

Colloque Cooccurrence 2012
09/02/2012 - Besançon

Longueur de branches et arbres de mots

Philippe Gambette, Nuria Gala, Alexis Nasr, Alain Guénoche

LIGM

Université Paris-Est
Marne-la-Vallée

LIF

Université Aix-Marseille

IML

CNRS



Plan

- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- Résultats
- Visualisations
- Perspectives

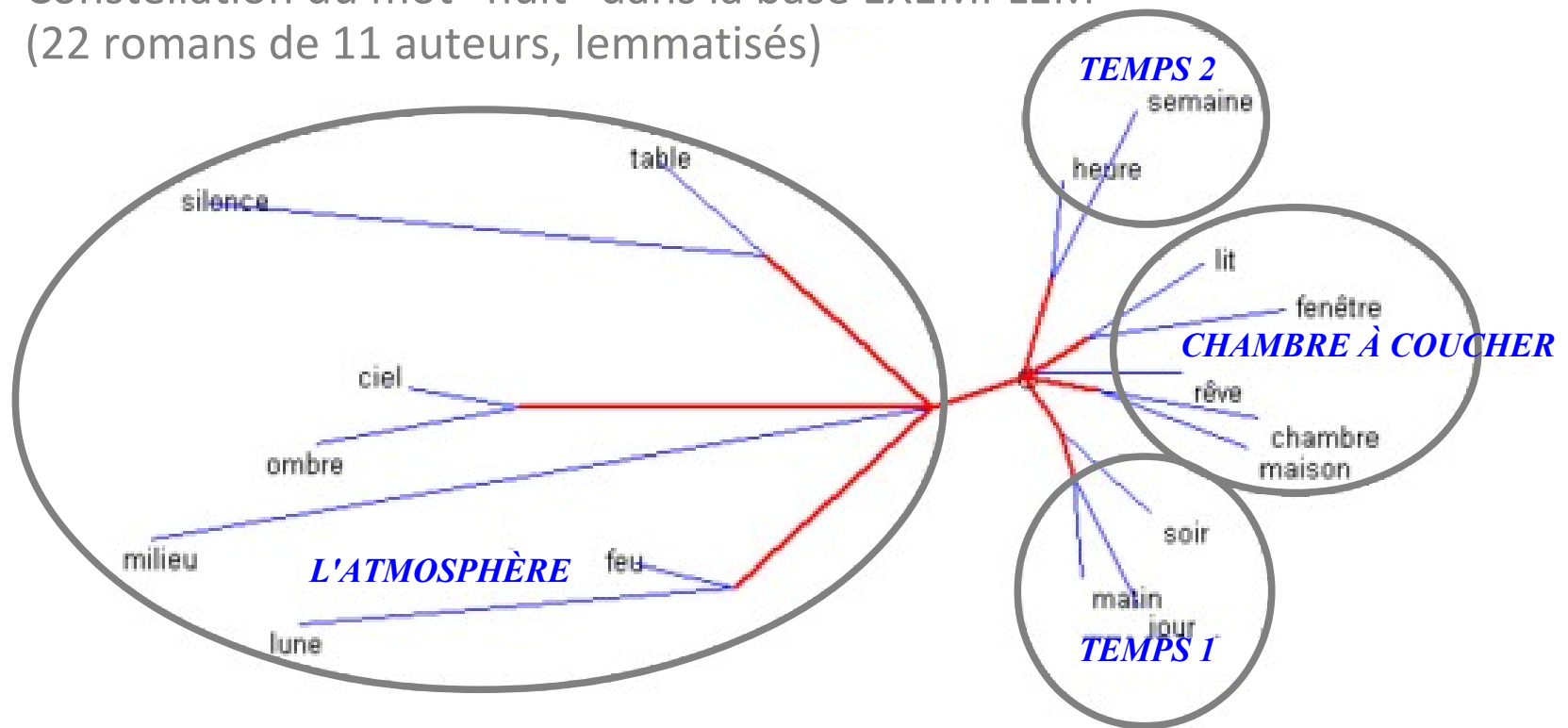
Plan

- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- Résultats
- Visualisations
- Perspectives

Analyses arborées

Rapprochement des mots d'un texte selon leur degré de cooccurrence dans le texte

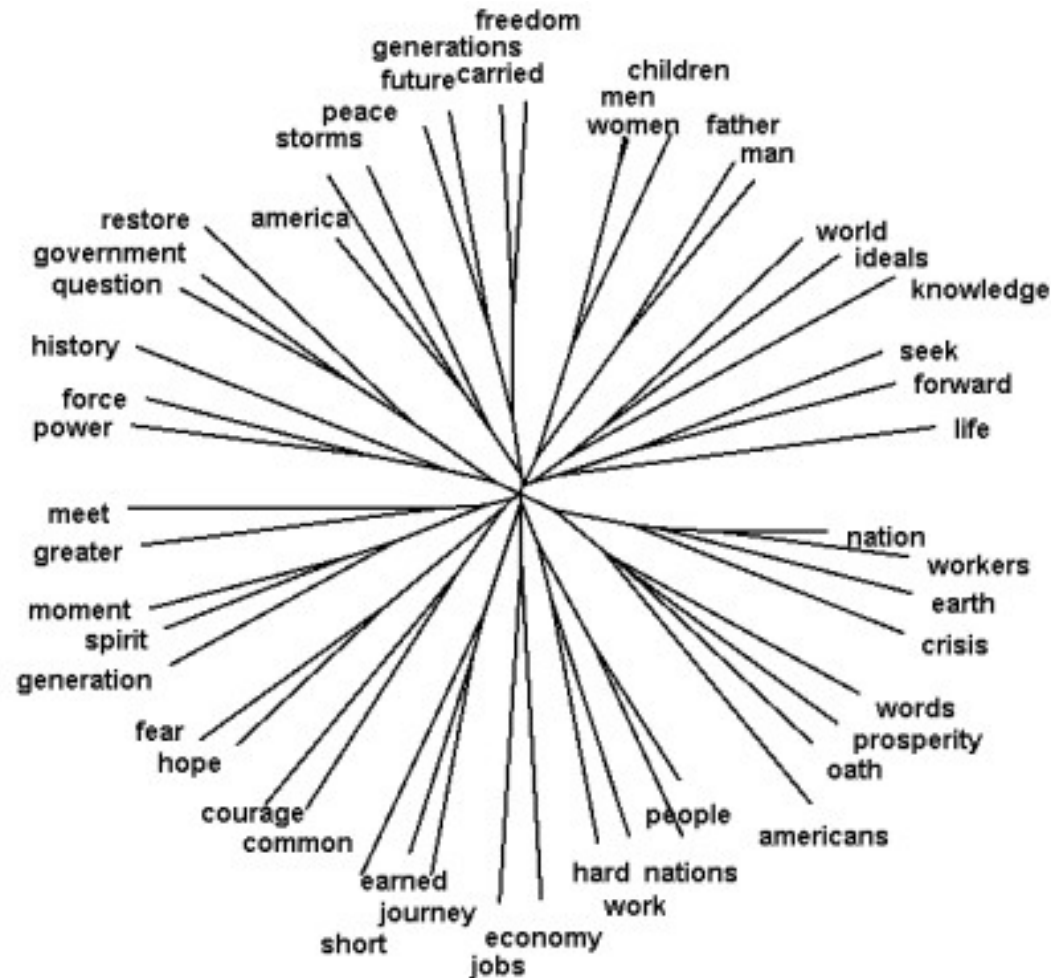
Constellation du mot "nuit" dans la base EXEMPLEM
(22 romans de 11 auteurs, lemmatisés)



Plan

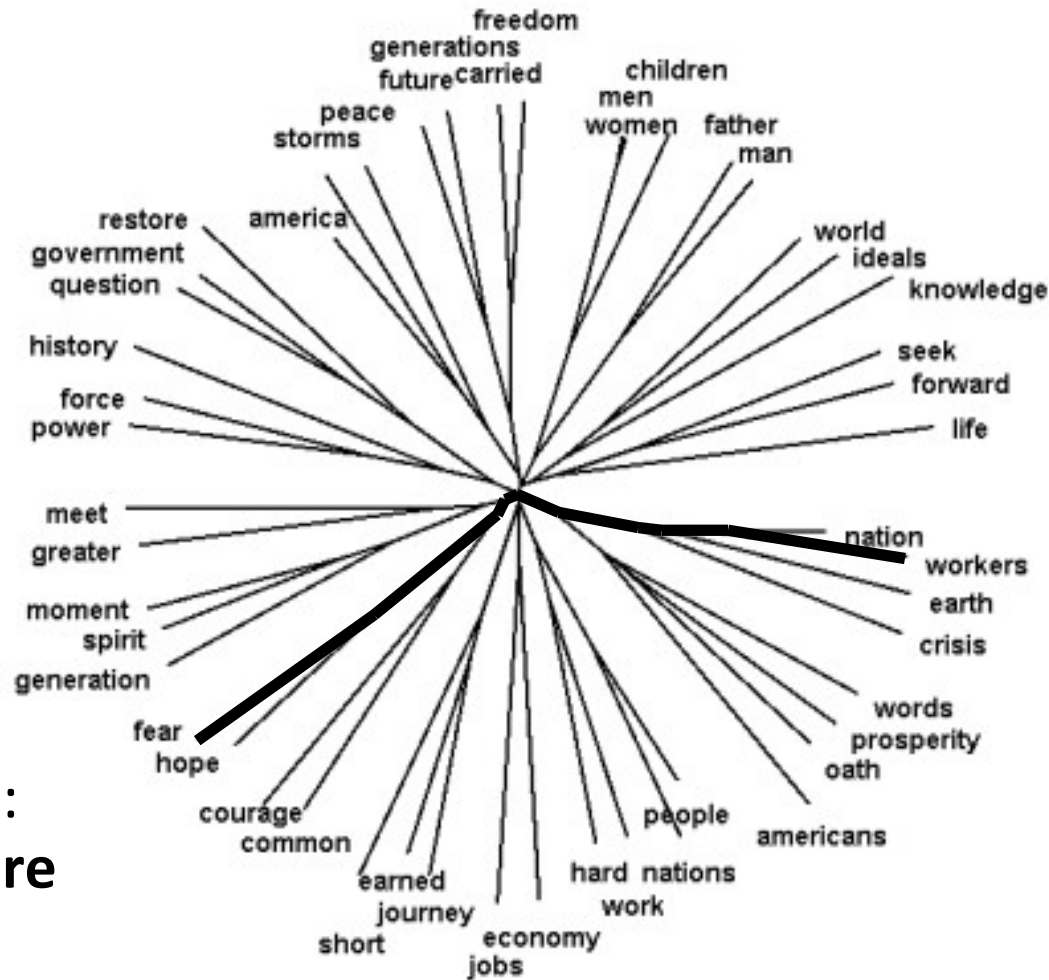
- Analyses et nuages arborés
- **Interprétation visuelle**
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- Résultats
- Visualisations
- Perspectives

Interprétation réelle

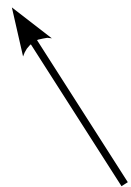


Les **distances dans l'arbre** entre deux mots reflètent *au mieux* le **degré de cooccurrence** entre ces deux mots

Interprétation réelle

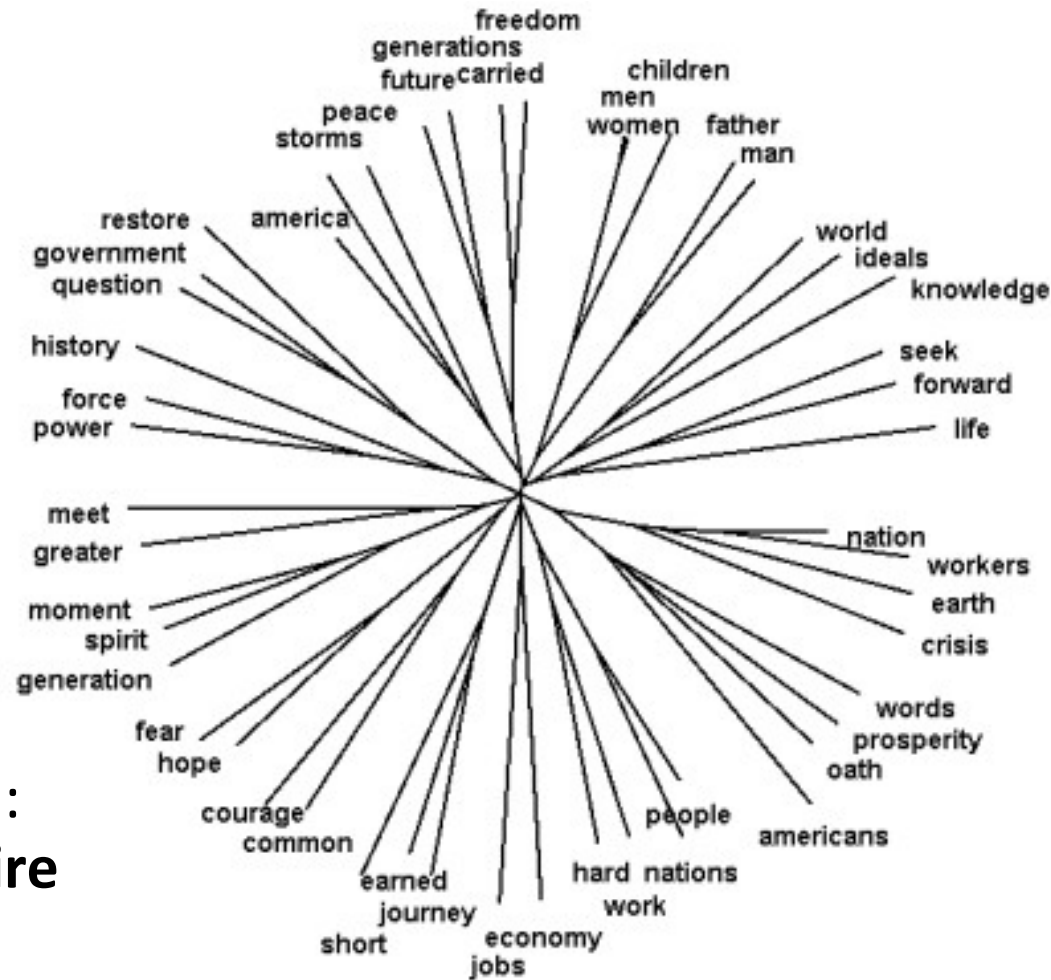


Problème 1 :
difficiles à lire



Les **distances** dans l'arbre entre deux mots reflètent *au mieux* le **degré de cooccurrence** entre ces deux mots

Interprétation réelle

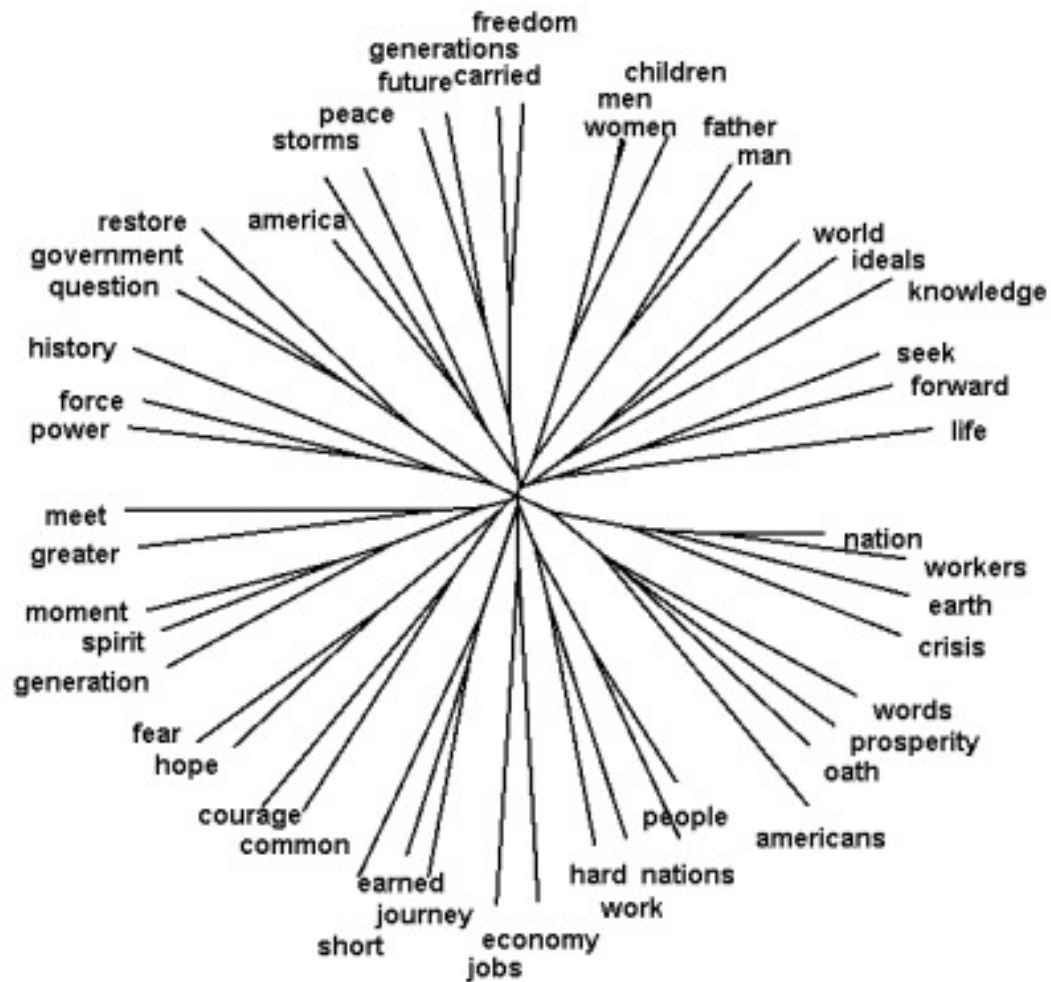


Problème 1 :
difficiles à lire

Problème 2 :
peu fiables

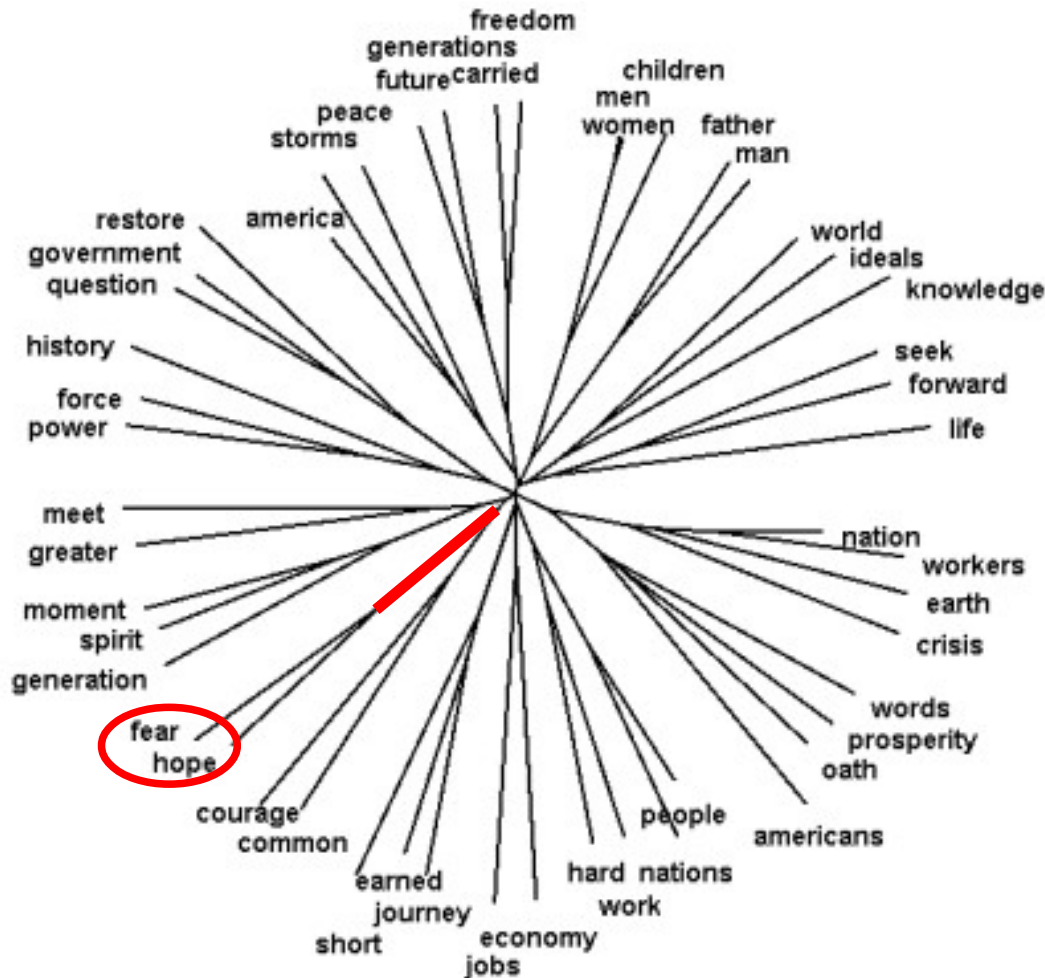
Les **distances dans l'arbre** entre deux mots reflètent *au mieux* le **degré de cooccurrence** entre ces deux mots

Interprétation pratique



arbre de distances
utilisé comme
classification

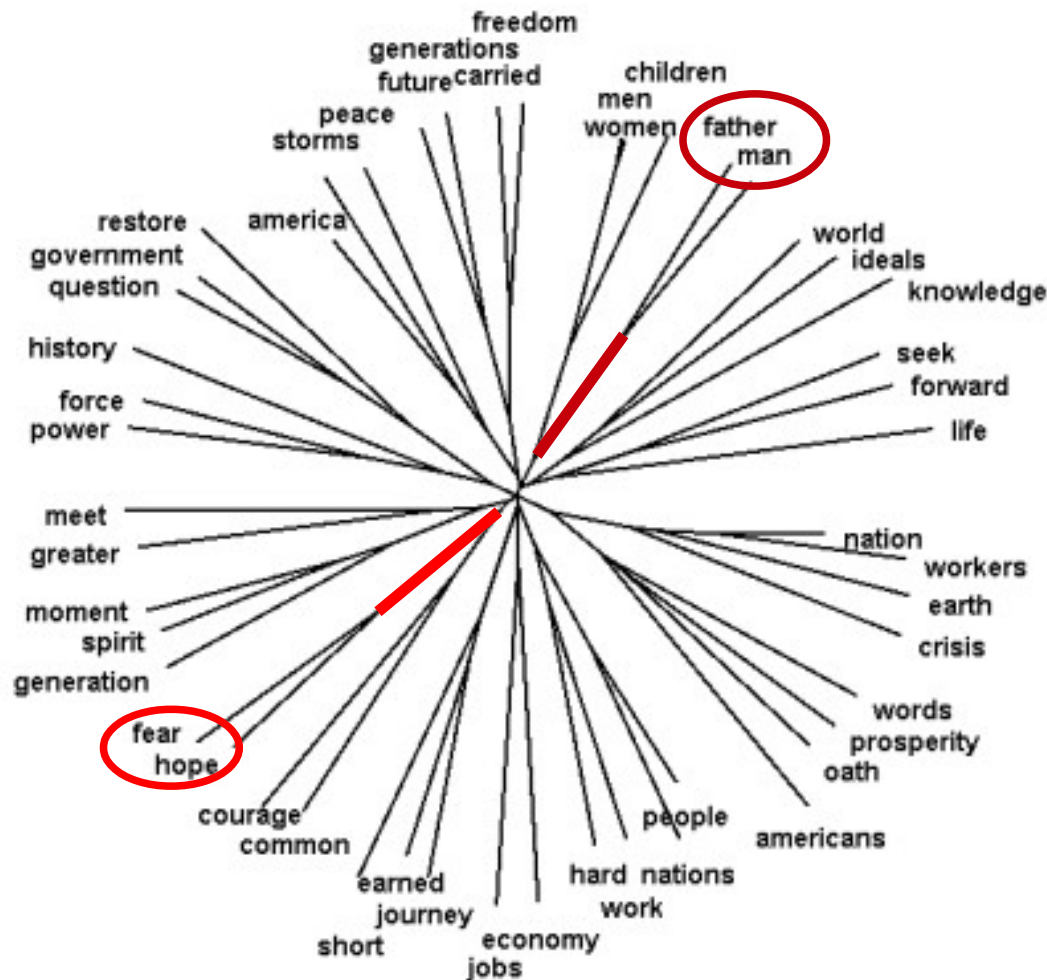
Interprétation pratique



arbre de distances
utilisé comme
classification

Les mots d'un **même sous-arbre** bien séparé du reste de l'arbre
constituent une classe de mots

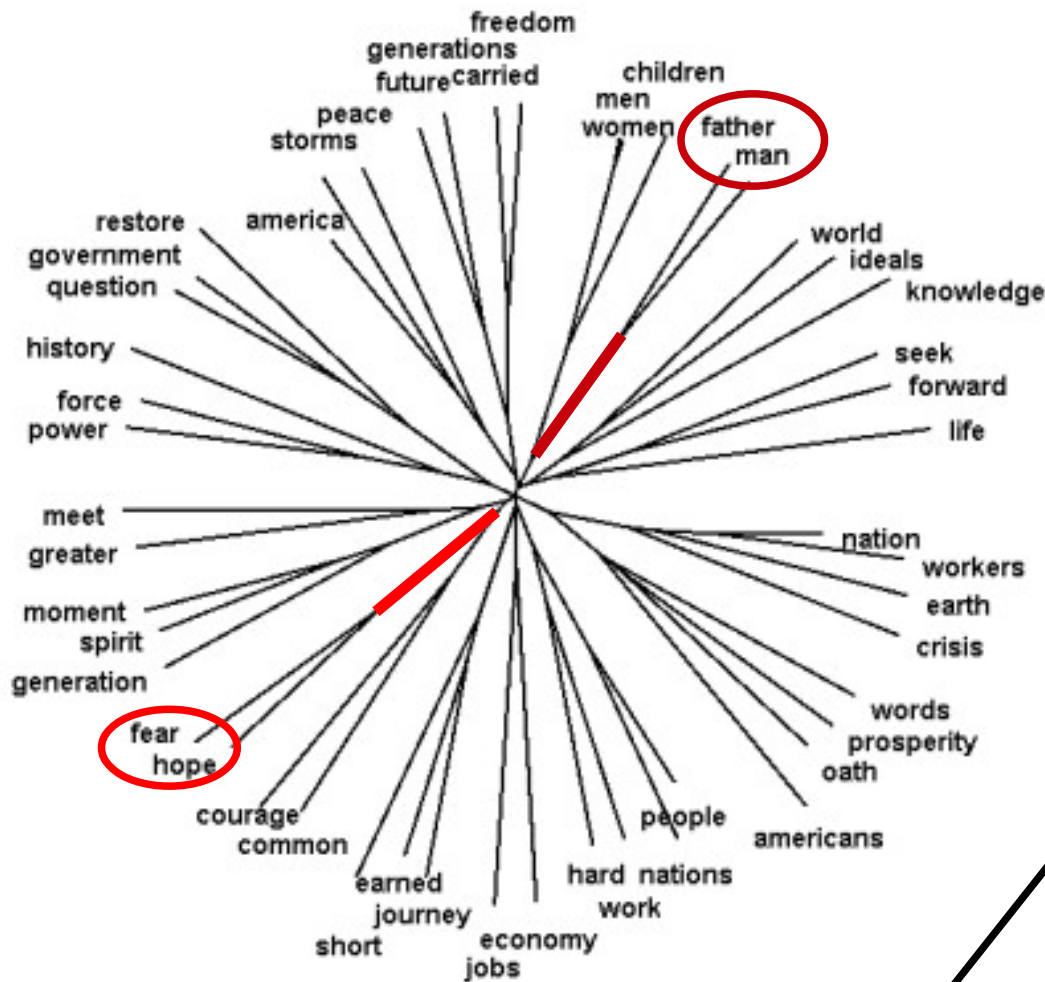
Interprétation pratique



arbre de distances
utilisé comme
classification

Les mots d'un **même sous-arbre** bien séparé du reste de l'arbre
constituent une classe de mots

Interprétation pratique



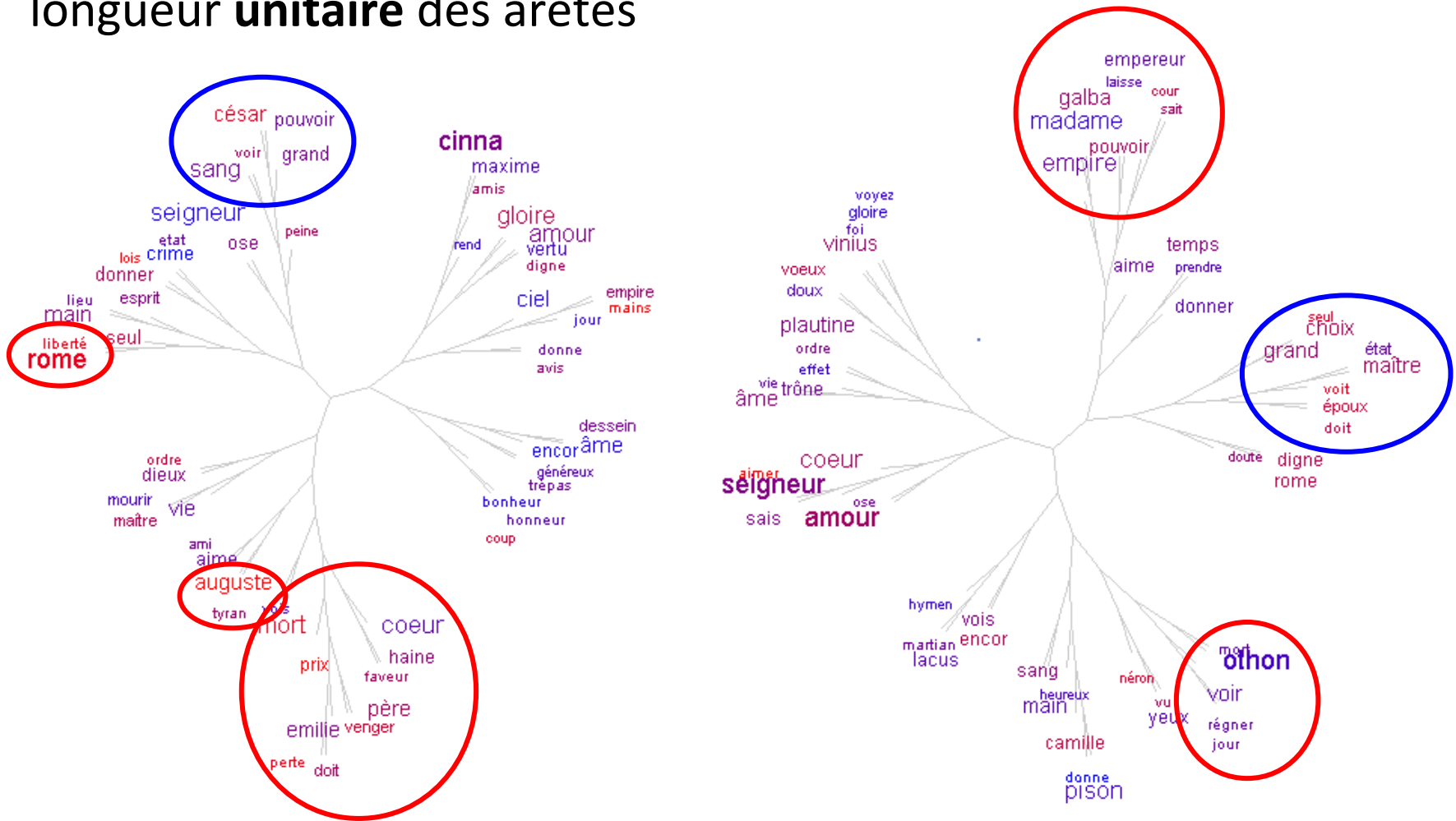
arbre de distances
utilisé comme
classification

Problème : **toujours
peu lisible** (longueur
des arêtes externes)
et **peu fiable**

Les mots d'un **même sous-arbre** bien séparé du reste de l'arbre
constituent une classe de mots

Interprétation pratique

Astuce de visualisation pour améliorer la lisibilité :
longueur **unitaire** des arêtes

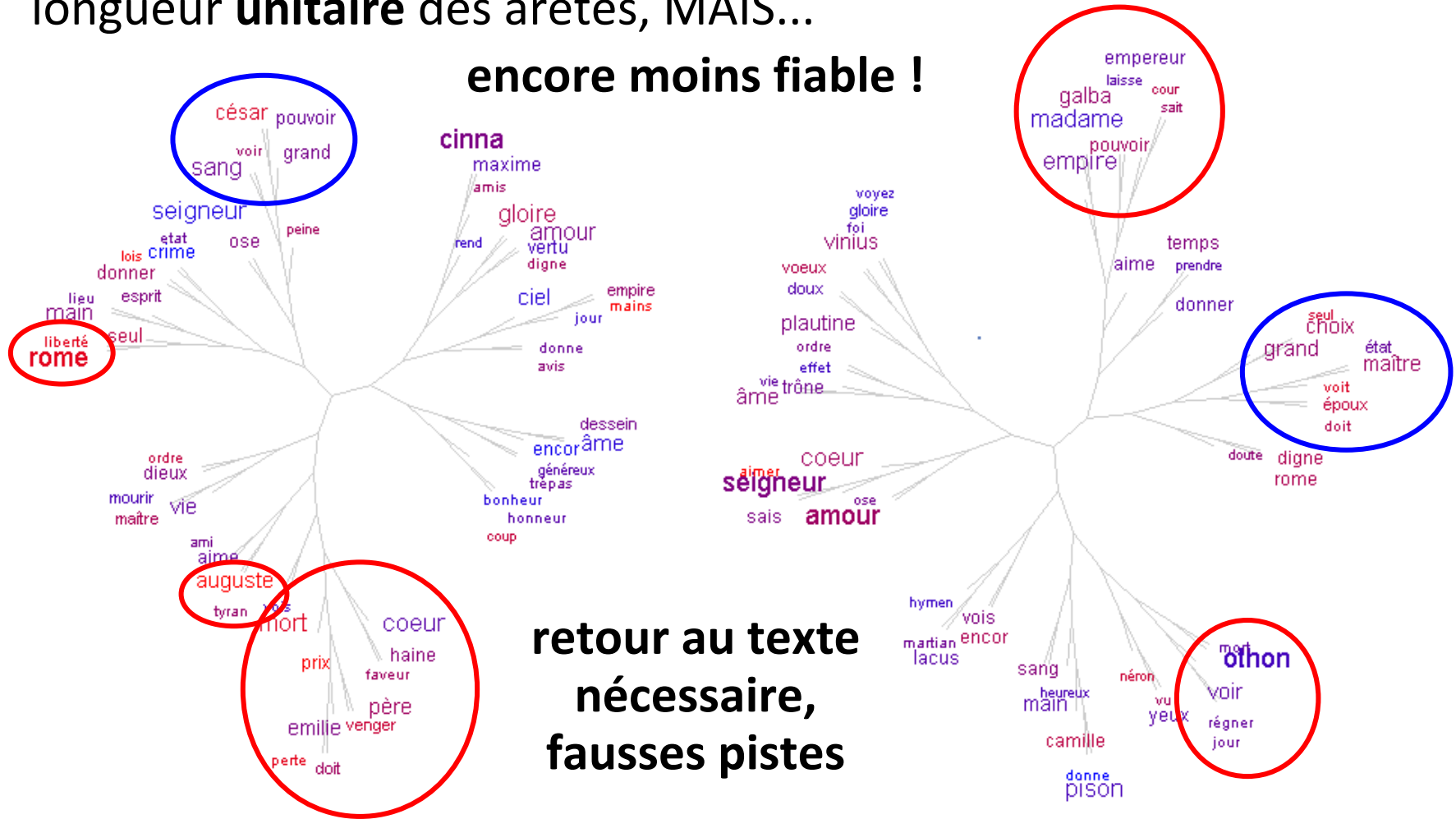


Nuages arborés globaux des 60 mots les plus fréquents dans Cinna et Othon (distance Liddell, fenêtre de largeur 20), colorés chronologiquement (rouge au début, bleu à la fin)

Interprétation pratique

Astuce de visualisation pour améliorer la lisibilité :
longueur **unitaire** des arêtes, MAIS...

encore moins fiable !



**retour au texte
nécessaire,
fausses pistes**

Nuages arborés globaux des 60 mots les plus fréquents dans *Cinna* et *Othon* (distance Liddell, fenêtre de largeur 20), colorés chronologiquement (rouge au début, bleu à la fin)

Interprétation pratique

Problème :

Comment calculer les longueurs des arêtes de l'arbre pour une **interprétation fiable des classes** ?

Arête longue = classe de mots significative (proches les uns des autres, bien séparés du reste)

Arête courte = classe de mots peu significative

Plan

- Analyses et nuages arborés
- Interprétation visuelle
- **Formules de longueurs d'arêtes**
- Protocole d'évaluation
- Résultats
- Visualisations
- Perspectives

Formules de longueurs d'arêtes

Post-calcul des longueurs d'arêtes après la construction de l'arbre,
pour que :

arêtes les plus longues \leftrightarrow classes de mots **les plus significatives**
 \leftrightarrow classes de mots **bien séparées**
d'après la **distance de cooccurrence**

Formules de longueurs d'arêtes

Post-calcul des longueurs d'arêtes après la construction de l'arbre, pour que :

arêtes les plus longues \leftrightarrow **classes de mots les plus significatives**
 \leftrightarrow classes de mots **bien séparées**
d'après la **distance de cooccurrence**

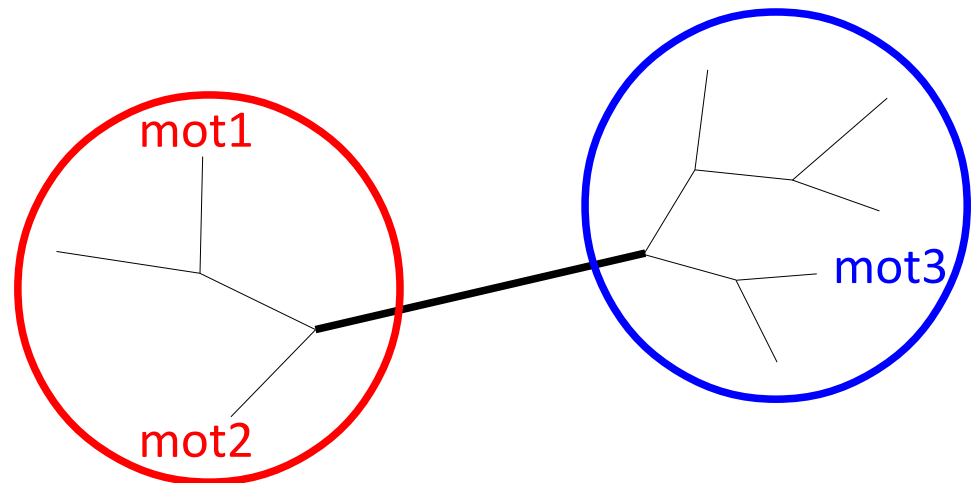
Formule du **ratio des bons triplets** (“triples”) :

Si **mot1** et **mot2** d'un côté de l'arête, **mot3** de l'autre côté,

“**bon triplet**” si

$\text{distance}(\text{mot1}, \text{mot2}) <$
 $\text{min}(\text{distance}(\text{mot1}, \text{mot3}),$
 $\text{distance}(\text{mot2}, \text{mot3}))$

ratio espéré proche de 1



Formules de longueurs d'arêtes

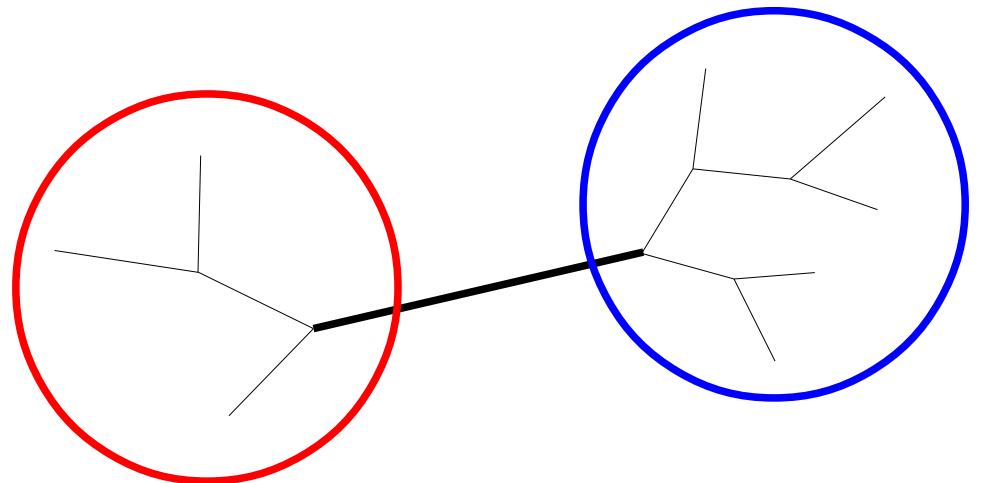
Post-calcul des longueurs d'arêtes après la construction de l'arbre, pour que :

arêtes les plus longues ↔ **classes de mots les plus significatives**
↔ **classes de mots bien séparées**
d'après la **distance de cooccurrence**

Formule du **ratio des distances moyennes** (“distanceRatio”) :

$$\frac{\text{moyenne}(\text{distances } \text{inter-classes})}{\text{moyenne}(\text{distances } \text{intra-classes})}$$

ratio espéré supérieur à 1



Formules de longueurs d'arêtes

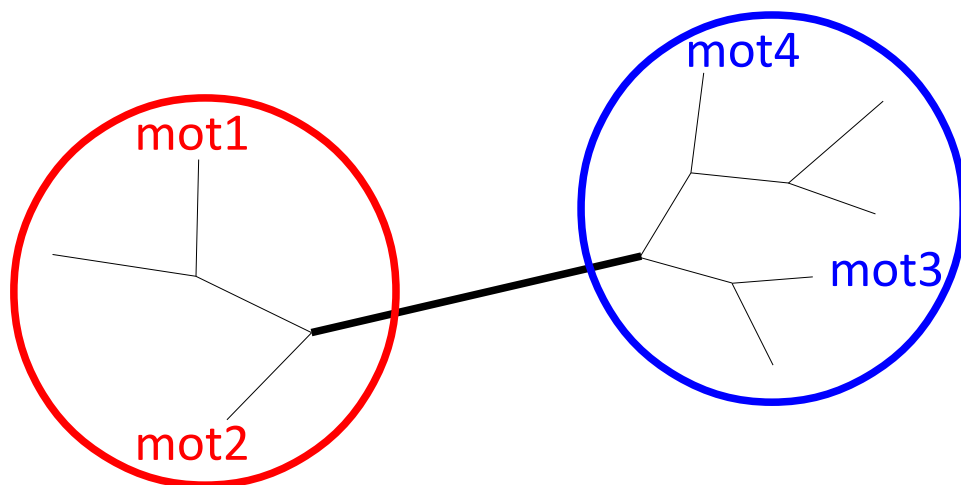
Post-calcul des longueurs d'arêtes après la construction de l'arbre, pour que :

arêtes les plus longues \leftrightarrow classes de mots les plus significatives
 \leftrightarrow classes de mots bien séparées
d'après la distance de cooccurrence

Formule du ratio des bons quadruplets (“quartets”) :

Si **mot1** et **mot2** d'un côté de l'arête, **mot3** et **mot4** de l'autre côté,

“bon quadruplet” si
 $\text{distance}(\text{mot1}, \text{mot2}) +$
 $\text{distance}(\text{mot2}, \text{mot3}) <$
 $\min(\text{distance}(\text{mot1}, \text{mot3}) +$
 $\text{distance}(\text{mot2}, \text{mot4}) +,$
 $\text{distance}(\text{mot1}, \text{mot4}) +$
 $\text{distance}(\text{mot2}, \text{mot3}))$



ratio espéré proche de 1

Plan

- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- **Protocole d'évaluation**
- Résultats
- Visualisations
- Perspectives

Protocole d'évaluation

Post-calcul des longueurs d'arêtes après la construction de l'arbre, puis :

Vérification que les classes de mots **les mieux séparées**
(d'après ces longueurs) sont **significatives**

Protocole d'évaluation

Post-calcul des longueurs d'arêtes après la construction de l'arbre, puis :

Vérification que les classes de mots
les mieux séparées (d'après ces
longueurs) sont **significatives**



Partition obtenue en **découpant les
arêtes les plus longues** comparée
avec une partition de référence

Protocole d'évaluation

Post-calcul des longueurs d'arêtes après la construction de l'arbre, puis :

Vérification que les classes de mots
les mieux séparées (d'après ces
longueurs) sont **significatives**



Partition obtenue en **découpant les**
arêtes les plus longues comparée
avec une partition de référence,

quelles
données ?

Protocole d'évaluation

Base de données Polymots

Base lexicale de familles morpho-phonologiques

20 000 mots, 2000 familles

The screenshot displays the Polymots website interface. At the top, the word 'POLYMOTS' is written in a stylized font. Below it, there are four search buttons: 'Recherche', 'Recherche alphabétique', 'Recherche par sens', and 'Recherche par type'. A 'Recherche simple' section follows, with a text input field containing 'art' and a 'Lancer' button. Below the search bar, a 'Résultats' section shows a list of 15 words: art, art, artifice, artificiel, artificiellement, artificier, artillerie, artilleur, artisan (highlighted), artisanal, artisanalement, artisanat, artiste, artistique, and artistiquement. To the right, a 'Fiche détaillée de "artisan"' section provides details: 'Mot base : art', 'Type : transparent', 'Nombre de mots dérivés contenant le mot base : 14', and 'Productivité du mot base : 0.70 %'. It also includes a 'Sens' list (aider, art, artisanat, automatiser, compter, créateur, exercer, général, manuel, métier, personne, pratique, propre) and an 'Affixes' list (an, is).

Protocole d'évaluation

Base de données Polymots

Base lexicale de familles morpho-phonologiques

20 000 mots, 2000 familles

+ **partitions sémantiques** des familles de 20 mots

(arbre, art, boule, carte, corde, dent, dict, fil, fusée, lune,
meuble, mode, onde, paille, penser, pot, presse, tenir, terre, val).

Protocole d'évaluation

Base de données Polymots

Base lexicale de familles morpho-phonologiques

20 000 mots, 2000 familles

+ **partitions sémantiques** des familles de 20 mots

(arbre, art, boule, carte, corde, dent, dict, fil, fusée, lune, meuble, mode, onde, paille, penser, pot, presse, tenir, terre, val).

Exemple pour la famille de **art** :

{ {artifice, artificiel, artificiellement, artificier},
{artillerie, artilleur},
{artisan, artisanal, artisanalement, artisanat},
{artiste, artistique, artistiquement, art} }

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?

Distance utilisée pour le calcul de la représentation arborée ?

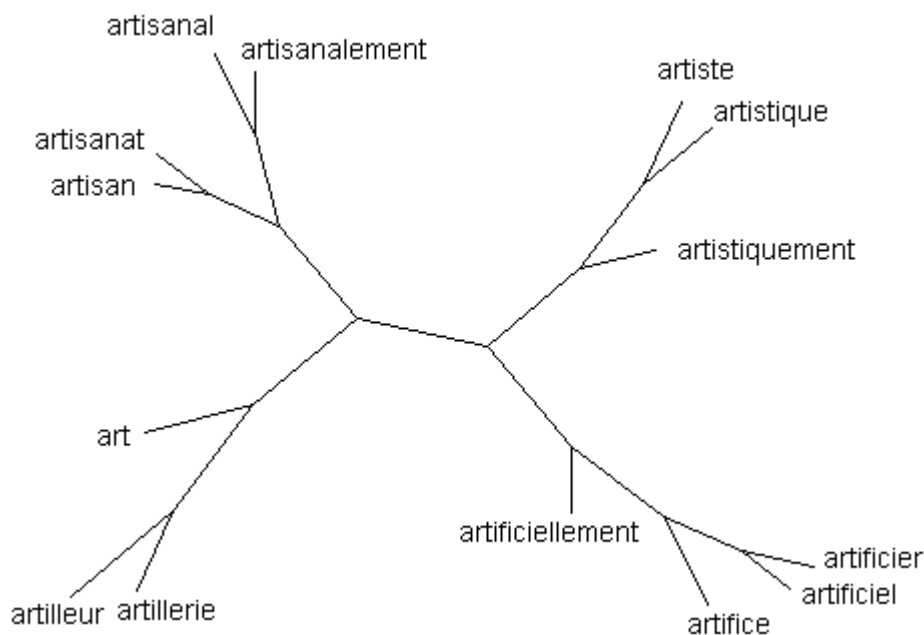
Distance composite entre :

- nombre d'affixes communs
- degré de cooccurrence dans **Le Trésor
de la Langue
Française
informatisé**

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir Pk
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

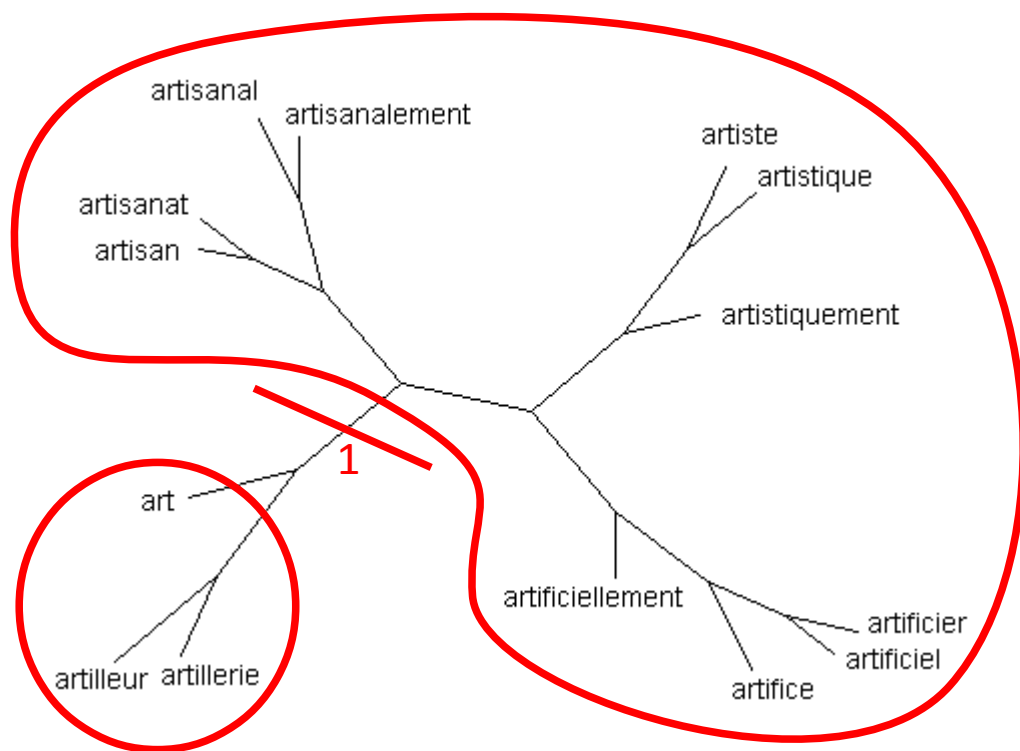
$P_0 = \{\{\text{artisan, artisanat, artisanal, artisanalelement, artiste, artistique, artistiquement, artificier, artificiel, artifice, artificiellement, artillerie, artilleur, art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalelement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

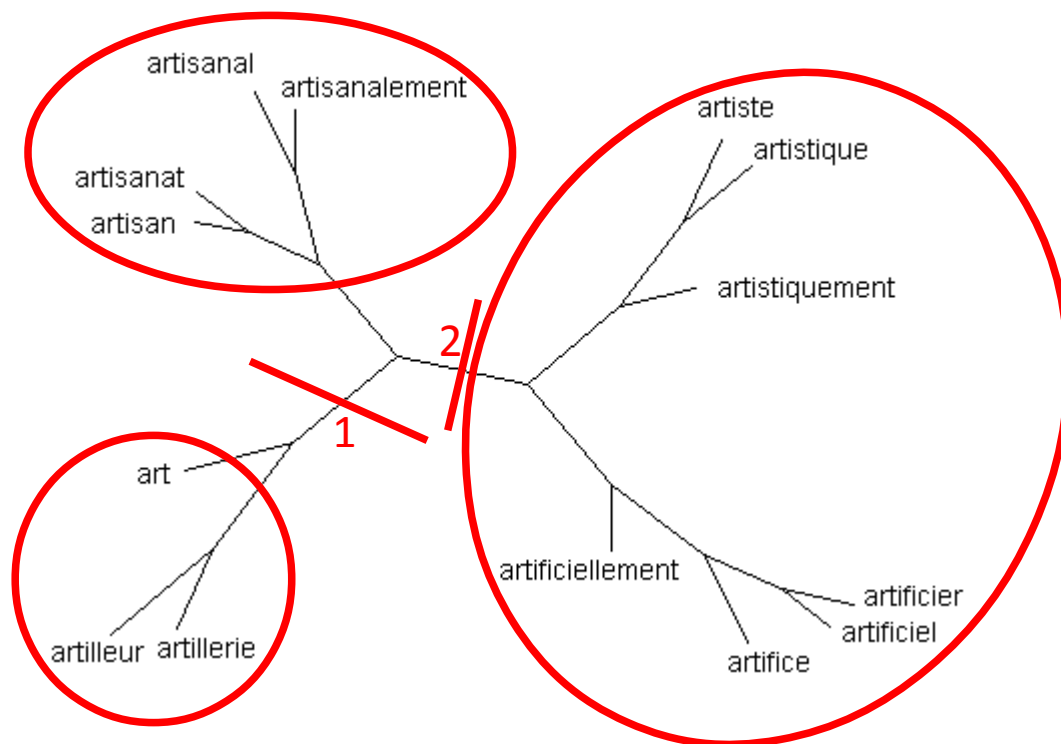
P1 = {{artisan, artisanat, artisanal, artisanalement, artiste, artistique, artistiquement, artificier, artificiel, artifice, artificiellement}, {artillerie, artilleur, art}}

Partition manuelle : **Pm** = {{artificier, artifice, artificiel, artificiellement}, {artillerie, artilleur}, {artisan, artisanal, artisanalement, artisanat}, {artiste, artistique, artistiquement, art}}

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

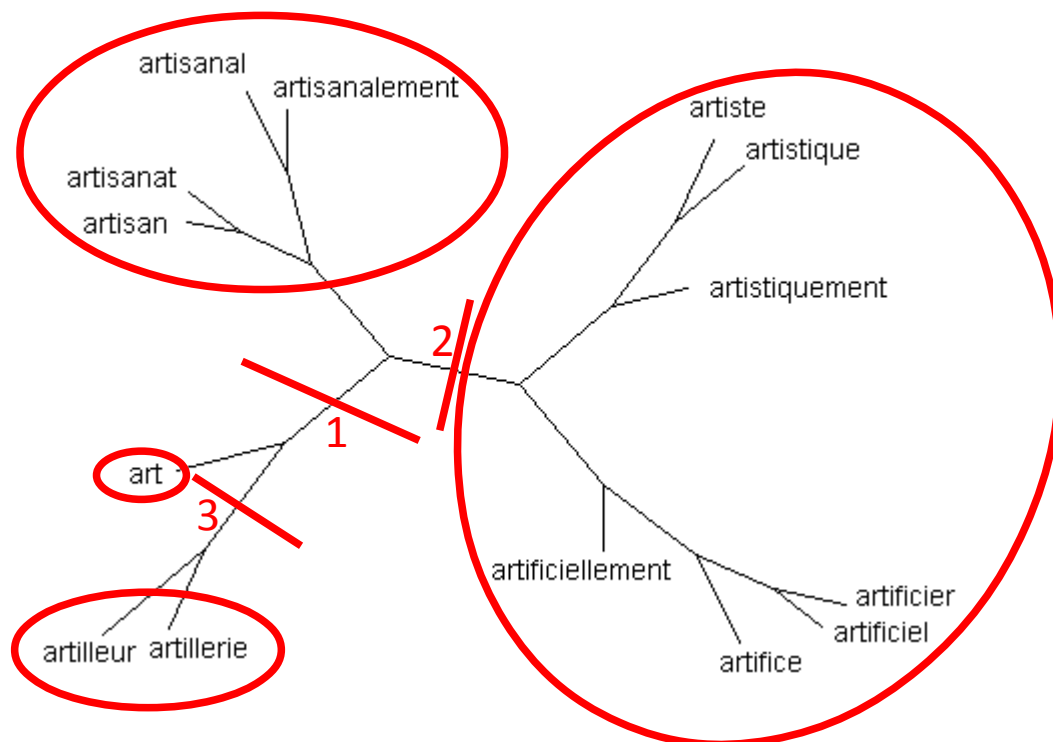
$P_2 = \{\text{artisan, artisanat, artisanal, artisanalement}\},$
 $\{\text{artiste, artistique, artistiquement, artificier, artificiel, artifice, artificiellement}\},$
 $\{\text{artillerie, artilleur, art}\}$

Partition manuelle : $P_m = \{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\},$
 $\{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

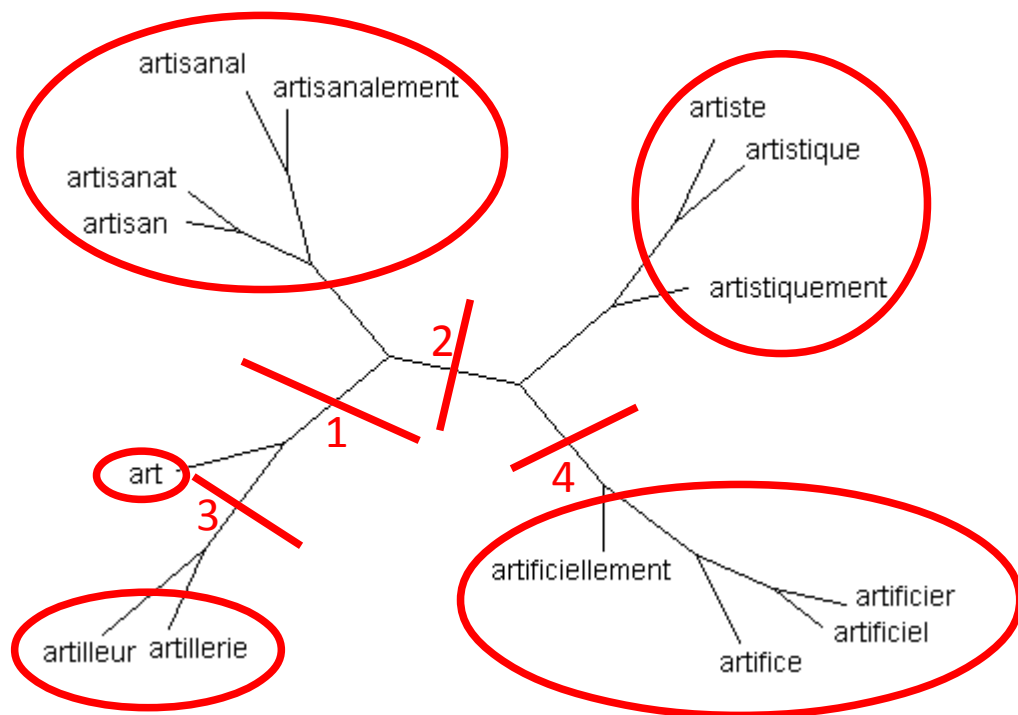
$P_3 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement, artificier, artificiel, artifice, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

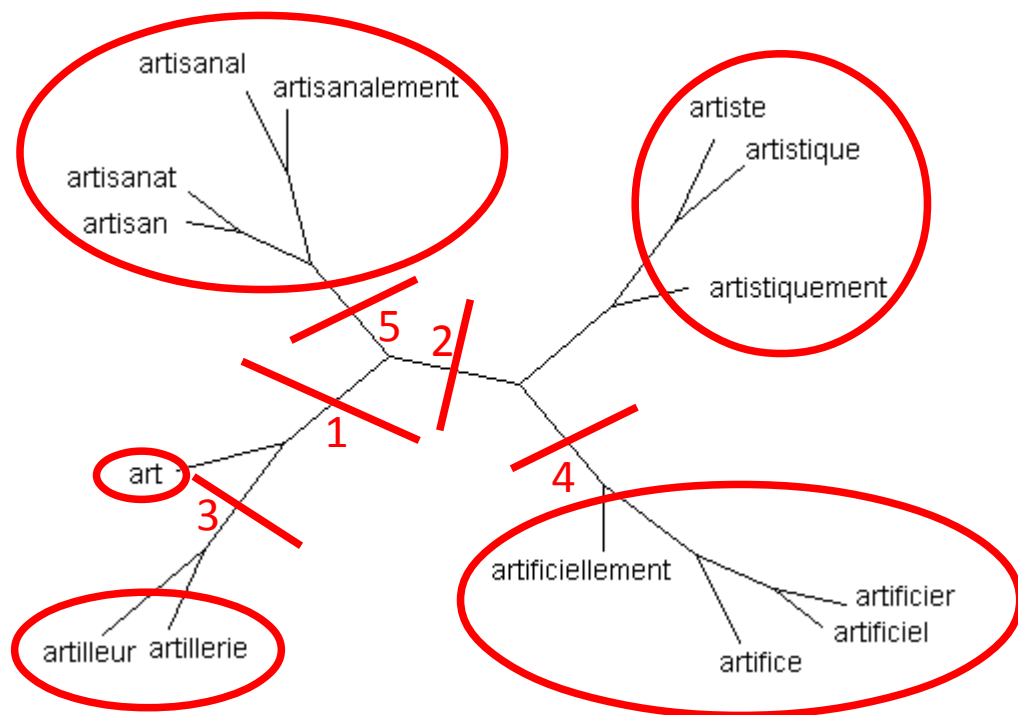
$P_4 = \{\{\text{artisan, artisanat, artisanal, artisansalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisansalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

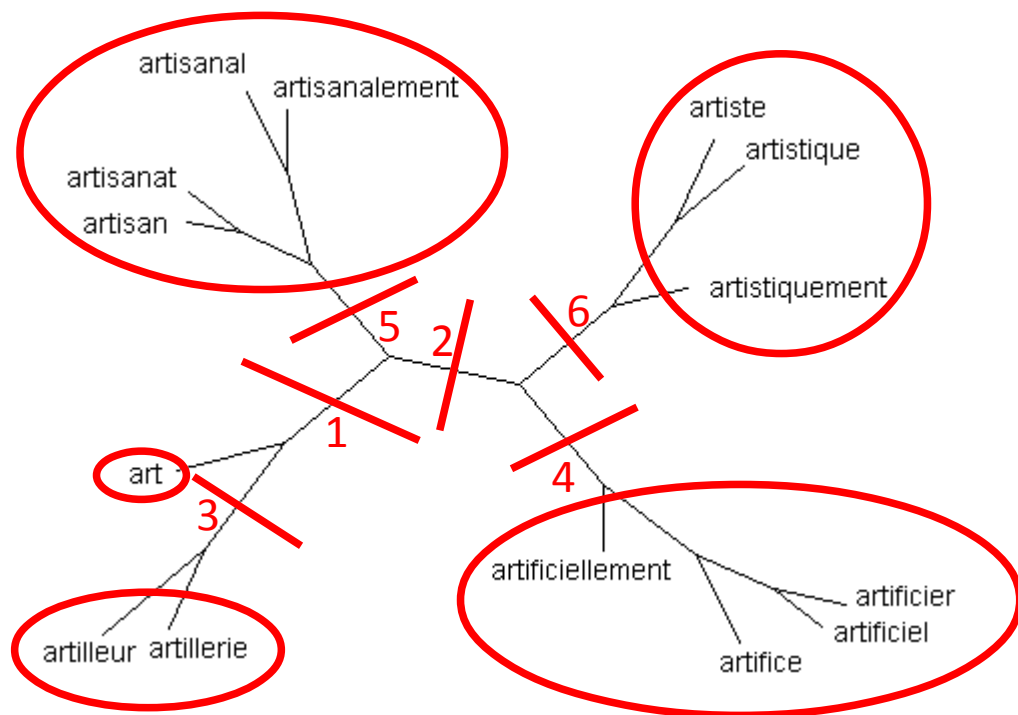
$P_5 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

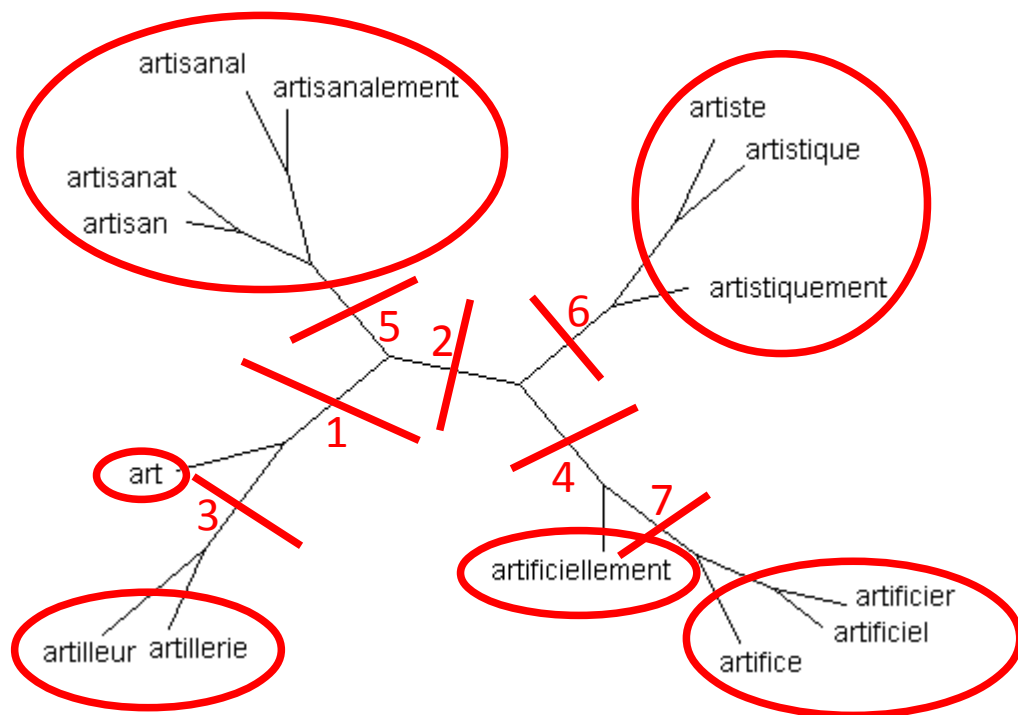
$P_6 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

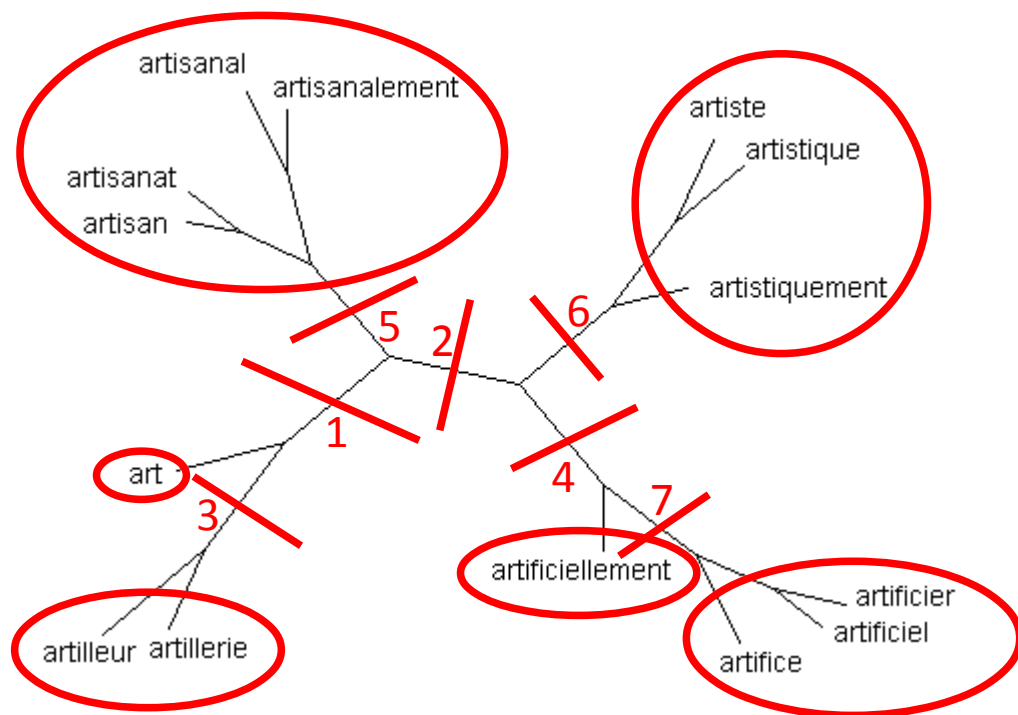
$P_7 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice}\}, \{\text{artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

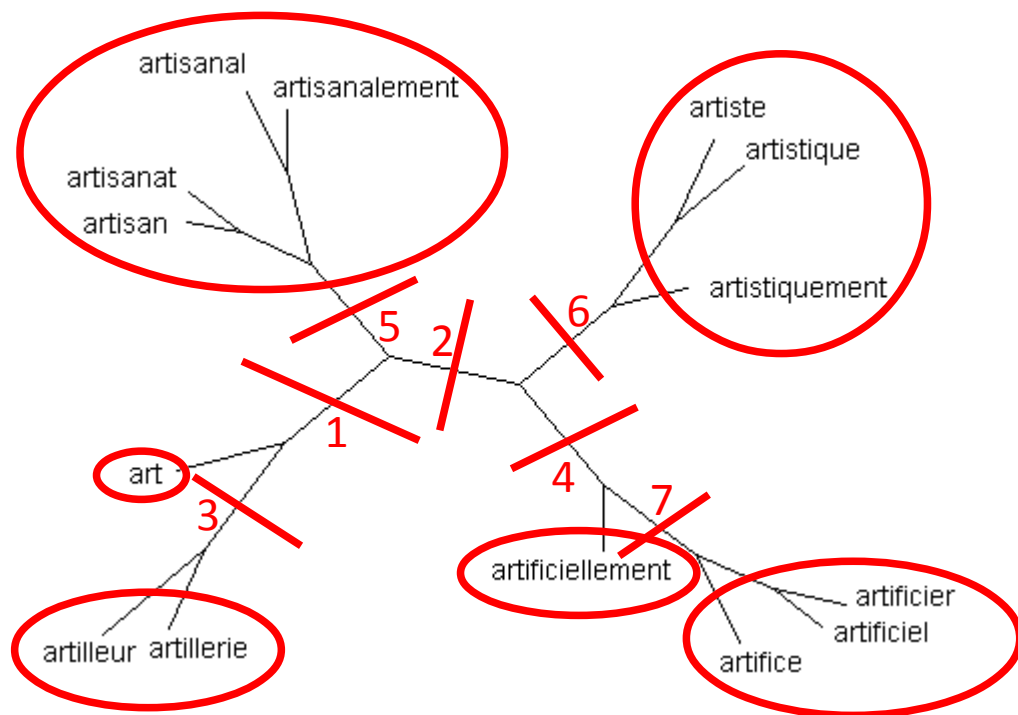
$P_7 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice}\}, \{\text{artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

$P_7 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice}\}, \{\text{artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

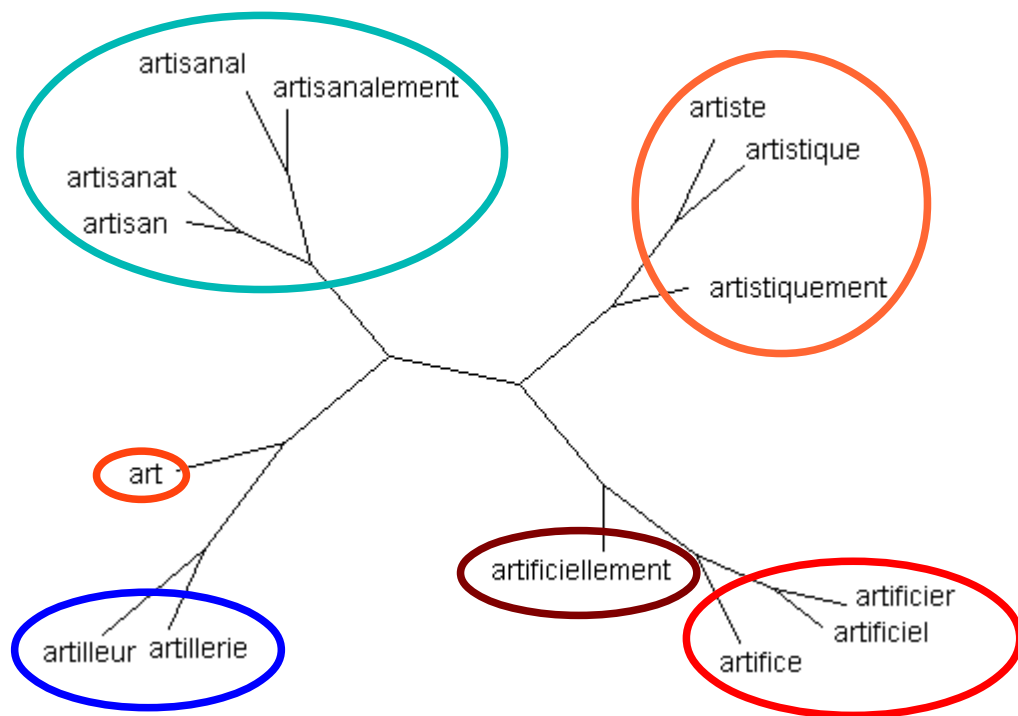
**Comparer les partitions !
(indice de Rand, Rand corrigé)**

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

$P_7 = \{\{\text{artisan, artisanat, artisanal, artisanalelement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice}\}, \{\text{artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

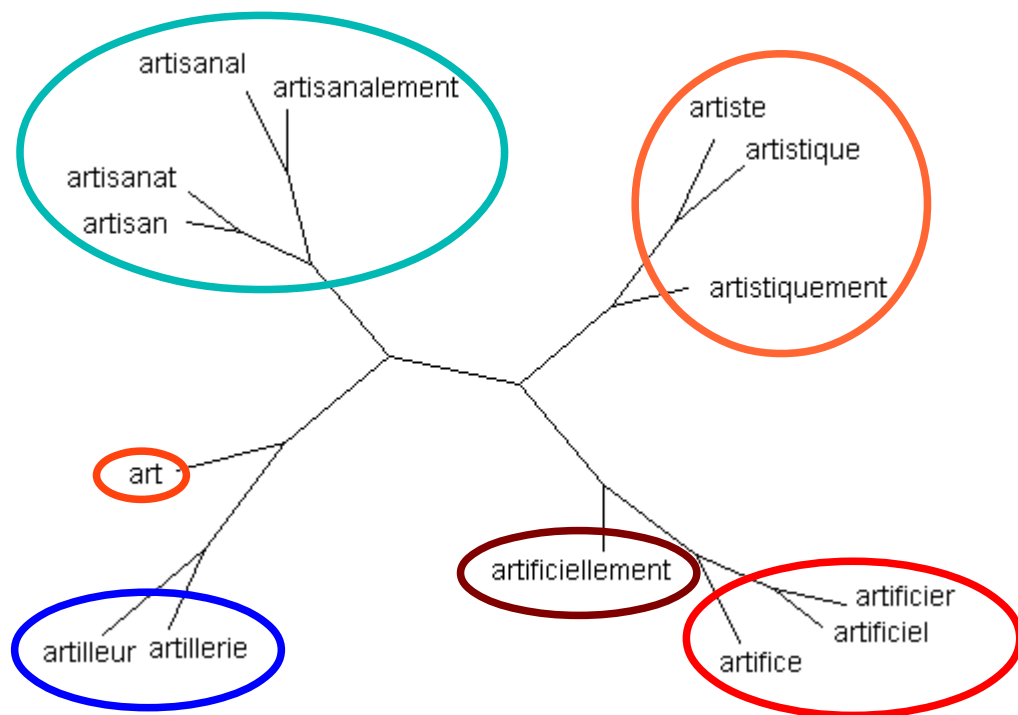
**Comparer les partitions !
(indice de Rand, Rand corrigé)**

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalelement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

$P_7 = \{\{\text{artisan, artisanat, artisanal, artisanalement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice}\}, \{\text{artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

$\text{rand}(P_m, P_7) = 0.934$

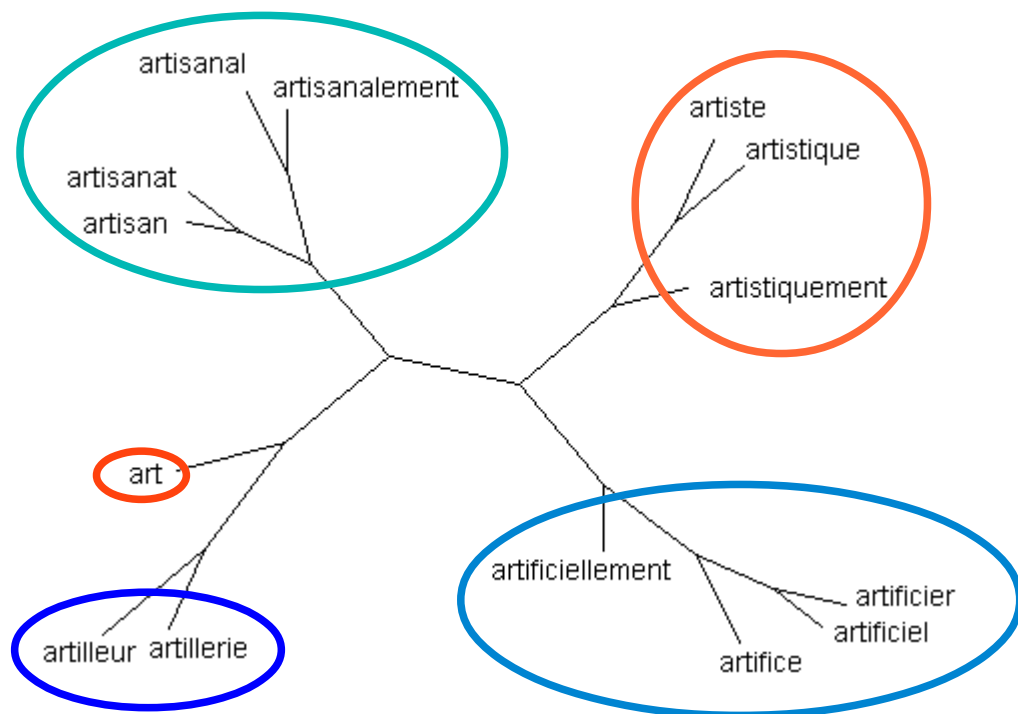
$\text{aRand}(P_m, P_7) = 0.774$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

Protocole d'évaluation

Idée :

- Construire une représentation arborée des mots de la famille
- Découper les k arêtes les plus longues de l'arbre pour obtenir P_k
- La partition obtenue est-elle proche de la partition "manuelle" ?



Partition automatique :

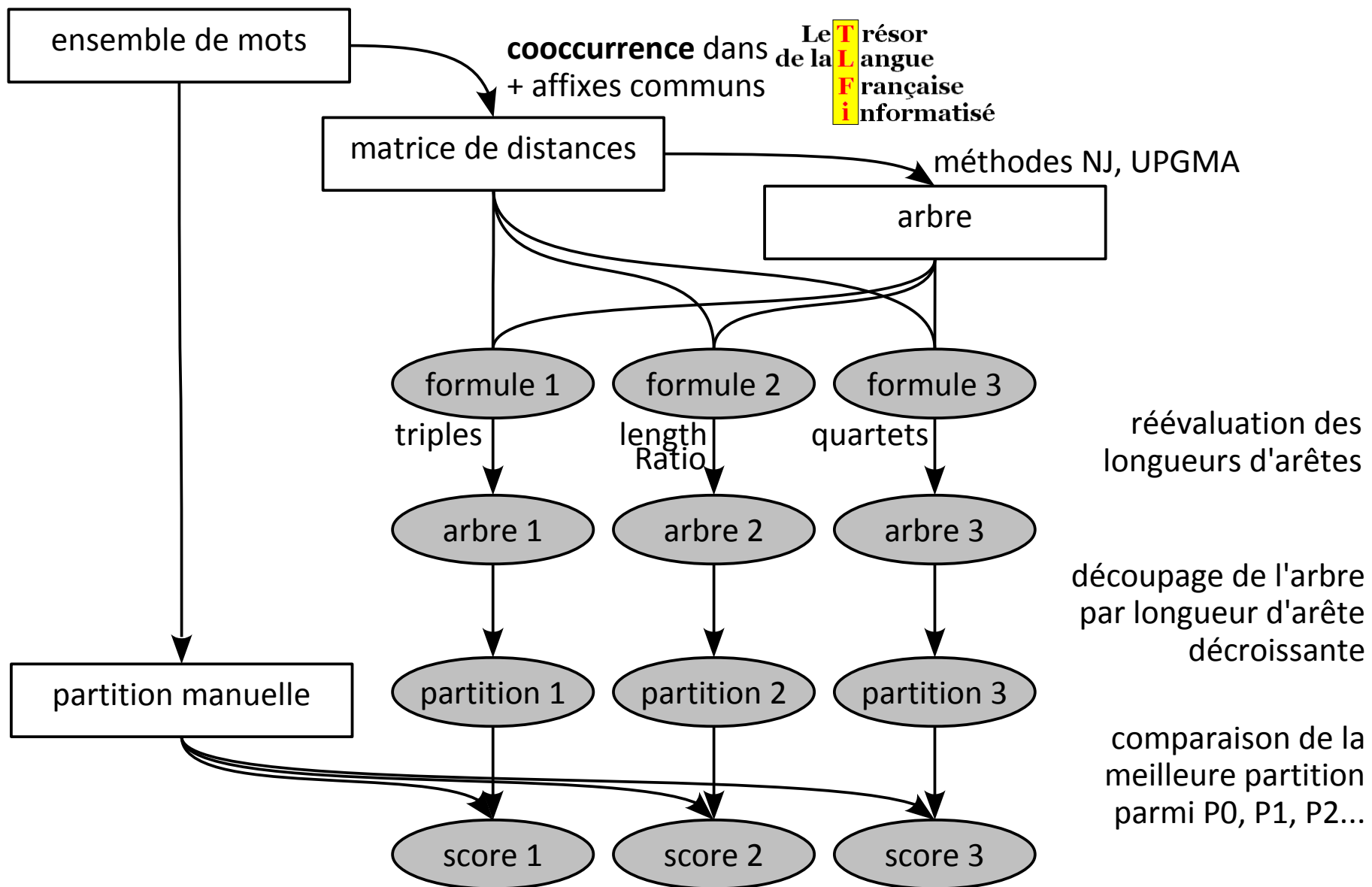
$P_4 = \{\{\text{artisan, artisanat, artisanal, artisanalelement}\}, \{\text{artiste, artistique, artistiquement}\}, \{\text{artificier, artificiel, artifice, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{art}\}\}$

$\text{rand}(P_m, P_4) = 0.967$

$\text{aRand}(P_m, P_4) = 0.894$

Partition manuelle : $P_m = \{\{\text{artificier, artifice, artificiel, artificiellement}\}, \{\text{artillerie, artilleur}\}, \{\text{artisan, artisanal, artisanalelement, artisanat}\}, \{\text{artiste, artistique, artistiquement, art}\}\}$

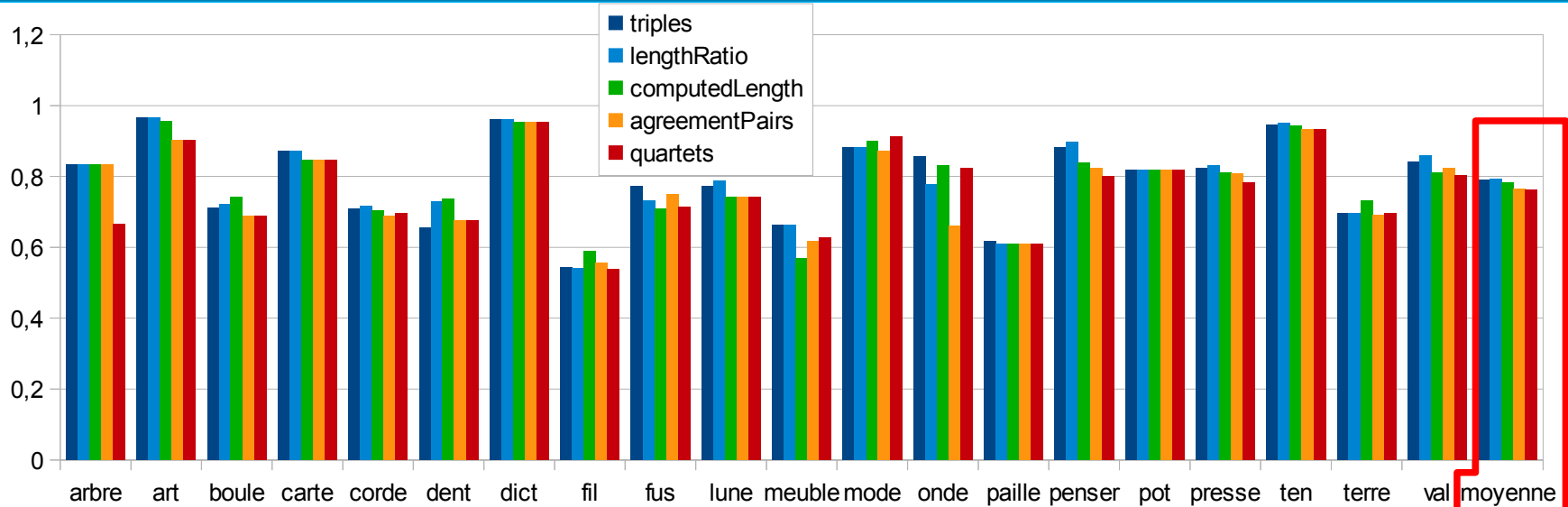
Protocole d'évaluation



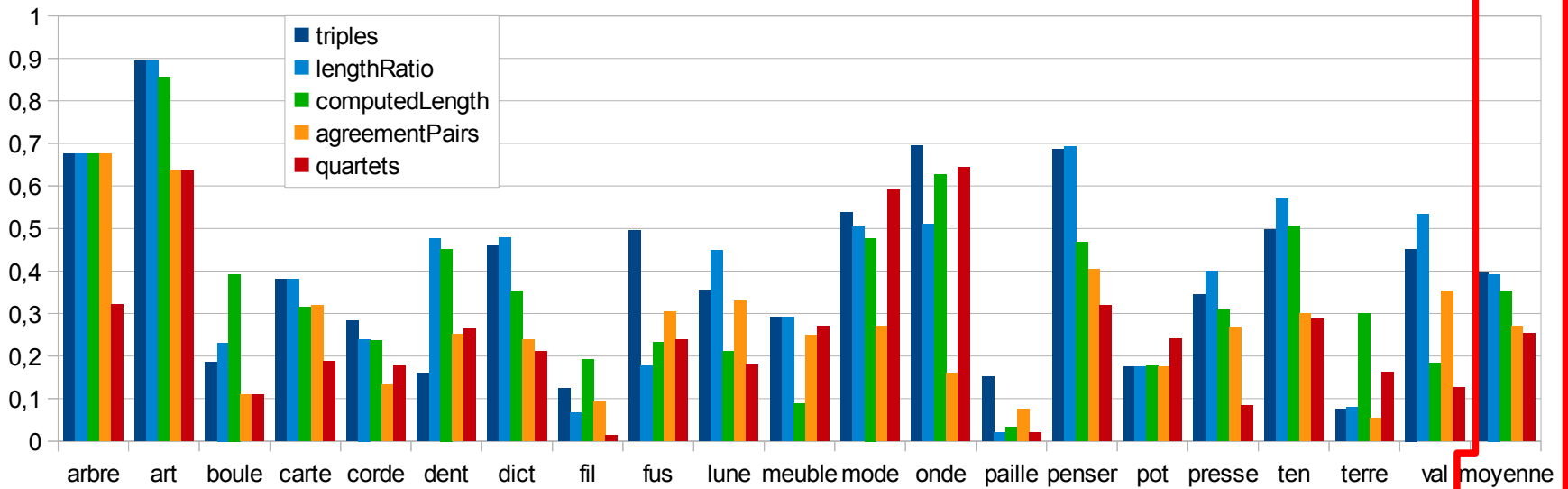
Plan

- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- **Résultats**
- Visualisations
- Perspectives

Scores de chaque formule

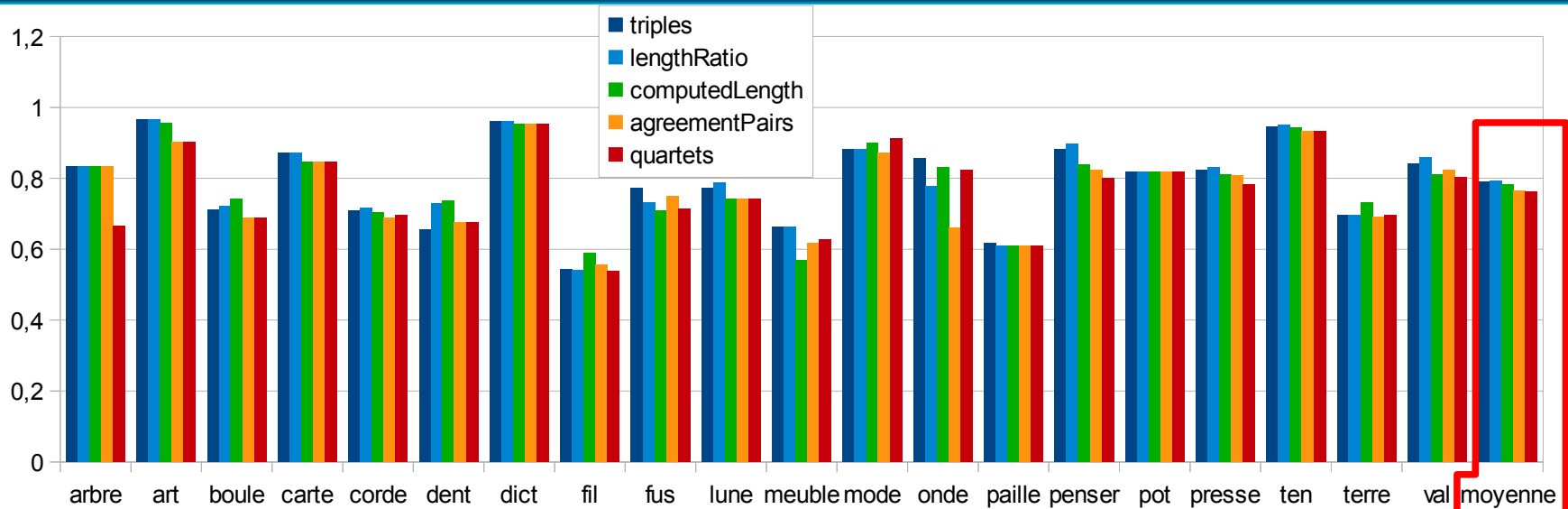


Score Rand de la meilleure partition trouvée automatiquement

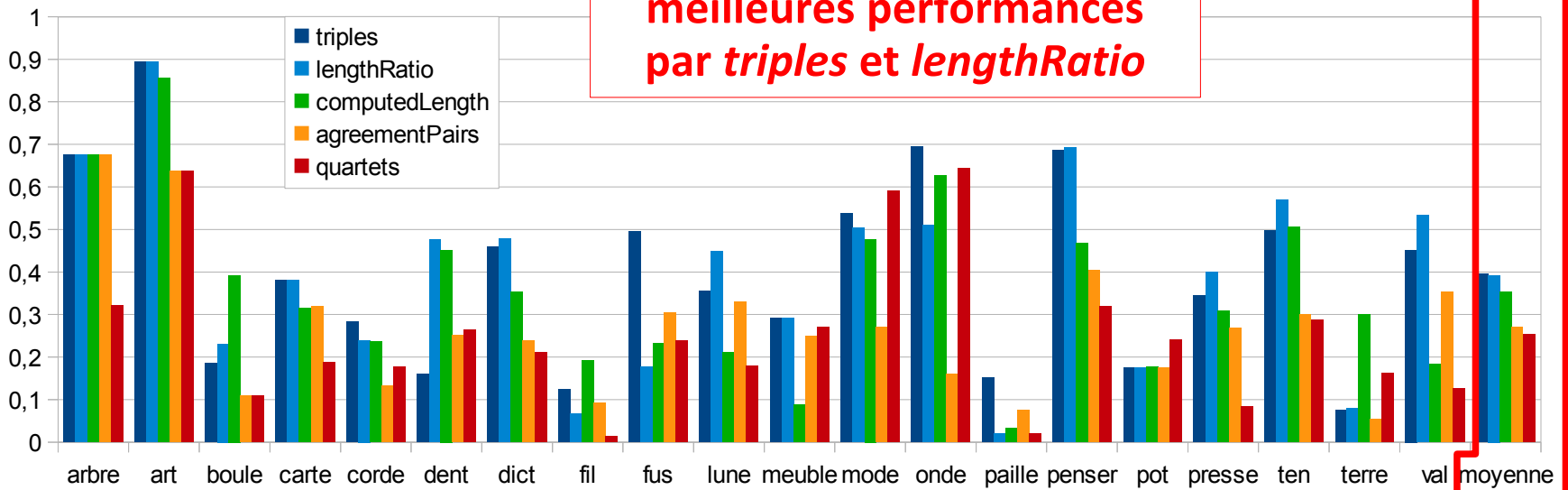


Score Rand corrigé de la meilleure partition trouvée automatiquement

Scores de chaque formule



Score Rand de la meilleure partition trouvée automatiquement



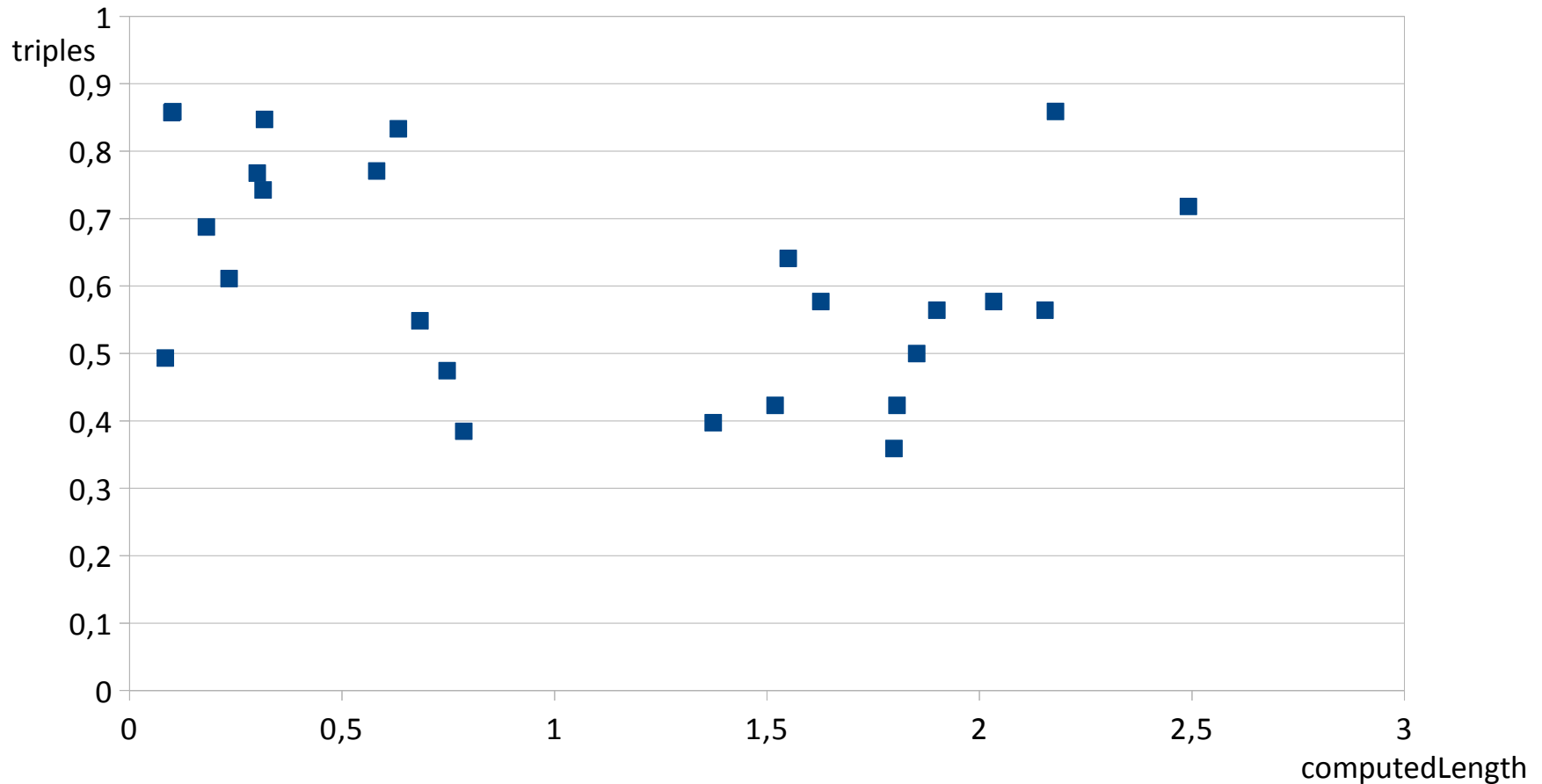
Score Rand corrigé de la meilleure partition trouvée automatiquement

Scores de chaque formule

Les formules de longueur d'arête sont-elles **cohérentes** ?

Scores de chaque formule

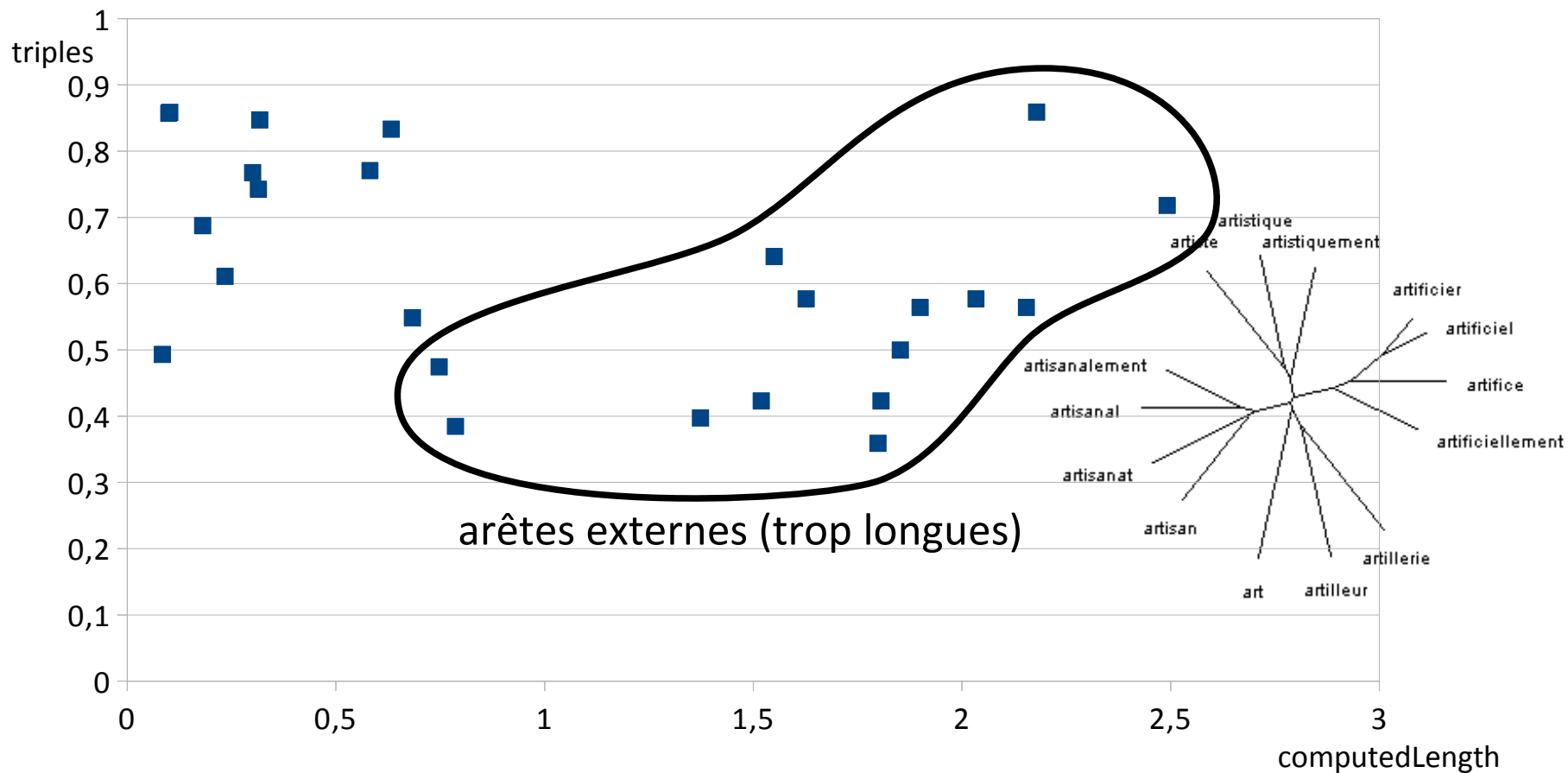
Les formules de longueur d'arête sont-elles **cohérentes** ?



Longueur selon la formule *triples* en fonction de la longueur originale de l'arête pour l'arbre de la famille de *art*

Scores de chaque formule

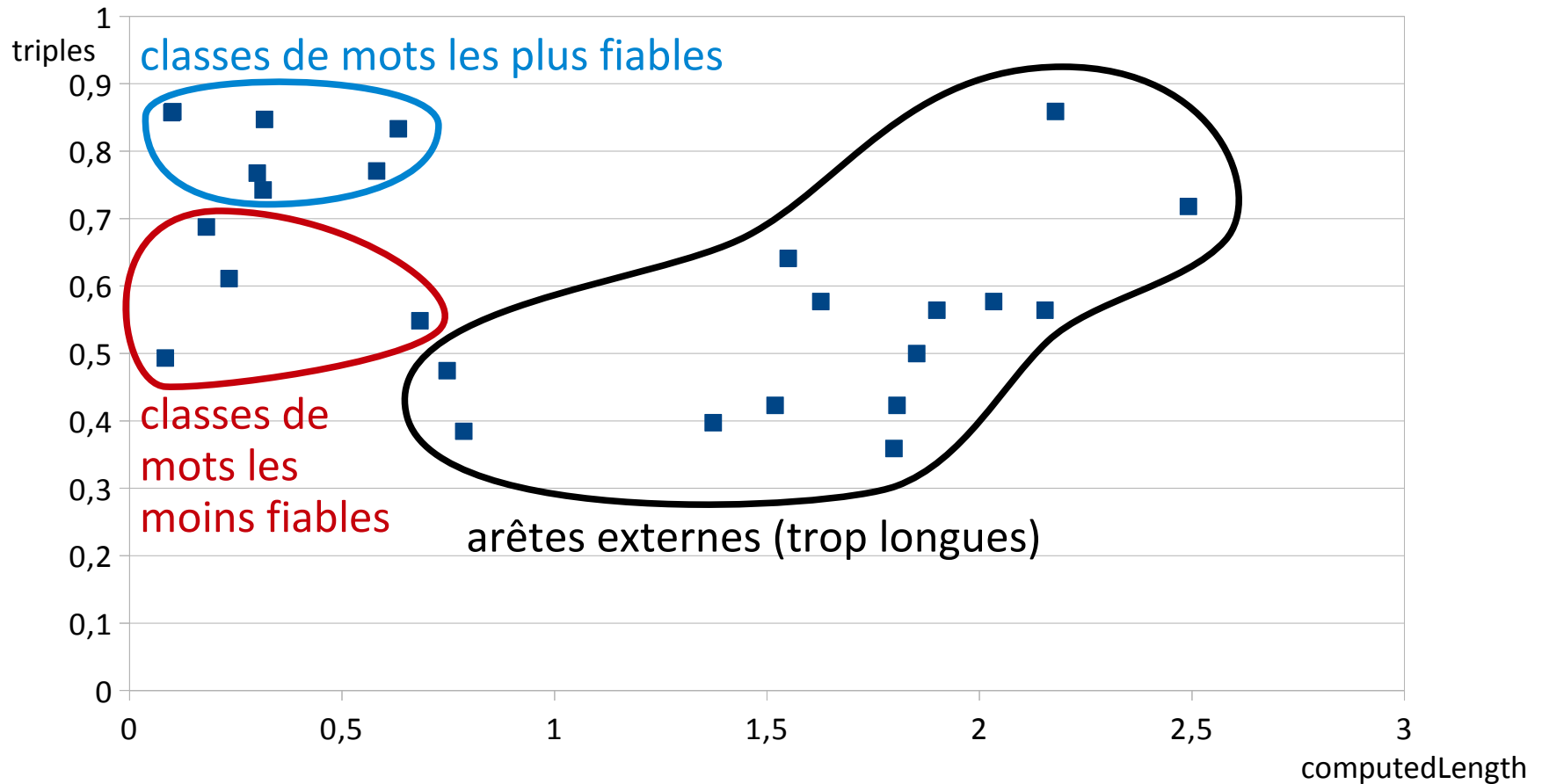
Les formules de longueur d'arête sont-elles **cohérentes** ?



Longueur selon la formule *triples* en fonction de la longueur originale de l'arête pour l'arbre de la famille de *art*

Scores de chaque formule

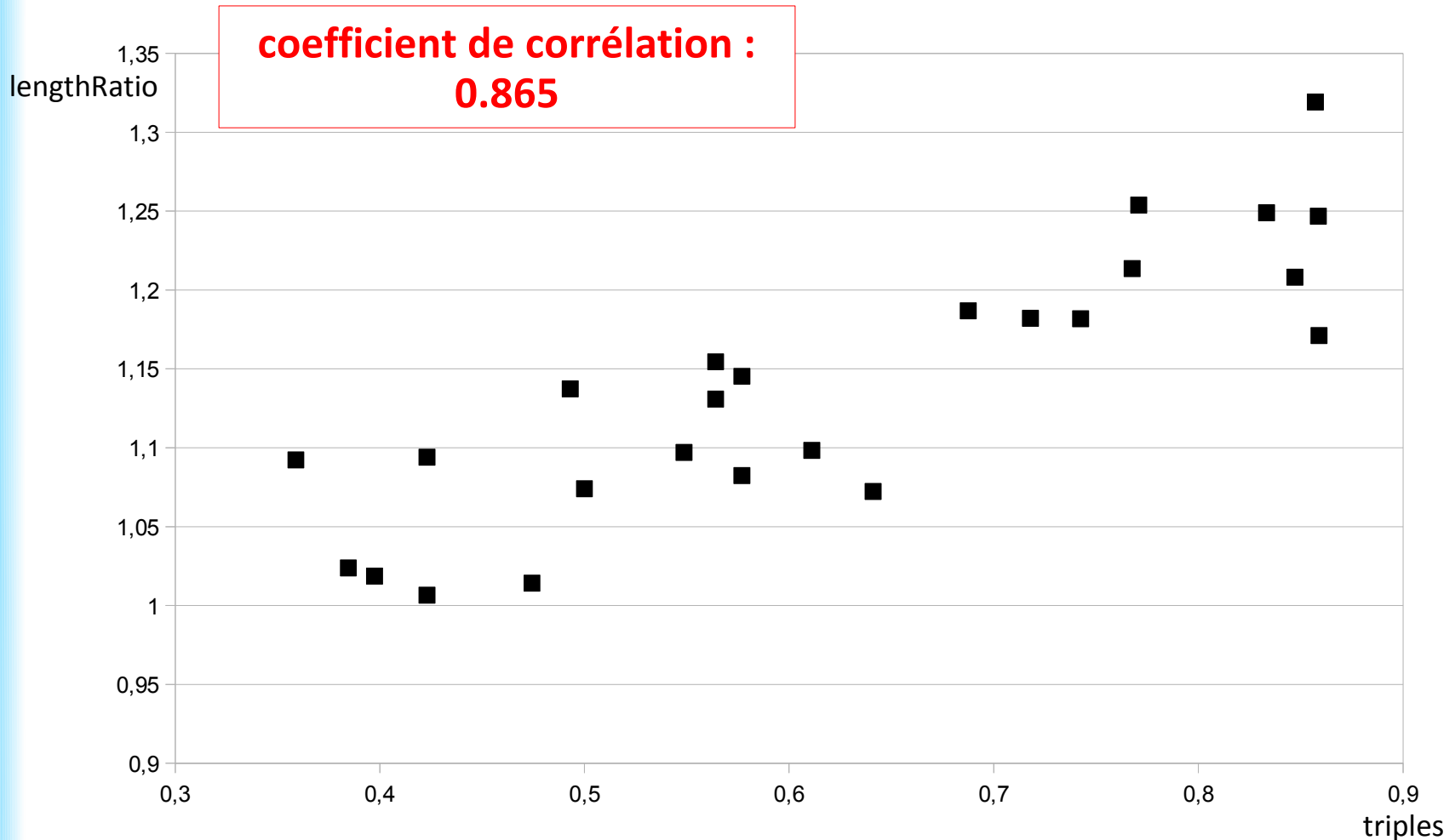
Les formules de longueur d'arête sont-elles **cohérentes** ?



Longueur selon la formule *triples* en fonction de la longueur originale de l'arête pour l'arbre de la famille de *art*

Scores de chaque formule

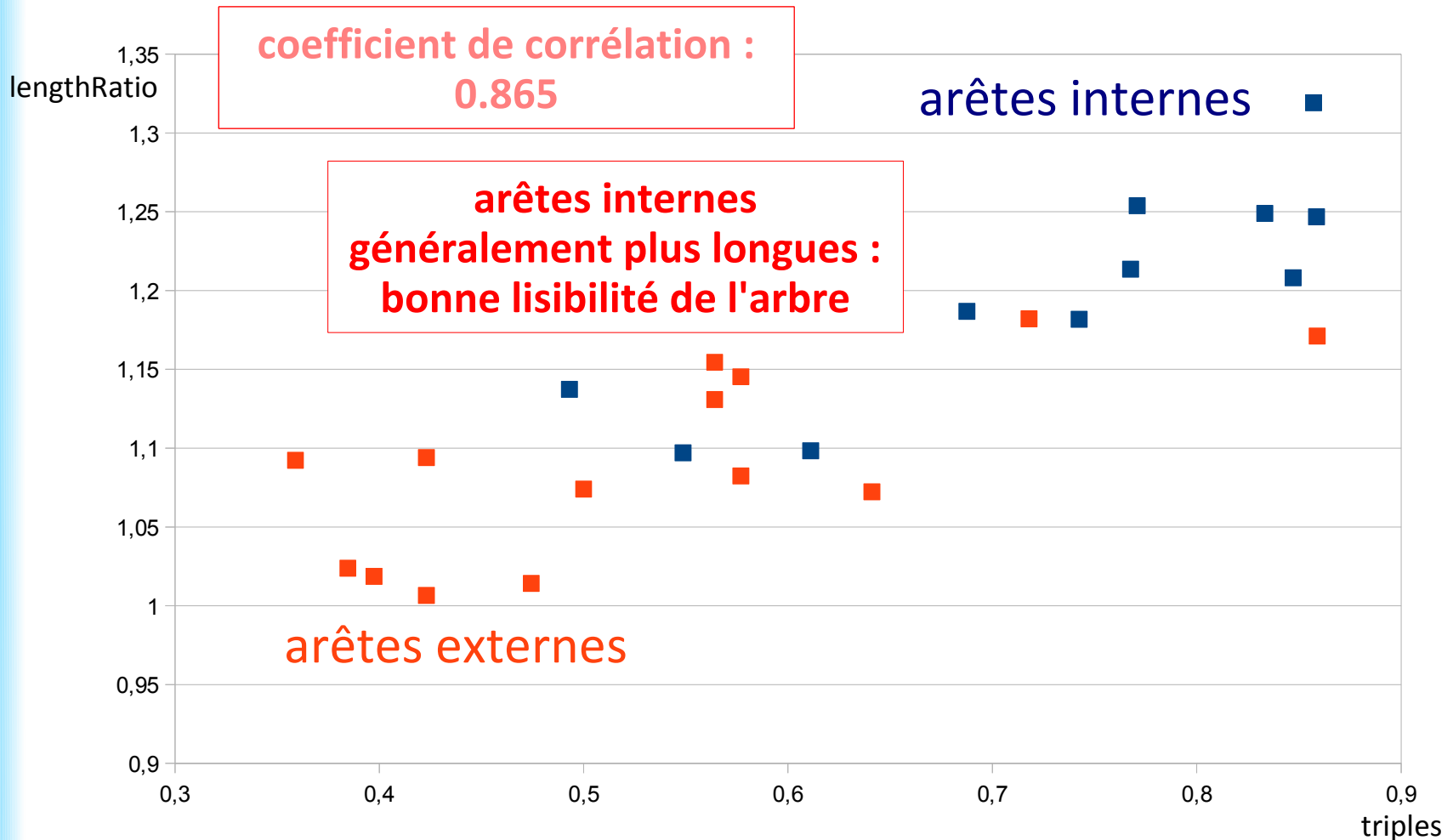
Les formules de longueur d'arête sont-elles **cohérentes** ?



Longueur selon la formule *lengthRatio* en fonction de celle selon la formule *triples* pour l'arbre de la famille de *art*

Scores de chaque formule

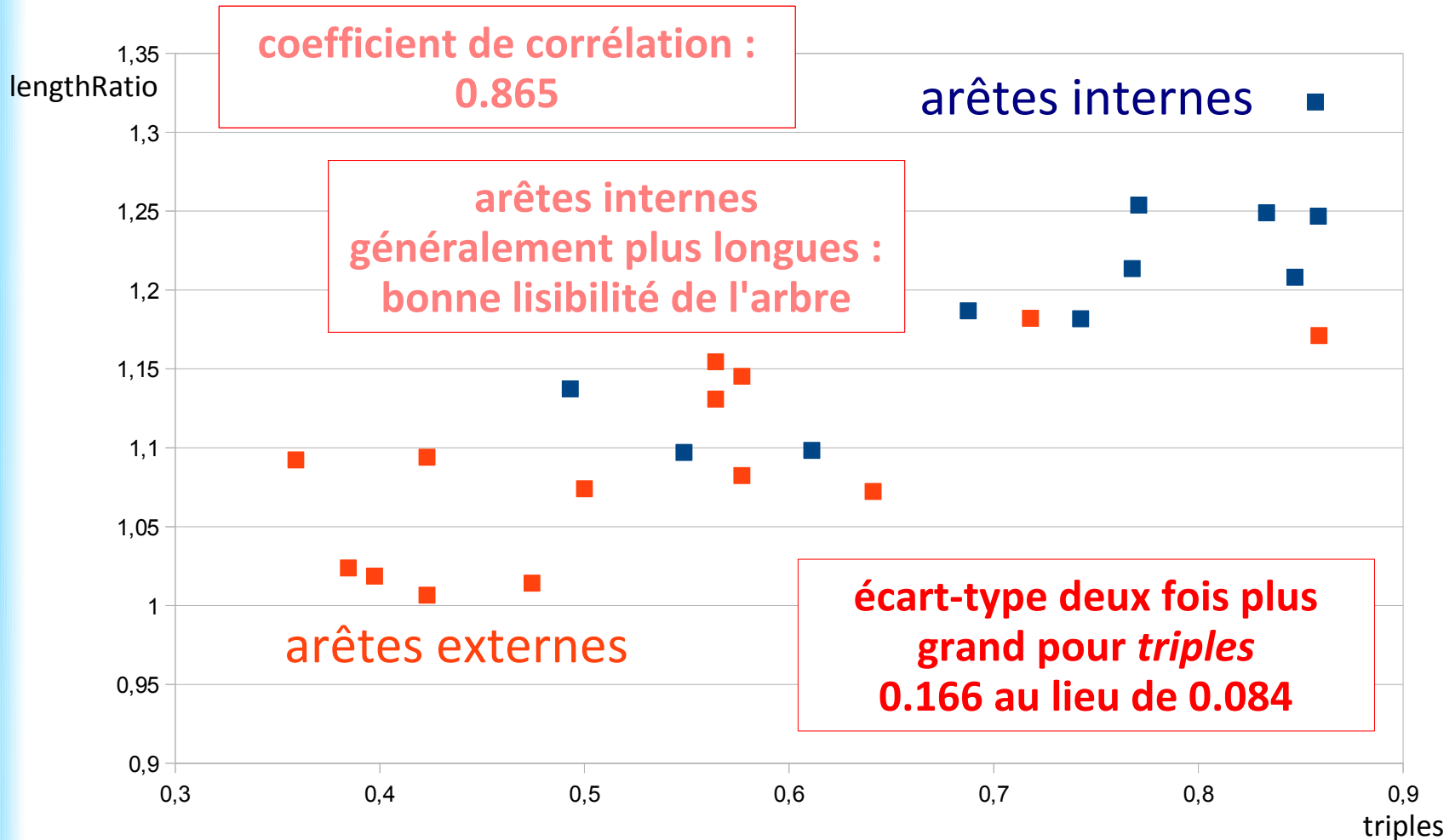
Les formules de longueur d'arête sont-elles **cohérentes** ?



Longueur selon la formule *lengthRatio* en fonction de celle selon la formule *triples* pour l'arbre de la famille de *art*

Scores de chaque formule

Les formules de longueur d'arête sont-elles **cohérentes** ?



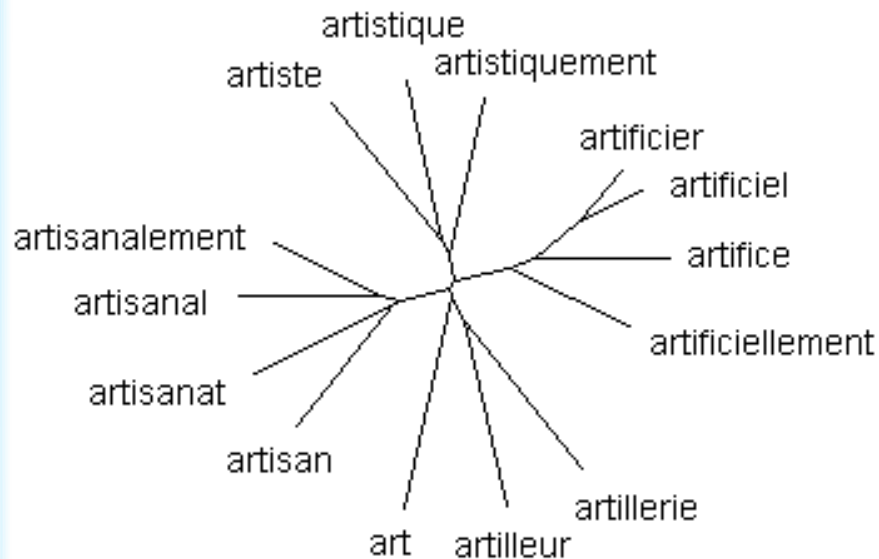
Longueur selon la formule *lengthRatio* en fonction de celle selon la formule *triples* pour l'arbre de la famille de *art*

Plan

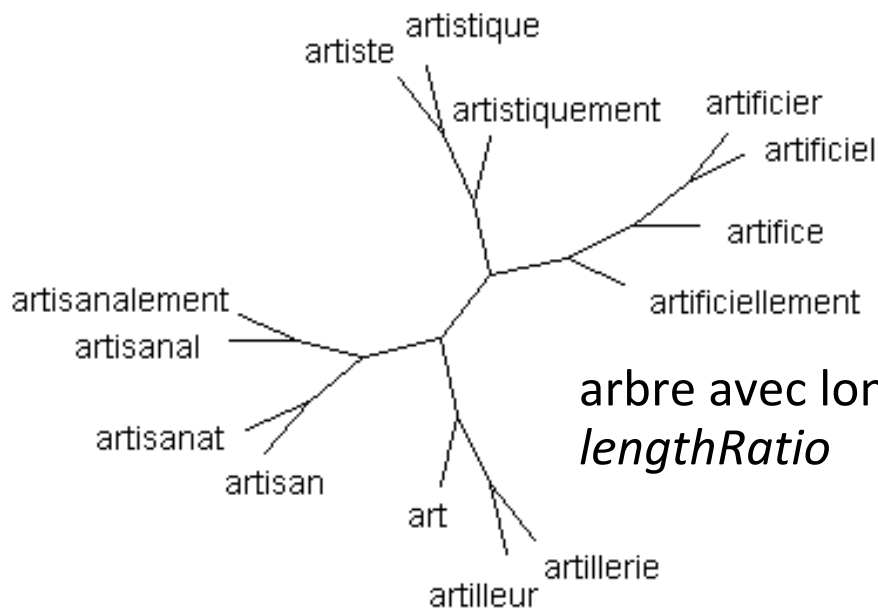
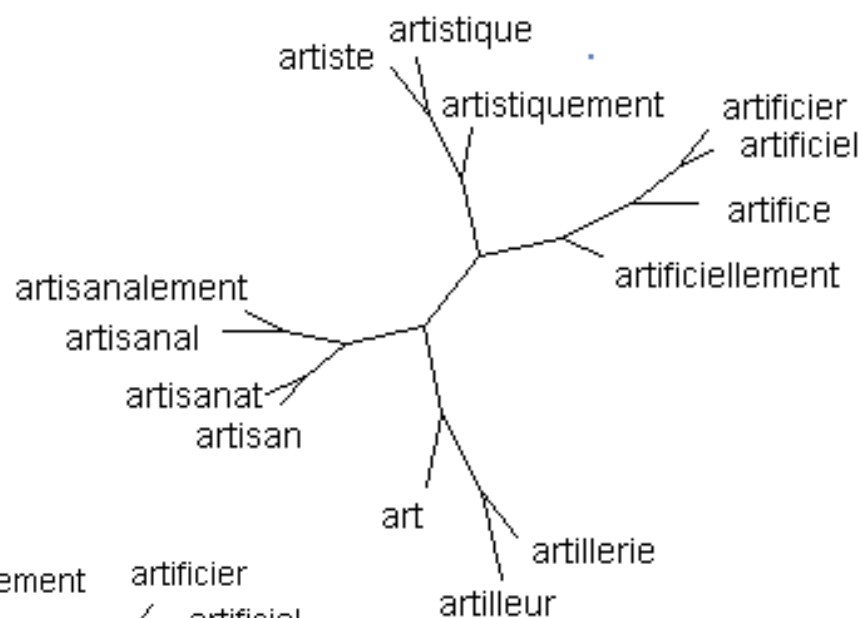
- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- Résultats
- **Visualisations**
- Perspectives

Visualisations

arbre original *computedLength*



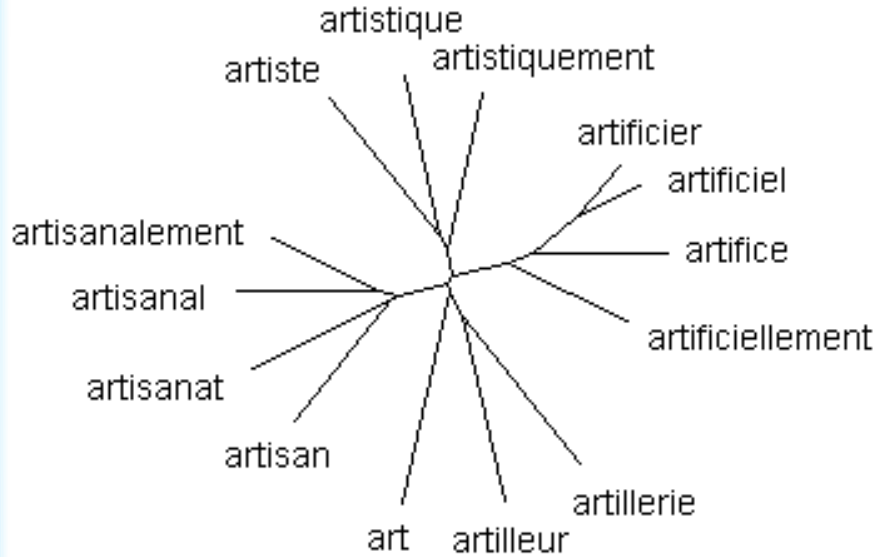
arbre avec longueurs d'arêtes *triples*



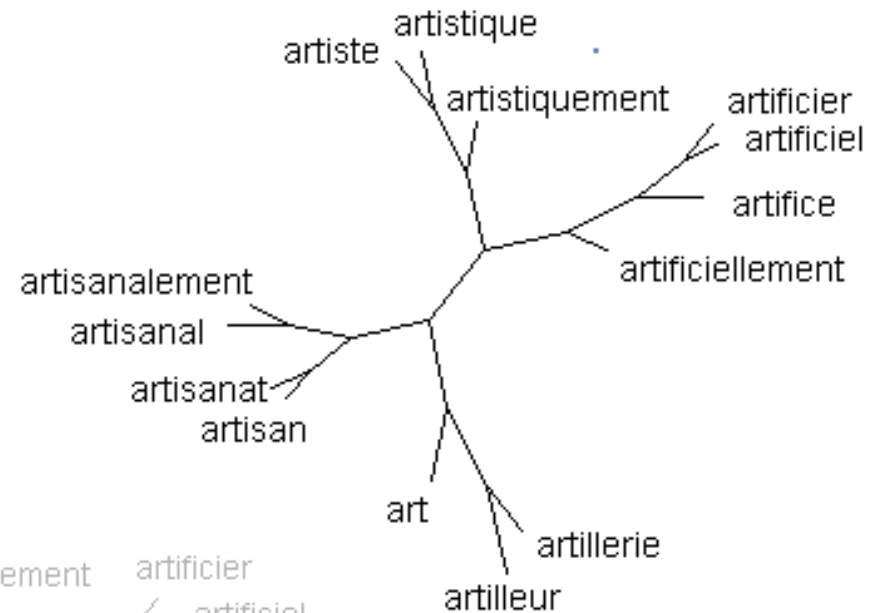
arbre avec longueurs d'arêtes *lengthRatio*

Visualisations

arbre original *computedLength*



arbre avec longueurs d'arêtes *triples*



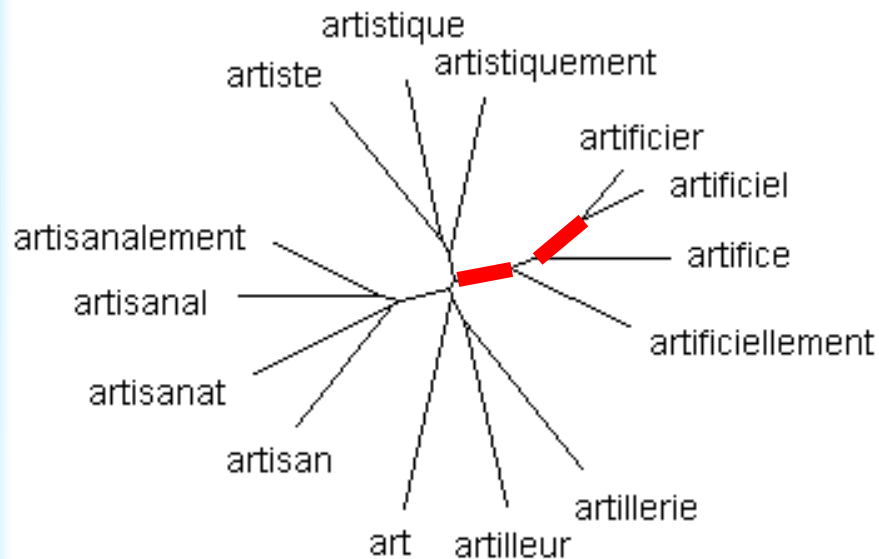
**variance trop faible des
longueurs d'arêtes !**

arbre avec longueurs d'arêtes
lengthRatio

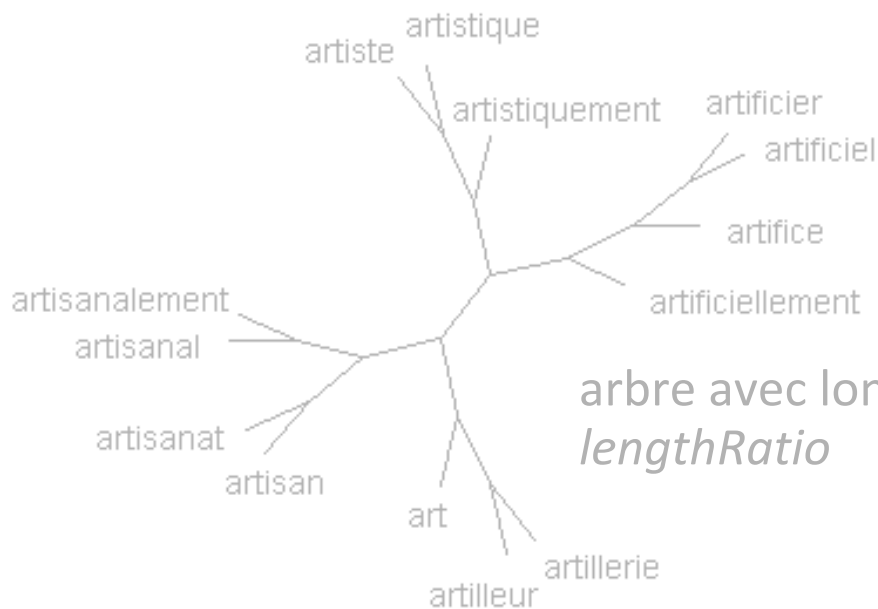
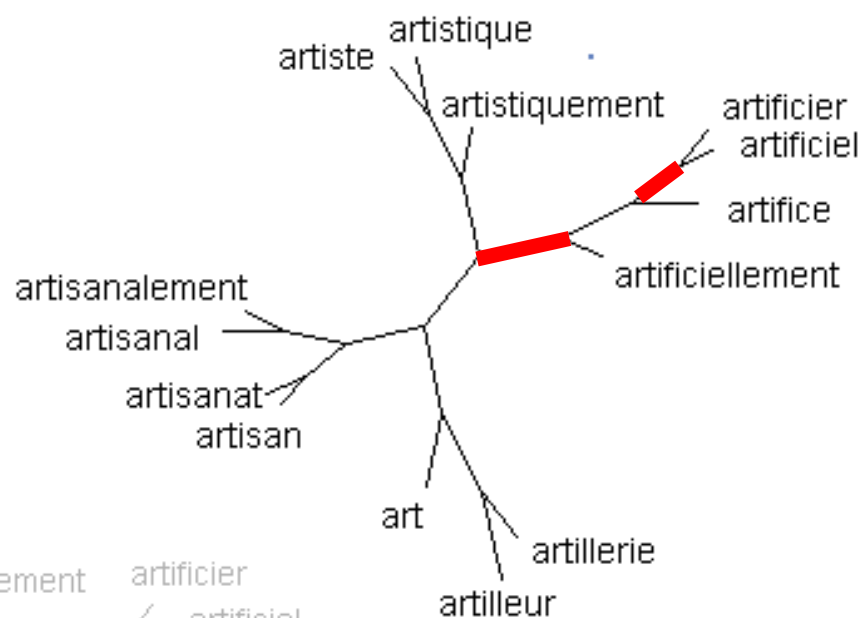


Visualisations

arbre original *computedLength*



arbre avec longueurs d'arêtes *triples*



arbre avec longueurs d'arêtes *lengthRatio*

Plan

- Analyses et nuages arborés
- Interprétation visuelle
- Formules de longueurs d'arêtes
- Protocole d'évaluation
- Résultats
- Visualisations
- **Perspectives**

Perspectives

- intégration des données de **Jeux de mots** dans le protocole, et plus généralement en textométrie
 - réseau de plus de 200 000 mots et 1 200 000 liens pondérés
 - cooccurrences dans la production spontanée de mots par rapport à un mot cible
 - cohérence avec les distances de cooccurrence calculées à partir d'un texte ?

Lafourcade, *JADT'08*

- intégration de la **visualisation en nuages arborés** avec longueurs de branches post-calculées :
 - dans les outils de textométrie existants
 - par des interfaces d'import/export adaptées
 - pour faciliter le retour au texte

Questions ?



Créer! Téléchargements Galerie A propos FAQ

Créez vos propres nuages arborés !

Collez votre texte dans le cadre ci-dessous, puis cliquez sur *Envoyer* ! Attention, l'utilisateur suivant verra votre texte quand il se connectera au site, si vous ne voulez pas faire apparaître vos textes, installez plutôt [TreeCloud](http://www.treecloud.org) sur votre machine.

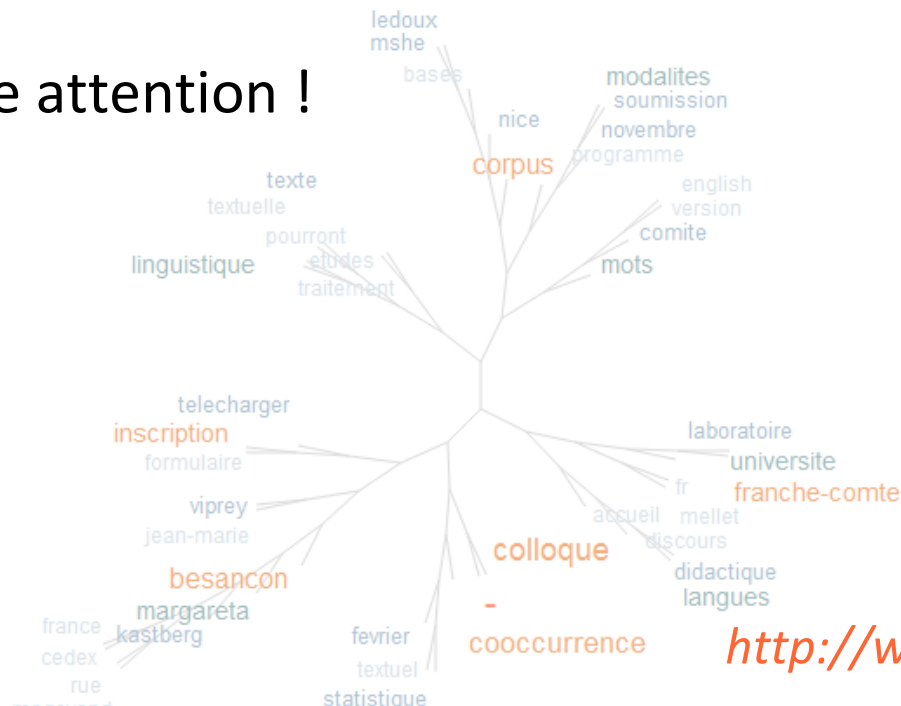
Texte :

universite
franche-comte
laboratoire
semiotique
linguistique

Envoyer

Vous pouvez déplacer les étiquettes par cliquer-glisser, l'étiquette reprend sa place d'origine lors d'un nouveau clic. L'infobulle indique le nombre d'occurrences du mot.

Merci pour votre attention !



<http://www.treecloud.org>