquête sur divers effets des contraintes dissimilatives en français. In: M. Roché, G. Boyé, N. Hathout, S. Lignon & M. Plénat (eds), *Des unités morphologiques au lexique*, chapitre IV. Paris : Hermès Lavoisier, 145-190.
- Prince A. & Smolensky P. (1993). *Optimality Theory. Constraint Interaction in Generative Grammar*. Piscataway : Rutgers Center for Cognitive Sciences – Rutgers University.
- Roché M. & Plénat M. (2014). Le jeu des contraintes dans la sélection du thème présuffixal. *CMLF 2014, SHS Web of Conferences* 8, 1863-1878.
- Roché M. (2011). Quel traitement unifié pour les suffixations en *-isme* et en *-iste* ? In : M. Roché, G. Boyé, N. Hathout, S. Lignon & M. Plénat (eds), *Des unités morphologiques au lexique*, Chapitre III. Paris : Hermès Lavoisier, 69-139.
- Štekauer P. (2014). Derivational Paradigms. In: R. Lieber & P. Štekauer (eds), *The Oxford Handbook of Derivational Morphology*. Oxford: Oxford University Press, 354-369.
- Villoing F. (2012). Contraintes de taille dans les mots composés : quand la phonologie entre en concurrence avec les contraintes morphologiques. *CMLF 2012, SHS Web of Conferences* 1, 1425-1440.