



HAL
open science

Modélisation graphique de la polysémie évolutive - 2nd bilan d'étape

Laurette Chardon, Jacques François, Justine Reynaud

► **To cite this version:**

Laurette Chardon, Jacques François, Justine Reynaud. Modélisation graphique de la polysémie évolutive - 2nd bilan d'étape. 2024. hal-04735960

HAL Id: hal-04735960

<https://hal.science/hal-04735960v1>

Preprint submitted on 14 Oct 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - ShareAlike 4.0 International License

PROJET DE
MODÉLISATION GRAPHIQUE
DE LA POLYSÉMIE ÉVOLUTIVE

2^{ème} bilan d'étape
(automne 2023 – été 2024)

Laurette **Jacques** **Justine**
CHARDON **FRANÇOIS** **REYNAUD¹**
CRISCO, EA 4255 CRISCO, EA 4255 GREYC, UMR CNRS
Université de Caen-Normandie

Le document qui suit est volumineux, passablement indigeste et donc peu adapté à une consultation rapide. Nous l'avons conçu, pour nous-mêmes et pour l'ATILF, comme un memento des résultats auxquels nous sommes parvenus en un an, entre la diffusion de notre 1^{er} bilan d'étape en septembre 2023 et celle de ce second bilan en octobre 2024. Pour une première approche en diagonale, nous énumérons ci-après les pages qui devraient retenir l'attention en priorité.

¹ Depuis septembre 2024, Justine REYNAUD a quitté le projet après nous avoir offert des données statistiques d'intérêt majeur, résumés dans la section 5 du Cahier 37 du CRISCO (p. 25-31) et dans la section 3.1 du présent bilan. Désormais, Mathieu GOUX, MC au département de linguistique de l'université de Caen et expert en humanités numériques au CRISCO, entre dans le projet comme spécialiste du traitement numérique des textes médiévaux et au-delà.

p. 8-12	Présentation de la 1 ^{ère} composante des 500 entrées de la version de démonstration, le tableau de conversion graphique des données historico-étymologiques [H-É] distillées (rectifié par rapport à la version de 2023)
p. 12-14	Présentation de la 2 ^{ème} composante, la nouvelle interface graphique (à peine évoquée en 2023)
p. 15-16	Profilage des entrées H-É à l'aide de l'amplitude (calcul de la dissémination des rubriques) et de la moyenne séculaire (localisation sur l'axe du temps) de chacune.
p. 17-19	Éventail de paires {tableau de conversion & interface graphique} illustrant des positionnements extrêmes
p. 20	Conclusion sur la version de démonstration
p. 21-30	Résultats quantitatifs sur la totalité des entrées dotées d'une polysémie évolutive
p. 38-41	Conclusion générale, notamment la question de la motivation des changements sémantaxiques au fil des siècles.

SOMMAIRE

1	RAPPEL DES DEUX PHASES PRÉCÉDENTES DU PROJET	4
2	VERSION DE DÉMONSTRATION (500 entrées)	7
2.1	L'élaboration des tableaux de conversion graphique	8
2.2	Le format des interfaces graphiques	12
2.3	Deux critères de profilage des entrées H-É (localisation et dissémination des sous-entrées)	15
2.4	Illustration des profils extrêmes parmi les 500 entrées	17
2.5	Conclusion	20
3	AU-DELÀ DE LA VERSION DE DÉMONSTRATION	21
3.1	Volume et distribution des entrées H-É polysémiques	21
3.2	La polysémie évolutive du n.m. <i>corps</i>	31
3.3	La polysémie évolutive de l'adj. et adv. <i>même</i>	35
4	CONCLUSION GÉNÉRALE	38
	RÉFÉRENCES BIBLIOGRAPHIQUES	42
	INDEX DES 12 VOCABLES INSPECTÉS DE PRÈS	44
	GLOSSAIRE	44

1. RAPPEL DES DEUX PHASES PRÉCÉDENTES DU PROJET

◆ Phase 1 (2020-2021) : examen approfondi des entrées historiques du TLFi

Les grandes lignes du projet de **MODÉLISATION GRAPHIQUE DE LA POLYSÉMIE ÉVOLUTIVE** (MGPE) ont été esquissées par Jacques François au début de 2020 dans le prolongement d'études antérieures sur le processus des **extensions de sens** (cf. François 2008, 2009, 2012).

Les bases de ce questionnement ont été jetées notamment par Hermann Paul (à partir de 1880) en Allemagne, puis Arsène Darmesteter (1887) et Michel Bréal (1897) en France à la fin du 19^e siècle avant les études plus approfondies de Krystoffer Nyrop (1913), Johannes Sperber (1923), Gustav Stern (1931) et Stephen Ullmann (1951, 1957, 1962) dans la première moitié du 20^e siècle, jusqu'à Dirk Geeraerts (1997, 2010) et Andreas Blank (1997) au tournant du 21^e siècle.

L'idée de base de ce projet tenait au constat que les entrées **ÉTYMOLOGIE ET HISTOIRE** (abrégées ici en « entrées H-É ») des articles du TLFi sont une mine d'informations historiques accumulée entre ±1970 et ±1990 par un groupe de rédacteurs de l'INaLF, notamment à partir du *Französisches Etymologisches Wörterbuch* élaboré initialement par Wilhelm von Wartburg, mais que cette ressource, très clairement documentée, se présente dans un format qui entrave son exploitation effective.

L'informatisation des entrées lexicographiques du TLFi est **fonctionnelle**, tandis que celle des entrées H-É est seulement **formelle**. En d'autres termes, les premières sont annotées dans le langage SGML (progressivement transformé en HTML, puis XML) de manière à attribuer une fonction à chaque segment, ce qui permet une **exploitation transversale**, tandis que les secondes sont annotées de manière à attribuer à chaque segment un format typographique (caractères romains, italiques, gras et petites capitales).

L'obtention d'une **annotation fonctionnelle** (par ex. date de 1^{ère} attestation, forme de mot ou collocation, etc.) à partir d'une **annotation formelle** ne serait pas problématique, s'il existait une projection bijective entre les deux ensembles. Mais c'est loin d'être le cas, ce qui permet de comprendre pourquoi l'informatisation des entrées historiques a débouché sur un instrument largement inexploitable (cf. François 2020a pour plus de détails).

Cependant un faisceau de régularités conditionnelles² a pu être dégagé qui s'avère suffisant pour l'attribution à chaque entrée d'un tableau de conversion graphique adapté à ce qui peut être raisonnablement visualisé dans l'interface graphique.

◆ Phase 2 (2022-2023) : Début de la segmentation fonctionnelle des entrées H-É

Le projet aurait pu en rester aux considérations théoriques contenues dans François (2020a, 2020b et 2021) si Justine REYNAUD, suivie par Laurette CHARDON n'avaient décidé, au début de 2022, d'apporter au projet leur compétence d'informaticienne afin de tester la possibilité de l'informatisation, au moins partielle, des entrées H-É du TLFi et de l'élaboration d'une interface graphique destinée aux internautes intéressés par l'histoire de la langue française.

Les résultats de cette seconde phase ont porté en priorité sur la nomenclature complète des entrées et l'enregistrement de plusieurs de leurs propriétés distinctives, notamment le **nombre** des rubriques de chaque entrée, la **localisation** (précoce, médiane ou tardive) sur l'axe chronologique de la 1^{ère} attestation du sens véhiculé par la 1^{ère} et par la dernière rubrique de chaque entrée, et le degré de **dissémination** des rubriques (compacte, fragmentée ou distendue).

Et sur cette base, il est apparu que le mode de visualisation du profil évolutif de la polysémie de chaque entrée le plus commode consiste à déployer un **arbre d'enchaînement des extensions de sens** (et de constructions syntaxiques) plaqué sur une **grille évolutive** où les siècles figurent en colonnes et les rubriques en lignes. Les principaux résultats de cette seconde phase du projet sont publiés dans deux formats : Chardon, Reynaud et François (Cahier du CRISCO 37, 2023) accorde une place notable à la dimension informatique, tandis que François, Chardon et Reynaud (BSL 117, 2023, mais publié en mai 2024) développe la dimension de sémantique historique.

² On appellera « **régularités conditionnelles** » des correspondances régulières entre une annotation formelle et une interprétation fonctionnelle sous réserve d'un contexte donné. Ainsi par ex. l'usage des italiques est double dans les entrées H-É pour identifier soit des formes de mot ou des collocations dans le corps d'une rubrique, après la datation de la 1^{ère} attestation et avant la définition entre guillemets, soit le titre d'une publication de référence dans le contexte d'un segment entre parenthèses courbes. L'un des problèmes pour l'informatisation de ces entrées consiste à s'assurer d'un contexte **conditionnant** l'une ou l'autre fonction. Ainsi, la composante Étymologie n'est pas clairement séparée de la composante Histoire, si bien qu'à l'heure actuelle, des segments relevant de l'étymologie figurent encore par erreur dans la colonne « forme de mot, collocation » du tableau de conversion graphique. Ces incongruités sont en cours de correction.

◆ **Précision terminologique : polysémie évolutive ou cumulative ?**

Jusqu'à présent nous avons placé ce rapport de recherche, comme dans les différentes publications antérieures de ce projet depuis 2020 sous le signe de la « polysémie évolutive ». Dans plusieurs de ses premiers travaux sur ce sujet J. François (cf. 2008 : 9-11 et 2020 : 86) a cherché à repérer l'époque du déclin de certains sens (initial ou extensions successives), mais les entrées H-É du TLFi sont muettes sur ce point.

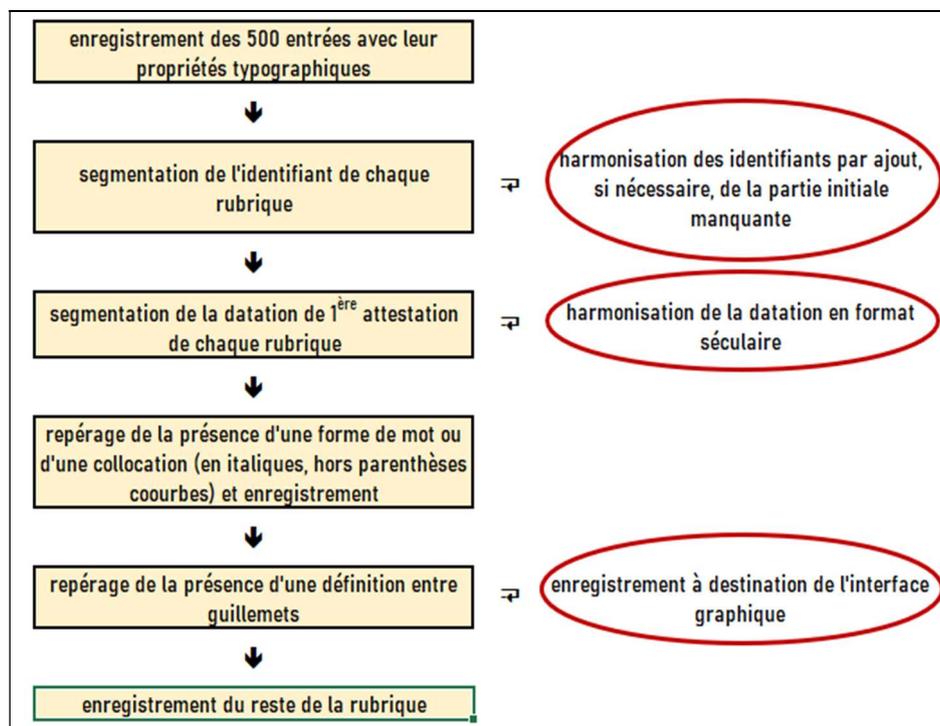
Seul un corpus diachronique couvrant toute l'histoire de la langue française, comme le fait FRANTEXT, permettrait d'enregistrer la progression et le déclin d'un sens particulier d'un vocable, et à condition qu'il soit suffisamment présent tout au long de cette histoire afin de pouvoir mesurer sa fréquence relative au nombre d'occurrences du vocable de siècle en siècle. D'autre part, il se trouvera toujours, dans un corpus littéraire, des auteurs adeptes d'un style archaïsant qui continueront à employer un vocable dans un sens ancien qu'une large part de ses lecteurs devra deviner à partir du contexte.

Nous précisons donc, afin de clarifier les limites de notre entreprise, dans le sillage des entrées H-É du TLFi que l'expression « polysémie **évolutive** » s'entend ici dans un sens limité qui fait l'impasse sur la question des emplois en voie de déclin. Les expressions « polysémie **cumulative** » ou « **profil cumulatif de la polysémie d'un vocable** » seraient plus précises.

2. VERSION DE DÉMONSTRATION (500 entrées)

Avant de passer à l'informatisation, ciblée et donc partielle, des entrées H-É du TLFi, nous avons résolu au début de 2024 de tester une version de démonstration, destinée à faire éventuellement apparaître des carences majeures, sur un ensemble de 500 entrées aussi différentes que possible³. La figure 1 représente l'enchaînement des opérations.

Figure 1 : Diagramme des principales opérations destinées au renseignement des cellules du tableau de conversion graphique des 500 entrées H-É



³ La sélection des 500 entrées n'est pas complètement aléatoire. En premier lieu, ces entrées comptent toutes entre 3 et 10 rubriques dans leur tableau de conversion (le nombre de nœuds dans l'interface graphique étant souvent supérieur en raison de l'introduction de nœuds auxiliaires pour les rubriques superordonnées absentes de l'entrée originale). Par ailleurs, de nombreux usagers réguliers de notre *Dictionnaire Électronique des Synonymes* nous ont demandé d'ajouter de nouveaux liens synonymiques et nous avons emprunté la grande majorité des 500 entrées à ces demandes. Les autres sont des vocables dont nous avons eu l'occasion d'examiner la polysémie dans des études précédentes. Il s'agit de *chemin, pince, lentille, citerne, lien, cirieux, -euse, chiqueter, deux, château, gracieux, -euse, impersonnel, -elle, insensible, pacifique, vulnérable* et deux entrées auxquelles nous avons consacré une interface graphique explicative en format powerpoint (cf. p. 13, notes 4-5).

2.1. L'élaboration des tableaux de conversion graphique

Nous prendrons un exemple simple, celui de l'entrée H-É du n.f. *ardeur*, qui compte cinq rubriques. Le premier stade d'enregistrement de cette entrée est la version fournie par le CNRTL-ORTOLANG ► <https://www.cnrtl.fr/etymologie/ardeur> :

Figure 2 : Format de l'entrée du n.f. *ardeur* dans l'onglet étymologie du TLFi (accès par le portail lexical du CNRTL-ORTOLANG)

Ortolang
Outils et Ressources pour un Traitement Optimisé de la LANGue

Accueil | Portail lexical | Corpus | Lexiques | Dictionnaires | Métalexigraphie | Outils | Contact

Morphologie | Lexicographie | **Etymologie** | Synonymie | Antonymie | Proxémie | Concordance | Aide

Entrez une forme : ardeur Chercher

notices corrigées catégorie : toutes

ARDEUR, subst. fém.

ÉTYMOL. ET HIST. – 1. 1120-50 « chaleur brûlante » (*Gd mal fit Adam*, éd. H. Suchier, I, 105 ds T.-L. : *Senz repos ardor aux enfers*); 2. au fig. **a)** 1170-91 « désir charnel » (Bérroul, *Tristan*, éd. E. Muret, 1195, *ibid.* : *Sire, en nos les lépreux a si grant ardor*, *Soz ciel n'a dame qui un jor Pèust sofrir nostre convers*); **b)** 1203 « passion amoureuse » (*Chastelain de Coucy*, éd. Crapelet, 239, *ibid.* : *Moult souvent souspiroit d'ardour*); **c)** mil. xiv^es. « désir violent » (*Li Batard de Bouillon*, éd. A. Scheleer, Bruxelles 1877, 170, *ibid.* : *de combattre as Turs avoient grant ardour*). **Rem.** Nombreuses attest. au masc., au xvi^es. (Hug.). Empr. au lat. *ardor* attesté dep. Accius au sens 1 (*Trag.* 581 ds TLL s.v., 489, 48), dep. Lucrece au sens 2 b (4, 1086 ds OLD, I, 165) et dep. Cicéron au sens 2 c (*Orat.*, 108, *ibid.*).

Le second est la version XML de la même entrée (Colonne de gauche du tableau p. 5). Pour en rendre la lecture plus aisée, nous reproduisons dans la colonne de droite l'entrée 'publique' ci-dessus dans un format linéarisé et avec les identifiants complets des rubriques. Et nous surlignons dans les deux parties

- en **jaune** les définitions
- en **bleu** les segments en gras
- en **vert** les segments en italiques
- et en **gris** les segments en petites capitales.

Comme nous l'avons déjà montré dans nos publications précédentes (François 2020a, François, Chardon & Reynaud 2023 ; Chardon, Reynaud & François 2023), et à titre d'exemple, les caractères gras ne sont instructifs que si l'on a pratiqué préalablement une segmentation apte à repérer parmi ceux-ci les identifiants des rubriques et la vedette dans les citations. Cela présuppose

- [1] le repérage du signe < ;> avant l'un des cinq débuts possibles d'un identifiant (par ordre hiérarchique décroissant : chiffre romain > lettre majuscule > chiffre arabe > lettre minuscule > lettre grecque) ou
- [2] la présence d'un segment en caractères romains précédé du signe <:> pour la vedette d'une citation

Tableau 1 : Formats XML et public comparés de l'entrée H-É du n.f. *ardeur*

<p><entry> <form>ARDEUR,</form> <gram>subst. fém.</gram> <etymology> <hi rend="bold">ÉTYMOL. ET HIST. _ 1.</hi> 1120-50 « chaleur brûlante » (<hi rend="italic">Gd mal fit Adam,</hi> éd. H. Suchier, I, 105 ds T.-L. : Senz repos <hi rend="bold">ardor</hi> aux enfers); <hi rend="bold">2.</hi> au fig. <hi rend="bold">a)</hi> 1170-91 « désir charnel » (<hi rend="smallcaps">Béroul</hi> , <hi rend="italic">Tristan,</hi> éd. E. Muret, 1195, <hi rend="italic">ibid.</hi> : Sire, en nos les lépreux a si grant <hi rend="bold">ardor</hi> , Soz ciel n'a dame qui un jor Pëust sofrir nostre con- vers); <hi rend="bold">b)</hi> 1203 « passion amoureuse » (<hi rend="italic">Chastelain de Coucy,</hi> éd. Crapelet, 239, <hi rend="italic">ibid.</hi> : Moult souvent souspiroit d' <hi rend="bold">ardour</hi>); <hi rend="bold">c)</hi> mil. <hi rend="smallcaps">xiv</hi> <R/> es. « désir violent » (<hi rend="italic">Li Batard de Bouillon,</hi> éd. A. Scheler, Bruxelles 1877, 170, <hi rend="italic">ibid.</hi> : de combattre as Turs avoient grant <hi rend="bold">ardour</hi>). <hi rend="bold">Rem.</hi> Nombreuses attest. au masc., au <hi rend="smallcaps">xvi</hi> <R/> es. (<hi rend="smallcaps">Hug.</hi> ' [...] </etymology> </entry></p>	<p>ARDEUR, subst.fém ÉTYMOL. ET HIST. _ 1. 1120-50 « chaleur brûlante » (<i>Gd mal fit Adam</i>, éd. H. Suchier, I, 105 ds T.-L. : Senz repos ardor aux enfers); 2. au fig. 2.a) 1170-91 « désir charnel » (BÉROUL , <i>Tristan</i>, éd. E. Muret, 1195, <i>ibid.</i> : Sire, en nos les lépreux a si grant Ardor , Soz ciel n'a dame qui un jor Pëust sofrir nostre convers); 2.b) 1203 « passion amoureuse » (<i>Chastelain de Coucy</i>, éd. Crapelet, 239, <i>ibid.</i> : Moult souvent souspiroit d' Ardour) ; 2.c) mil. Xiv es. « désir violent » (<i>Li Batard de Bouillon</i>, éd. A. Scheler, Bruxelles 1877, 170, <i>ibid.</i> : de combattre as Turs avoient grant ardour). Rem. Nombreuses attest. au masc., au Xvi es. (Hug.).</p>
---	--

Pour modéliser la polysémie évolutive (ou ‘cumulative’) d’une entrée H-É, il est indispensable de sélectionner les données qui méritent de figurer dans le tableau de conversion graphique au côté de l’interface graphique. Nous avons retenu ⇨

- [1] les **identifiants de rubriques**,
- [2] les **datations** de première attestation,
- [3] **les formes de mot ou collocations** en italiques (hors parenthèses courbes) et
- [4] le reste de la (1^{ère} partie de) la rubrique, avec (sauf en cas d’absence, laquelle renvoie implicitement à la définition de la rubrique immédiatement antérieure) la **définition** entre guillemets,

ce qui définit *in fine* le format de chaque tableau de conversion en quatre colonnes.

Toutefois, une partie importante des identifiants de rubriques sont incomplets, le premier caractère n’étant pas répété dans les rubriques subordonnées. Il a donc fallu programmer la restitution des identifiants complets. Et les datations de première attestation ont une variété de formats (cf. François 2020a : 59). Un second programme a permis de transformer tous ces formats dans le format le plus général, celui du siècle.

S’agissant de l’entrée H-É du n.f. *ardeur*, le tableau 2 subdivise son contenu entre trois types de segments, à gauche **enregistrés**, au centre **reformulés** et à droite **écartés**. Il est à noter que le contenu de (la 1^{ère} partie de) certaines rubriques dépasse les limites d’une cellule du tableau de conversion. Pour prendre connaissance de la totalité de ce contenu (cf. p. 4), il suffit de faire glisser la souris sur la cellule en cause. La totalité des données est téléchargeable en cliquant sur l’onglet **EXPORT**. Et l’on peut vérifier que, pour la rubrique 2.c, la remarque complémentaire a été enregistrée.

Tableau 2 : Élaboration du tableau de conversion des données principales

segments enregistrés	segments à reformuler	reformulation
1.	1120-1250	⇒ 12 ^e s.
(<i>Gd mal fit Adam</i> , éd. H. Suchier, I, 105 ds T.-L. : Senz repos ardor aux enfers);		
« chaleur brûlante »		
2. au fig.		
	a)	⇒ 2.a
« désir charnel »	1170-91.	⇒ 12 ^e s
(Bérout, <i>Tristan</i> , éd. E. Muret, 1195, <i>ibid.</i> : Sire, en nos les lépreux a si grant ardor, Soz ciel n'a dame qui un jor Pëust sofrir nostre convers);		
	b)	⇒ 2.b
	1203	⇒ 13 ^e s.
« passion amoureuse »		
(<i>Chastelain de Coucy</i> , éd. Crapelet, 239, <i>ibid.</i> : Mout souvent souspiroit d' ardour);		
	c)	⇒ 2.c
mil.	xiv ^e s.	⇒ 14 ^e s.
« désir violent »		
(<i>Li Batard de Bouillon</i> , éd. A. Scheler, Bruxelles 1877, 170, <i>ibid.</i> : de combattre as Turs avoient grant ar- dour).		
Rem. Nombreuses attest. au masc., au xvi ^e s. (Hug.).		

Au final, le tableau de conversion se présente ainsi :

Tableau 3 : de conversion graphique des données principales de l'entrée H-É du n.f. *ardeur*

↓ téléchargement intégral

Export			
Niveaux	Siècle	Forme/colloc	Histoire
1	12		1120-50 « chaleur brûlante » (Gd mal fit Adam, éd. H. Suchier, I, 105 ds
2			au fig.
2.a	12		1170-91 « désir charnel » (Bérout, Tristan, éd. E. Muret, 1195, ibid. :
2.b	13		1203 « passion amoureuse » (Chastelain de Coucy, éd. Crapelet,
2.c	14		mil. xives. « désir violent » (Li Batard de Bouillon, éd. A. Scheler,

2.2. Le format des interfaces graphiques

Plusieurs formats ont été testés dans nos publications antérieures. Nous ne reviendrons pas ici sur leurs propriétés distinctives. Il suffira de les évoquer succinctement.

- a) Visualisation de la polysémie évolutive d'une entrée H-É par une **série d'instantanés** siècle par siècle : c'est la première sorte qui a été testée (cf. François 2020a ; diaporama⁴ du n.f. *campagne*), mais elle a trois faiblesses : elle prend beaucoup de place, chaque nouveau sens est disposé dans la zone du diagramme en rapport avec le siècle de sa 1^{ère} attestation, mais l'orientation vers le haut, le bas, la gauche ou la droite du diagramme n'est pas fixée, ce qui alourdit le processus de construction automatisée.
- b) Visualisation sous la forme d'un **arbre disposé sur un quadrillage orthogonal** avec les identifiants des rubriques en ordonnées et les siècles en abscisses (cf. François 2020b, diaporama⁵ historique du v. *gagner*). Contrairement au type (a), ce second type de représentation ne permet pas de se faire une idée précise de la structure de la polysémie du vocable à chaque siècle, mais il se laisse plus aisément programmer (notamment pour des entrées riches en rubriques généralement hiérarchisées), c'est celui que nous avons retenu.
- c) Un troisième type de visualisation a été testé (cf. François 2021). Il dérive du précédent, mais chaque sens y est enveloppé dans une ellipse

⁴ Cf. <https://hal.science/hal-04505752> (diaporama du n.f. *campagne*)

⁵ Cf. <https://hal.science/hal-04505936> (diaporama du v. *gagner*)

qui vise à mettre en évidence sa « **durée de vie** ». Il n'a pas pu être retenu, car les entrées H-É du TLF se concentrent sur la 1^{ère} attestation de chaque sens et laissent de côté leur éventuel déclin avant le début du 21^e siècle.

Pour l'entrée du n.f. *ardeur*, l'interface graphique est présentée à la page suivante.

URL de consultation : <<https://crisco3.unicaen.fr/noticestlfi/>>

Elle est complétée par celle de l'entrée n.f. *méthode*, pour montrer qu'en cas de rubrique superordonnée absente dans la version publique et dans le tableau de conversion, un nœud lui est cependant attribué dans l'interface graphique afin de dessiner correctement les arêtes de l'arbre. En effet, le tableau de conversion de méthode ne mentionne pas de rubrique 2, il introduit immédiatement les rubriques subordonnées 2.a-c.

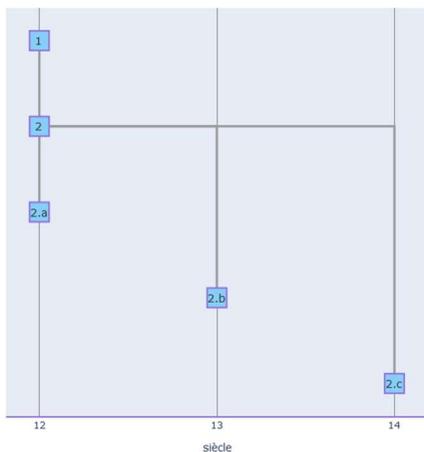
2.3. Deux critères de profilage des entrées H-É (localisation et dissémination des sous-entrées)

Le nombre de rubriques d'une entrée H-É est une donnée importante, ainsi que le siècle de 1^{ère} attestation des deux sens véhiculés respectivement par la 1^{ère} rubrique et la dernière rubrique (ou plus exactement la rubrique la plus précoce et la rubrique la plus tardive). On appellera **amplitude** le nombre de siècles (premier et dernier inclus) couvert par une entrée entre les attestations de ses différentes rubriques, et **moyenne séculaire** la somme du nombre de rubriques attestées dans chaque siècle, divisée par le nombre de rubriques.

Pour l'entrée du n.f. *ardeur*, le nombre de rubriques s'élève à 12, la première attestation étant datée du 12^e siècle et la dernière du 14^e. L'amplitude de l'entrée est donc de 3 siècles. 2 des 5 rubriques sont datées du 12^e, une du 13^e, une du 14^e siècle et la 2 n'est pas datée (sa datation est par défaut celle de 2a).

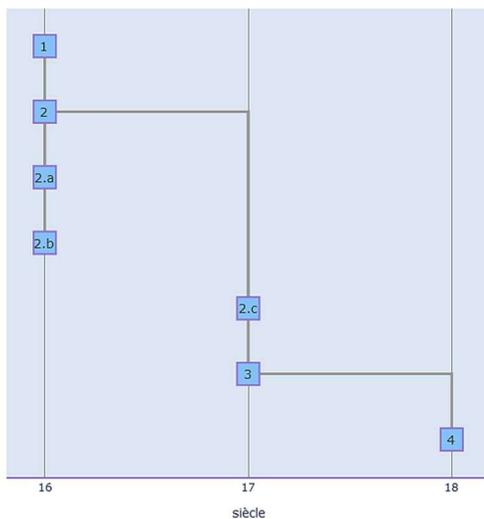
Le calcul de la **moyenne séculaire** (qui représente la localisation temporelle moyenne de l'entrée) consiste pour chaque siècle à multiplier le nombre de rubriques par le chiffre du siècle, à additionner le tout et à diviser la somme par le nombre de rubriques dont le siècle est renseigné. Ainsi pour *ardeur* on obtient : $(2*12) + (1*13) + (1*14) = 51$ et la moyenne séculaire s'élève à $51/4 = 12,75$.

ardeur



Niveaux	Siècle	Forme/colloc	Histoire
1	12		1120-50 « chaleur brûlante » (Gd mal fit Adam, éd. H. Suchier, I, 105 ds T.-L. : Senz repos ardor aux enfers);
2			au fig.
2.a	12		1170-91 « désir charnel » (Béroul, Tristan, éd. E. Muret, 1195, ibid. : Sire, en nos les lépreux a si grant ardor, Soz ciel n'a
2.b	13		1203 « passion amoureuse » (Chastelain de Coucy, éd. Crapelet, 239, ibid. : Moult souvent souspiroit d'ardour);
2.c	14		mil. xives. « désir violent » (Li Batard de Bouillon, éd. A. Scheler, Bruxelles 1877, 170, ibid. : de combattre as Turs avoient

méthode



Niveaux	Siècle	Forme/colloc	Histoire
1	16	cf. méthode curative ou	1537 méd. «manière particulière d'appliquer une médication» (J. Canappe, 4eLivre de Thérapeutique de Galien d'apr. Chauvelot ds Fr.
2.a	16		1546 «ensemble de procédés raisonnés» (Rabelais, Tiers Livre, chapitre 8, éd. M. A. Screech, p. 72: hors toute méthode);
2.b	16		1547 «manière d'enseigner» (J. Martin, Arch. Vitruve, 10 B);
2.c	17		1643 «manière de faire, habitude» (Corneille, Le Menteur, acte V, scène I, éd. Regnier, t. 4, p. 222);
3	17		1637 spéc. terme de philos. (Descartes, Discours de la méthode pour bien conduire sa raison et chercher la verité dans les sciences);
4	18		1763 hist. nat. (Adanson, Familles des plantes, éd. F. A. Staflen, Préface, xciv: Les Botanistes ont distingué 2 espèces de Metodes,

Tableau 4 : Nombre de rubriques de l'entrée du n.f. *ardeur* par siècle
(calcul de la moyenne séculaire)

ARDEUR			
9 ^e s.	0	16 ^e s.	0
10 ^e s.	0	17 ^e s.	0
11 ^e s.	0	18 ^e s.	0
12 ^e s.	2	19 ^e s.	0
13 ^e s.	1	20 ^e s.	0
14 ^e s.	1	MoySéc	12,75
15 ^e s.	0		

Quant au degré de dissémination des rubriques de l'entrée, il est calculé par l'**écart-type** entre l'ensemble des valeurs. Pour *ardeur*, il égale **0,83**. Ce nombre n'a qu'une valeur comparative. Sachant que, pour l'ensemble des 500 rubriques, l'écart-type le plus bas égale 0 et le plus élevé égale 3,76, l'écart-type de *ardeur* (0,83) se situe donc dans la moitié inférieure de l'échelle.

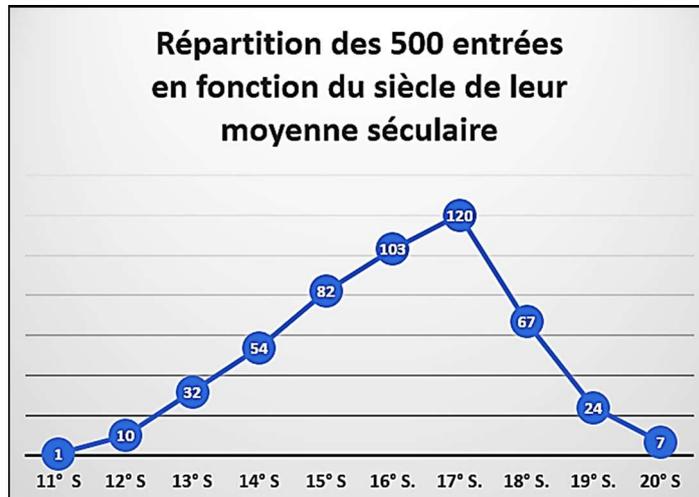
On trouvera sur le site du CRISCO ([URL](#)) pour chacune des 500 entrées deux tableaux mentionnant

- le **nombre de rubriques** (y compris les rubriques non datées), l'amplitude de l'entrée (le nombre de siècles couverts),
- la **moyenne séculaire pondérée** de l'entrée (c'est-à-dire sa localisation temporelle moyenne, quelle que soit la dissémination de ses rubriques)
- et l'**écart-type** (représentant le degré de dissémination entre la rubrique la plus précoce et la rubrique la plus tardive).

◆ **Classement des 500 entrées par leur localisation** (moyenne séculaire croissante)

Dans le premier tableau les lignes sont classées par ordre croissant de la moyenne séculaire (colonne colorée en ocre). Cette disposition permet d'enregistrer le nombre des entrées dont la moyenne séculaire se situe dans l'un des siècles représentés.

Figure 3



Les deux époques les plus riches en masses d'entrées sont le 16^e et le 17^e siècles avec plus de 100 entrées. Cependant ces chiffres doivent être pris avec précaution, d'abord parce que la sélection des 500 entrées est plus ou moins aléatoire, ensuite parce que le 16^e siècle est l'époque de développement rapide de l'imprimerie, des premiers grammairiens et lexicographes et le 17^e celui de l'émergence d'un français centralisé dont plusieurs dictionnaires enregistrent les vocabulaires.

◆ **Classement des 500 entrées par leur degré de dissémination** (écart-type croissant)

Ce second classement accorde la priorité à la croissance de l'écart-type (3^e colonne, sur fond ocre) de 0,00 à 3,72.

2.4. Illustration des entrées H-É présentant un profil extrême

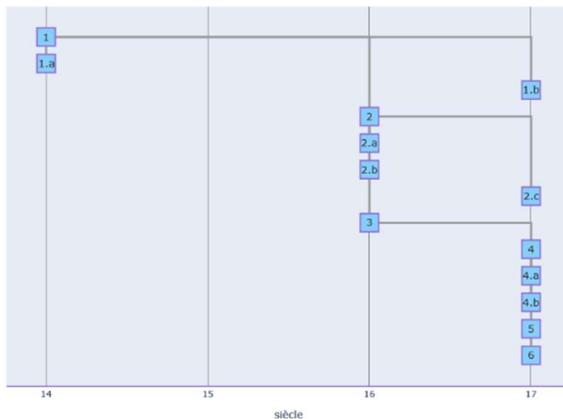
— *Nombre minimal de rubriques : avantage*



Niveaux	Siècle	Forme/colloc	Histoire
1	12		1160-74 avantage « ce qui produit une différence » (Wace, Rou, éd. H.
2	12		1160-74 « ce qui avance, ce que l'on a en plus d'autrui » (Wace, loc. cit,
3	12		ca 1175 « profit » (Chr. de Troyes, Chevalier Lion, éd. W. Foerster, 1321

— *Nombre maximal de rubriques : posséder*

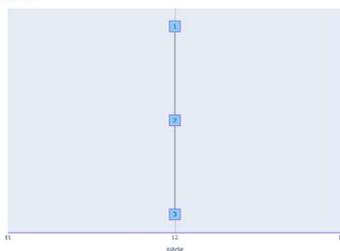
posséder



Niveaux	Siècle	Forme/colloc	Histoire
1.a	14		1364 «avoir la propriété de» (Arch. nat. ds Fagniez t.2, p.96);
1.b	17	possédante	1617 possédante «femme qui possède» (Crespin d'apr. FEW t.9, p.240) _1637 (ibid.); 1897 les possédants (Barrès, Déracinés, p.97);
2.a	16	posséder les reins de qq	1562 posséder les reins de qq «connaître et diriger les affections et pensées les plus secrètes (en parlant de Dieu qui dirige
2.b	16	possédé de maling esprit	1567 possédé de maling esprit (Amyot, Marcel., 31 ds Littre);
2.c	17	possédé	1666 possédé part. passé subst. (Boileau, Satires, IV, 81, éd. A. Cahen, p.68);
3	16		1580 (Montaigne, Essais, éd. P. Villey, I, chapitre 28: cette amitié qui possède l'âme);
4.a	17	posséder une femme	1637 posséder une femme «l'avoir pour épouse» (Corneille, Cid, III, 6);
4.b	17		1666 «avoir ses faveurs» (La Rochefoucauld, Maximes, éd. J. Truchet, p.301, var.);
5	17	se posséder	1652 se posséder «se dominer» (Corneille, Pertharite, III, 3);
6	17		1666 «connaître à fond» (Pascal, Pensées, éd. Lafuma, série III, fragment 427)

— *Amplitude minimale : célébrer*

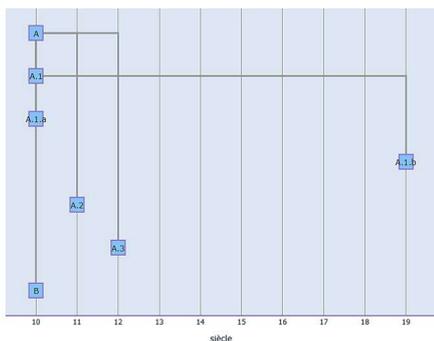
célébrer



Niveaux	Siècle	Forme/colloc	Histoire
1	12		Ca 1119 « accomplir (un rite religieux) solennellement » (Ph. de Thaon, Comput, 52 ds T.-L.); 1174 « id. (la messe) » (G. de Pont-
2	12		ca 1160 « marquer un événement par une cérémonie » (Eneas, éd. Salverda de Grave, 4645);
3	12		entre 1154-73 « honorer (qqn) » (B. de Ste-Maure, Troie, 4861 ds T.-L.)

— Amplitude maximale : **honneur**

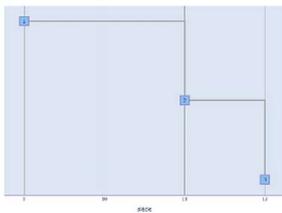
honneur



Niveaux	Siècle	Forme/colloc	Histoire
A.1.a	10		2 ^e moitié xes. « marque de vénération, de considération, d'honneur » (S. Léger, éd. J. Linskill, 2 [sing.], 7 [plur.]);
A.1.b	19	Votre Honneur	1835 Votre Honneur (en Angleterre) titre donné par respect à certaines pers. de qualité (Ac.);
A.2	11		mil. xies. « considération, estime, haut crédit dont on jouit » (Alexis, éd. Chr. Storey, 69); 1679 faire honneur à la lettre [de
A.3	12		ca 1100 « sentiment qu'on a de sa dignité » (Roland, éd. J. Bédier, 533); spéc. à propos d'une femme ca 1145 « sagesse du maintien, de
B	10		2 ^e moitié xes. « office, charge » (S. Léger, éd. citée, 120); ca 1100 « domaine, possession, fief » (Roland, éd. citée, 315).

— Localisation la plus précoce : **aimer**

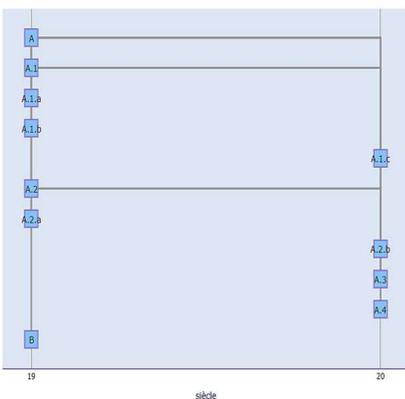
aimer



Niveaux	Siècle	Forme/colloc	Histoire
1	9		Fin ixes. « avoir du goût (pour qqc.) » (Eulalie, éd. Förster et Koschwitz, 10 : La polle sempre ñ amast lo dö menestrier);
2	11		mil. xies. « éprouver de l'amour (pour qqn, ici pour Dieu) » (Alexis, 50 e, éd. Paris : Plus aimet Deu que trestot son lignage);
3	12		1160 « attacher du prix (à qqc.) » (Roman de Troie, éd. Constans, 11226 ds T.-L. : jusqu'à poi, s'il n'a aïe, Porra petit preisier

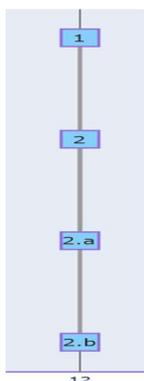
— Localisation la plus tardive : **performance**

performance



Niveaux	Siècle	Forme/colloc	Histoire
A.1.a	19		1839 plur. « résultats, actions accomplies par un cheval de course » (J. des Haras, XXIII, 191 ds Bonn., p. 105); 1872 sing. (Pearson,
A.1.b	19		1869 « manière de faire quelque chose, action accomplie (dans tout autre domaine) » (J. Faure, Des différences physiologiques et
A.1.c	20		1929 plur. « ensemble des possibilités optimales d'un appareil » (Guillemain, Constr., calcul et essais avions, p. 417);
A.2.a	19		1867 « résultat (sportif) remarquable, exploit » (Le Sport, 29 mai ds Petiot : Le turf a pris performance à la langue anglaise. Ce mot
A.2.b	20		1928 dans tout autre domaine (Dorgelès, Le Cabaret de la belle femme, p. 161 ds Rey-Gagnon Anglic.);
A.3	20		1953 psychol. (Delay, loc. cit.);
A.4	20		1966 ling. (N. Ruwet ds Langages, déc., pp. 6-7 d'apr. Rey-Gagnon Anglic.).
B	19		1869 « représentation » (Hugo, Homme qui rit, t. 2, p. 83)

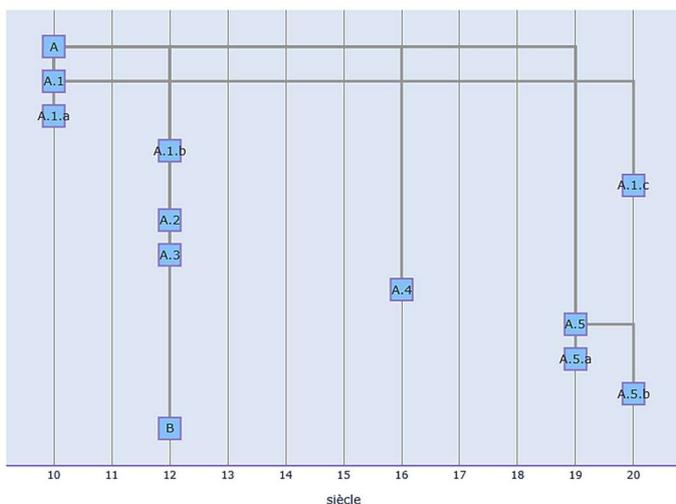
— *Dissémination minimale : fondement*



Niveaux	Siècle	Forme/colloc	Histoire
1	12	fundement	1119 fundement « anus » (Ph. de Thaon, Comput, 1740 ds T.-L.);
2.a	12		ca 1120 archit. (Psautier d'Oxford, 86, 1 ds T.-L.);
2.b	12		ca 1174-76 p. ext. « base, ce sur quoi repose quelque chose » (G. de

— *Dissémination maximale : défendre*

définir



Niveaux	Siècle	Forme/colloc	Histoire
A.1.a	10		Pronom. fin xes. « se protéger contre » (Passion, éd. d'Arco Silvio Avalle,
A.1.b	12		1180-85 pronom. « se protéger contre les effets nuisibles de quelque chose
A.1.c	20		1908 « résister aux effets du temps, de l'âge » supra, ex 11;
A.2	12		ca 1100 trans. « interdire à l'ennemi par des moyens militaires l'accès ou
A.3	12		mil. xiies. pronom. « nier ce dont on est accusé » (Lois de Guillaume le
A.4	16	définir qq en jugement	1559 trans. « prendre la défense, le parti de » défendre qq en jugement
A.5.a	19		1843 pronom. « (en parlant d'une chose) être valable, être justifié »
A.5.b	20		1920 fam. « (en parlant d'une pers.) se débrouiller habilement dans
B	12		Ca 1100 « enjoindre à quelqu'un de ne pas faire quelque chose » (Roland,

2.5. Conclusion

Dans son état actuel de finition, la version de démonstration du projet MGPE présente encore quelques bavures mineures que nous nous efforçons de traquer. Elles concernent

- a) Le **perfectionnement des tableaux de conversion graphique** : il ne porte que sur l'absence ou la présence erronée de formes de mot ou de collocations dans la colonne prévue à cette fin. Les formes de mot incongrues proviennent généralement de segments en italique situés dans l'étymologie, le procédé de séparation entre la partie historique et la partie étymologique (un double espace vide) n'étant pas absolument régularisé. La vérification de l'exactitude de cette troisième colonne est en cours.
- b) Le **perfectionnement des interfaces graphiques**. Il s'agit d'une part de nœuds auxiliaires omis, particulièrement quand deux nœuds auxiliaire successifs sont à insérer, par ex. quand la première rubrique d'une entrée est A.1.a et qu'il manque non seulement A, mais aussi A.1 avant A.1.a. D'autre part, la longueur de la partie verticale de certaines arêtes entre deux nœuds successifs (dans le même siècle ou entre deux siècles) varie d'une interface graphique à l'autre, ce qui oblige parfois l'internaute à dérouler longuement l'interface. Enfin le cheminement de certaines arêtes de l'arbre est difficile à suivre, mais c'est un biais inéluctable du mode de visualisation adopté.

En dépit de ces bavures mineures, la présentation associée du tableau de conversion et de l'interface graphique de chacune des 500 entrées peut être considérée comme **très satisfaisante** et à ce titre elle mérite d'être et déjà d'être diffusée non seulement sur le site du CRISCO mais aussi sur d'autres sites afin de recueillir des retours d'internautes intéressés par l'histoire du lexique français.

3. AU-DELÀ DE LA VERSION DE DÉMONSTRATION

3.1. Volume et distribution des entrées H-É polysémiques (Laurette Chardon et Justine Reynaud)

Selon l'état actuel des dénombrements de Laurette Chardon (mise à jour du 31 mai 2024), le nombre total d'entrées traitées s'élève à **49 862** sur les 49 876 des fichiers xml des entrées H-É exploités. Sur ce total, **20 314** entrées sont polysémiques (40,7%), c'est-à-dire comportent plus d'une rubrique.

Remarque sur les premières entrées H-É

Ces premières entrées (essentiellement **Ab...**) sont inexploitable, car les entrées H-É étaient conçues au tout premier temps selon une autre disposition, avec deux parties *Étymologie* et *Histoire* complètement séparées et une distinction supplémentaire dans cette seconde partie entre deux époques : avant et depuis la Révolution française. Plusieurs motifs ont probablement justifié l'abandon quasiment immédiat de cette disposition : un excès de détails, de nombreuses redondances entre les deux parties et une tentative avortée de datation du déclin des différents sens de chaque vocable.

◆ NOMBRE D'ENTRÉES EN FONCTION DU NOMBRE DE RUBRIQUES

L'étude macroscopique des 20 314 entrées polysémiques révèle une répartition intéressante de celles-ci. Le tableau 5 réunit les 20 premiers rangs par ordre de volume décroissant des rubriques, ceux des entrées les plus polysémiques. Comme on ne rencontre qu'une entrée à chaque rang, elles ont été distribuées entre quatre classes grammaticales.

Le nombre des entrées nominales (9) est inférieur à celui des entrées verbales (13), mais cela tient probablement au fait que Jacques François a exploré en priorité le lexique verbal par le passé (cf. François 2007). Plus intéressant est le constat du nombre restreint d'adjectifs (4), alors que les adjectifs plus courants sont souvent très polysémiques (ex. *clair, curieux, faible, fort, propre, sec, sensible*, etc.). Quant aux six « autres » morphèmes (conjonctions, déterminants, pronoms) ils figurent dans la sélection pour tester superficiellement le degré de polysémie des morphèmes grammaticaux.

Tableau 5 : Vingt premières entrées classées par rang (nombre de rubriques décroissant) et réparties par classes grammaticales

Rang	NbRu	NbEnt	N	V	Adj	autres
1	158	1		<i>tenir</i>		
2	93	1		<i>tomber1</i>		
3	89	1				<i>que</i>
4	84	1	<i>tête</i>			
5	75	1		<i>tourner</i>		
6	68	1				<i>tant</i>
7	62	2	<i>raison</i>			<i>rien</i>
8	61	1		<i>toucher1</i>		
9	60	1		<i>porter1</i>		
10	58	1	<i>vie</i>			
11	56	1	<i>volée</i>			
12	55	3	<i>oeil, trait</i>		<i>vieux</i>	
13	54	1		<i>rendre</i>		
14	52	1		<i>prendre</i>		
15	51	3		<i>monter</i>		<i>le1, sien</i>
16	50	2		<i>jeter</i>	<i>noir</i>	
17	49	3	<i>tambour</i>	<i>vivre1</i>		<i>si1</i>
18	48	3		<i>mettre</i>	<i>plein, tel</i>	
19	47	2	<i>part1, retour</i>			
20	46	2		<i>abattre, sauter</i>		
Nb total			9	13	4	6

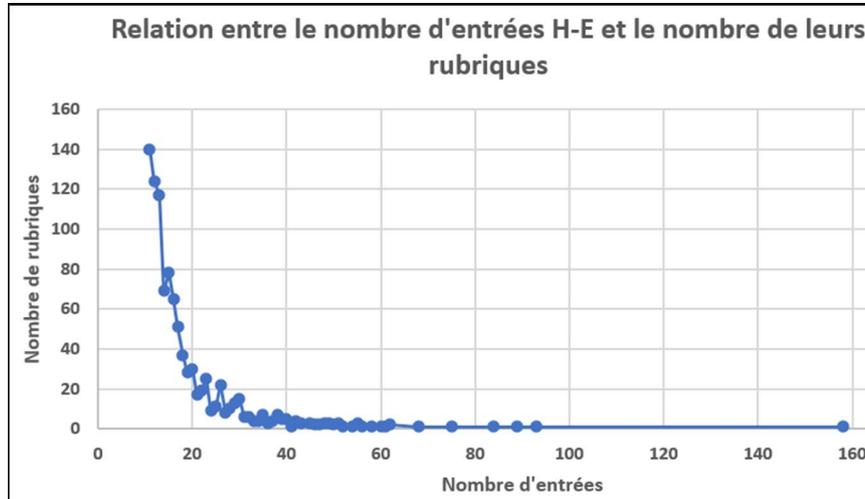
Dans le tableau 6 (rangs 21 à 54), le nombre des entrées concernées augmente graduellement, si bien qu'à partir du rang 42 (23 rubriques), seules les premières entrées sont mentionnées.

Tableau 6 : Rangs 21-54

21	45	3	<i>même, pied, vide</i>
22	43	3	<i>point1, retourner, voir</i>
23	42	4	<i>moins, ordre, retenir, touche</i>
24	41	1	<i>marquer</i>
25	40	5	<i>notre, nos, oreille, partie, rouler, vouloir1</i>
26	39	5	<i>ne, n', passer1, qui, soutenir, vous</i>
27	38	7	<i>maître2, maîtresse, réseau, rouge, suite, sur, tien, tienne, tirer</i>
28	37	4	<i>néant, revenir, rompre, train</i>
29	36	3	<i>jeu, remuer, sortir1</i>
30	35	7	<i>laisser1, net1, nette, piquer, porteur, -euse, remettre, retirer, voix</i>
31	34	4	<i>haut2, là, pour, sec, sèche</i>
32	33	4	<i>corps, long, longue, par1, voie</i>
33	32	6	<i>à, lieu1, mouvement, non, taper1, vert, verte</i>
34	31	6	<i>loin, présenter, tout1, toute, tous, toutes, trou, vérité, voler1</i>
35	30	15	<i>remonter, ni, mesure, juste, léger, travers, où, porte1, parler1, donner ...</i>
36	29	13	<i>reconnaître, jour, type, perdre, rentrer, tube, soi, maison, mort2 ...</i>
37	28	10	<i>mouche, mouiller, patte1, physique2, portée, suspendre, tampon ...</i>
38	27	8	<i>-ard, -arde, peu, rond1, ronde, séparer, service, souris2, suivre, sûr</i>
39	26	22	<i>partir2, terminer, si2, verser, profond, dire1, lever1, mais1, suspension, ...</i>
40	25	11	<i>simple, pousser, temps, sous, vue, mobile, volant1, le2, muet, aiguille, ...</i>
41	24	9	<i>abatage, abatage, ordinaire, raide, roide, rapport, refuser, rond2 ...</i>
42	23	25	<i>mourir, ligne, âme, or3, lâcher1, table, venir, révolution, réduire, ...</i>
43	22	19	<i>quartier1, ramasser1, unité, premier, charger, penser1, sécher, soleil ...</i>
44	21	17	<i>présence, verge, ton3, tour3, zone, nom, courir, main, moi, faire1, ...</i>
45	20	30	<i>rustique, pierre, sortie, saisir, vague1, aboyeur, mode1, parti, te, poser, ...</i>
46	19	28	<i>section, abordeur, plan3, degré, précieux, quel, subtil, sauver, souffle...</i>
47	18	37	<i>planche, portant, promettre, perte, livre1, sûreté, siffler, perle, loup, ...</i>
48	17	51	<i>gras, prétendre, union1, tablier, toit, montrer, rapide, régulier, pomme ...</i>
49	16	65	<i>sage, trancher, rouleau, poule1, unir, solide, siècle, têtard, station, ...</i>
50	15	78	<i>rampe, soutien, gris1, minute1, provision, tringle, ongle, scène, robe ...</i>
51	14	69	<i>rare, violet, orienter, sacrer1, image, petit, abaque, nettoyer, ...</i>
52	13	117	<i>appuyer, mouillage, tailler, négatif, transmission, suffire, sale ...</i>
53	12	124	<i>maîtrise, souffleur, sauce, prolongement, couvrir1, malle, sonder ...</i>
54	11	140	<i>reprise1, traîne, rive, fournir, manteau, transversal, agripper, clef ...</i>

Le graphique en figure 5 montre que la relation entre le nombre d'entrées H-É et le nombre de leurs rubriques se conforme globalement à une distribution de Zipf.

Figure 5



◆ RÉPARTITION CHRONOLOGIQUE DES RUBRIQUES DES DIX ENTRÉES LES PLUS POLYSÉMIQUES

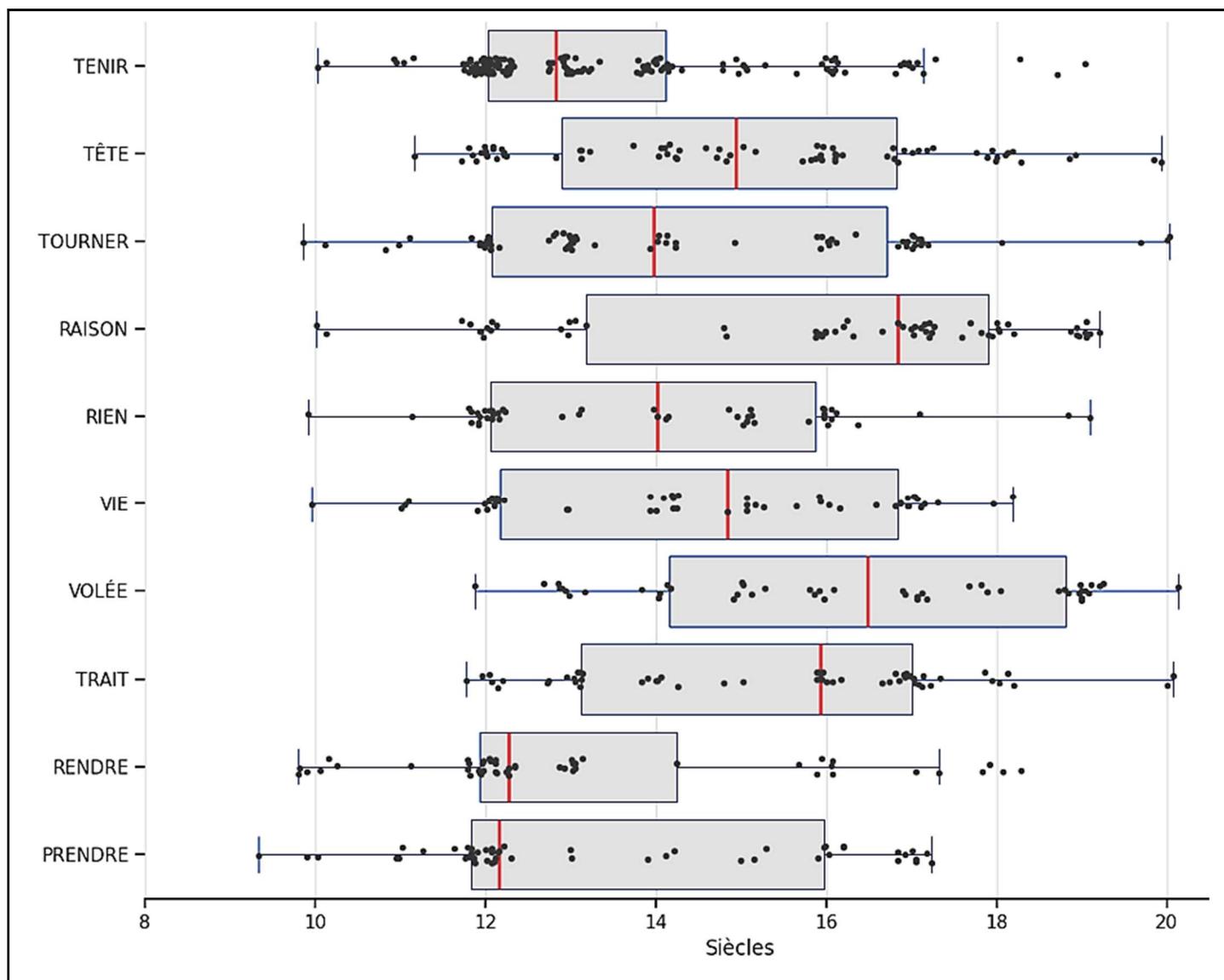
La figure 6 (p. 25) représente les *boîtes à moustaches* des dix entrées les plus polysémiques. Pour chaque entrée, un point représente une rubrique (pour des questions de lisibilité, les points sont légèrement dispersés autour de leur abscisse et ordonnée).

La barre rouge représente la **médiane** : une moitié des rubriques est apparue avant et l'autre moitié après le siècle indiqué par cette barre. Le rectangle gris indique les premier et troisième quartiles⁶, c'est-à-dire qu'il englobe la moitié des points (un quart de chaque côté de la médiane) et donne une idée de l'amplitude des données : plus le rectangle est étendu, plus l'**amplitude** est importante.

La position de la médiane dans le rectangle fournit une information sur la répartition des points : si la médiane est au centre du rectangle, les rubriques sont réparties de manière à peu près homogène avant et après celle-ci. C'est par exemple le cas du n.f. *vie*. Lorsque la médiane est proche de l'un des bords du rectangle, cela signifie que de nombreuses rubriques ont émergé à cette époque. Dans le cas du v. *prendre* par exemple, on observe de nombreuses apparitions au 12^e siècle. Le poids remarquable de ce 12^e siècle tient probablement à son statut fréquent de tête de collocation verbale, ex. *prendre en mains* (1160-74) ; *prendre essample de qqn* (ca 1180).

⁶ Définition du Robert : « Chacune des trois valeurs qui divisent les éléments d'une distribution statistique en quatre parties d'effectif égal ».

Figure 6 : Répartition chronologique des rubriques de dix des entrées les plus polysémiques



◆ **NOMBRE D'ENTRÉES EN FONCTION DU NOMBRE DE NIVEAUX**

Un autre type d'exploration macroscopique de la microstructure des entrées H-É porte sur les pratiques des rédacteurs de ces entrées en matière de hiérarchisation des rubriques. Il se peut que certains d'entre eux aient souhaité dégager un petit nombre de groupes de rubriques moyennement apparentées (entraînant une hiérarchie limitée de classes copieuses), tandis que d'autres ont préféré dégager un nombre élevé de petits groupes de rubriques étroitement apparentées (avec pour effet une hiérarchie des classes plus élevée).

Premier constat dressé dans le tableau 6 : la très grande majorité des entrées recourent à un seul niveau de hiérarchie des rubriques, bien entendu les 29 562 entrées monosémiques, mais également 12 811 entrées polysémiques sur 20 314, soit 63%. Il s'agit manifestement des entrées faiblement polysémiques, comptant deux ou trois rubriques.

Tableau 6 : Répartition de la totalité des entrées en fonction du nombre de niveaux hiérarchiques de leurs rubriques

Nombre de niveaux	Nombre d'entrées	Quelques exemples d'entrées
1	42373	<i>ursidés, guil, simoun, fureter, stréphopode, hormonothérapie, répréhension, hideux, sarcastique, poigne, ...</i>
2	5857	<i>digérer, spiritualité, vase2, barnache, équivoque, brioche, physiologiste, saisonnier, stepper2, gratte, ...</i>
3	1305	<i>second, filer, tombeur, taupe, sexuel, latin, garantie, souverain1, vésuvien, naturel, ...</i>
4	266	<i>jouer, siège, pied, plafond, tas, raison, sauver, nègre, vôtre, pavillon, ...</i>
5	53	<i>néant, rang, mouton, outre2, partie, tel, sien, tant, période, touche, ...</i>

Le tableau 7 complète le précédent en indiquant pour dix des entrées les plus polysémiques (avec plus de 50 rubriques) le nombre et la nature des niveaux hiérarchiques impliqués. Il en ressort que ces entrées tirent essentiellement parti des trois niveaux médians (A - 1 - a), deux y ajoutent le niveau le plus élevé, huit recourent aussi au niveau le plus bas (Z - 1 - a - α) et deux entrées exploitent les cinq niveaux (I - A - 1 - a - α).

Le constat le plus surprenant concerne la microstructure de l'entrée *tenir*, la plus copieuse avec 158 rubriques (et beaucoup plus de nœuds dans l'interface graphique en raison de l'ajout indispensable de nombreux nœuds auxiliaires) dont le rédacteur a réussi le tour de force de limiter la hiérarchie de ses rubriques à trois niveaux.

Tableau 7 : Les dix entrées les plus polysémiques et la hiérarchie de leurs rubriques

<i>entrée</i>	NbRub	l	A	1	a	α
<i>tenir</i>	158					
<i>tête</i>	84					
<i>tourner</i>	75					
<i>raison</i>	62					
<i>rien</i>	62					
<i>vie</i>	58					
<i>volée</i>	56					
<i>trait</i>	55					
<i>rendre</i>	54					
<i>prendre</i>	52					
Total		2	10	10	10	8

◆ **ENTRÉES MONOSÉMIQUES VS POLYSÉMIQUES ET DISPOSITION CHRONOLOGIQUE DES RUBRIQUES DE CES DERNIÈRES**

Le premier histogramme de la page suivante représente le nombre d'entrées monosémiques vs polysémiques par siècle. On observe un accroissement impressionnant du nombre et de la proportion des entrées polysémiques au 12^e siècle (près de dix fois supérieure à celle des entrées monosémiques). Ce nombre diminue graduellement jusqu'au 15^e siècle avant un nouvel accroissement au-delà de 10 000 entrées par siècle aux

16^e et 17^e. Enfin le 19^e siècle enclenche un troisième accroissement avec 8 000 nouvelles entrées monosémiques.

Ce profil a certainement des causes diverses :

- pour le 12^e siècle, c'est l'abondance de la **littérature** conservée et enregistrée par les philologues depuis le 18^e siècle,
- pour le 16^e, c'est la multiplication des textes publiés à bas coût avec l'invention de l'**imprimerie**, laquelle a permis à une couche plus étendue de la population lettrée de les acquérir.
- Pour le 17^e, c'est le développement rapide des **dictionnaires** (lui-même associé au recul des dialectes et à l'épanouissement d'une langue nationale unifiée) qui enregistrent les sens nouveaux de mots préexistants.
- Et pour le 19^e, c'est le développement des **écrits scientifiques** : de nombreux termes nouveaux voient le jour, mais pour alléger l'effort de mémorisation d'un lexique surabondant, on recourt fréquemment à la **néologie sémantique** quand le risque d'ambiguïté est limité (ex. **chambre**, apparu au milieu du 11^e siècle et donnant lieu au 19^e à *symphonie de chambre* (1814), *chambre chaude* (chimie, 186), *chambre pollinique* (fin 19^e), *chambre à air* (1891).

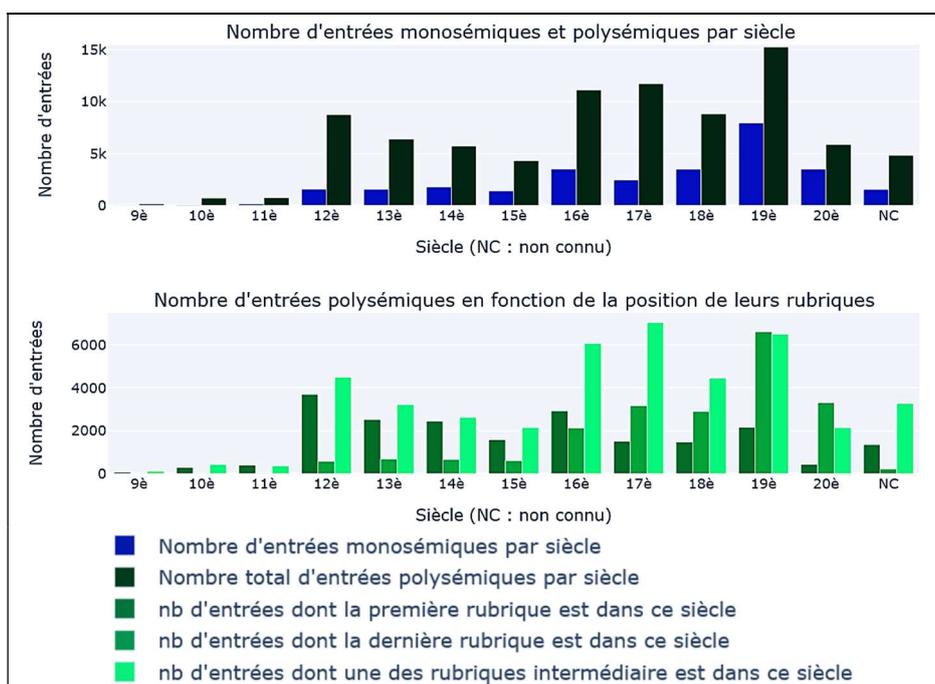
Le second graphique est d'une lecture plus délicate. Dans chaque siècle la première colonne représente le nombre d'entrées dont la 1^{ère} rubrique figure dans ce siècle, la seconde le même profil pour la dernière rubrique et la troisième le nombre d'entrées dont (au moins) une des rubriques intermédiaires figure dans ce siècle.

L'évolution comparée de la 1^{ère} et de la 2^{ème} colonne diffère de celle du graphique précédent essentiellement du 17^e au 20^e siècle : la proportion de sens nouveaux, supérieure jusque là à celle des sens anciens, s'inverse au 17^e siècle, avec un déclin marqué aux 19^e et 20^e siècles. Cela signifie que l'**âge d'or de la néologie sémantique** s'achève au 16^e siècle, même si l'introduction de nouveaux sens se poursuit au-delà.

Quant à l'évolution de la 3^{ème} colonne (nombre de rubriques figurant dans le siècle), elle s'inverse au 19^e siècle : jusque là elle est toujours supérieure à celle de la 2^{ème} (nombre de dernières rubriques localisées dans le siècle), mais au 19^{ème} et au 20^{ème} siècle elle devient légèrement, puis

largement inférieure. Cela signifie que, globalement, le gros de la polysémie évolutive s'est constitué **entre le 12^e et le 18^e siècle**.

Figure 7 : Répartition par siècle des entrées historiques mono- et polysémiques



Au-delà, c'est la fonction des grands dictionnaires du 19^e siècle (dictionnaires de langue comme le Bescherelle, le Littré et le Dictionnaire national de Hatzfeld et Darmesteter, et dictionnaires encyclopédiques comme le Larousse du 19^e siècle, important pour l'enregistrement des vocabulaires techniques nouveaux) d'enregistrer et de **fixer les acquis lexicologiques** des siècles précédents.

Les sections 3.2-3.4 sont consacrées à la présentation du tableau de conversion et de l'interface graphique de deux entrées représentatives de la classe des entrées très polysémiques, celle du n.m. *corps* (33 rubriques, 36 nœuds), et celle de l'adj. et adv. *même* (46 rubriques, 46 nœuds).

◆ **ÉTAT D'AVANCEMENT ACTUEL DE L'ÉLABORATION DES PAIRES** {tableau de conversion + interface graphique}

On peut actuellement distinguer trois classes de résultats :

- a) Les entrées que nous n'avons pas l'intention de traiter
- b) Les entrées pour lesquelles nous pouvons d'ores et déjà fournir une paire {tableau de conversion + interface graphique}
- c) Les **10 242 entrées en attente**, mais dont le traitement ne pose pas de problème particulier en dehors du contenu de la colonne « forme de mot / collocation »⁷.

Tableau 7 : Proportion d'entrées polysémiques traitées

NOMBRE DE RUBRIQUES	PAIRES TRAITÉES	REMARQUES
Entrées <i>Ab-</i>	non traitées	Composition différente des entrées historico-étymologiques
2	non traitées	11 731 entrées comportant 2 rubriques, d'un intérêt négligeable pour l'étude de la polysémie évolutive
3-10	507	10 242 entrées en attente
11	140	Inventaire complet
12-16	453	
17-20	146	
21-154	243	
	1489	Total des entrées traitées

⁷ La difficulté à surmonter prochainement est l'introduction dans cette colonne de tous les segments en caractères italiques qui figurent à la « bonne place », notamment ni dans des passages entre parenthèses courbes ou droites, ni dans la composante d'étymologie. C'est cette dernière difficulté qui est la plus délicate, car **LES RÉDACTEURS DES ENTRÉES HISTORIQUES N'ONT PAS SÉPARÉ DE MANIÈRE HOMOGENE LES DEUX PARTIES HISTORIQUE ET ÉTYMOLOGIQUE**. Dans la majorité des cas, l'étymologie figure après deux espaces blancs, mais cette convention (que nous avons repérée par hasard) n'est pas appliquée régulièrement.

3.2. La polysémie cumulative du n.m. *corps*

Le tableau de conversion graphique de l'entrée *corps* compte 33 lignes. L'interface graphique compte 33 nœuds substantiels auxquels s'ajoutent trois nœuds auxiliaires. Cela signifie que la grande majorité des rubriques de 1^{er} rang de cette entrée sont renseignées (une rubrique renseignée fournit canoniquement un nœud de l'arbre de distribution des sens. L'entrée H-É du n.m. *corps* a la microstructure suivante :

Tableau 8 : Microstructure de l'entrée H-É du n.m. *corps*

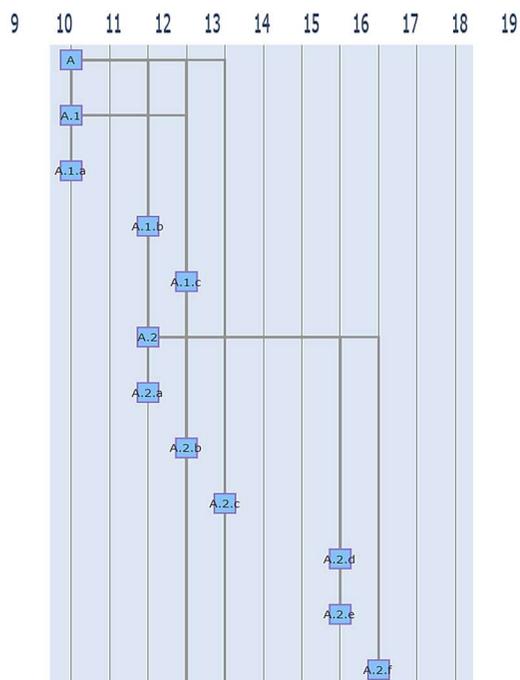
1 ^{er} rang	A Partie matérielle des êtres animés				B Partie principale				C Objet matériel					D Groupe (de personnes, de choses)					E Consistance	
	A.1	A.2	A.3	A.4	B.1	B.2	B.3	B.4	C.1	C.2	C.3	C.4	C.5	D.1	D.2	D.3	D.4	D.5	E.1	E.2
2 ^e rang	organisme humain / cadavre	personne / individu	tronc du corps	partie d'un vêtement couvrant le buste	CORPS DE LA CITÉ	CORPS DE LA MAISON	CORPS D'UNE LETTRE	CORPS DE DÉLIT	CORPS CÉLESTE	PETITZ CORPS =atomes	CORPS ESTRANGE	tout objet matériel	CORPS SIMPLE (chimie)	CORPS DE LOIS	CORS DE VILLE	groupe de soldats	CORPS DES ARTISANS,	CORPS DE BALLET	d'une teinture [qui a du corps]	DONNER DU CORPS (à une idée, etc.)
3 ^e rang	A.1.a A.1.b A.1.c	A.2.a A.2.b A.2.c A.2.d A.2.e A.2.f												D.2.a D.2.b						

Pour gagner de la place et éviter les redondances, les rédacteurs des notices H-É n'ont pas introduit une définition dans chaque rubrique. Les cellules surlignées en gris ont une définition, les cellules surlignées en rose en sont démunies.

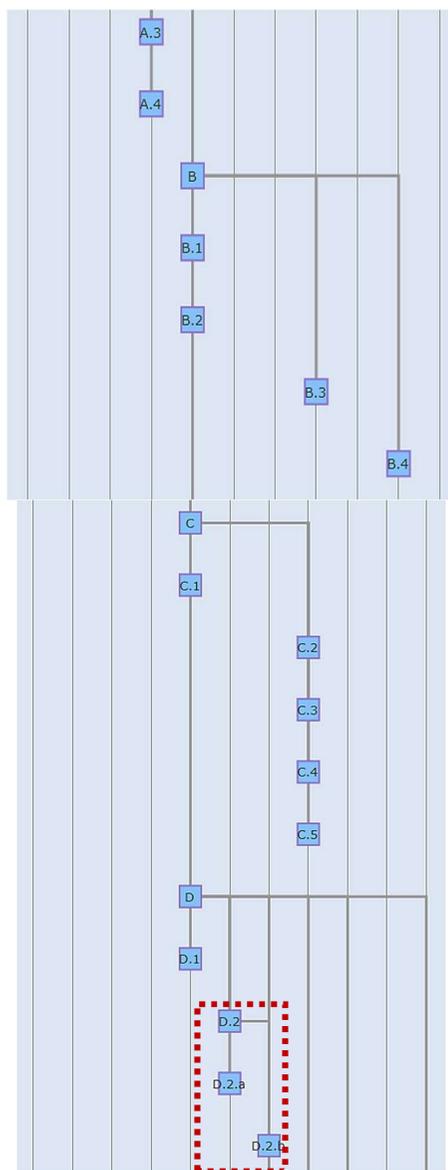
Au rang supérieur (A-E) le contenu de chaque rubrique se limite fréquemment à une définition générique (sans guillemets). Au second rang (A.1-E.2), six rubriques portent une définition contre 14 qui n'en portent pas. Le sens véhiculé par les rubriques dénuées de définition est à rechercher en général soit dans la rubrique immédiatement antérieure de même niveau, soit dans la première rubrique subordonnée.

Dans le groupe A, les rubriques A.3-4 dénuées de rubriques subordonnées sont définies, tandis que les rubriques A.1-2 dotées de rubriques subordonnées ne le sont pas. Dans cette configuration, on peut se reporter respectivement aux définitions des sous-rubriques A.1.a et A.2.a. Parmi les rubriques de 3^e rang de ce groupe, A.2.c et A.2.f ne sont pas définies. Par défaut les sens de A.2.c et A.2.f sont respectivement similaires à ceux de A.2.b et A.2.f. En outre une indication de domaine thématique ou de relation rhétorique peut se substituer à une définition.

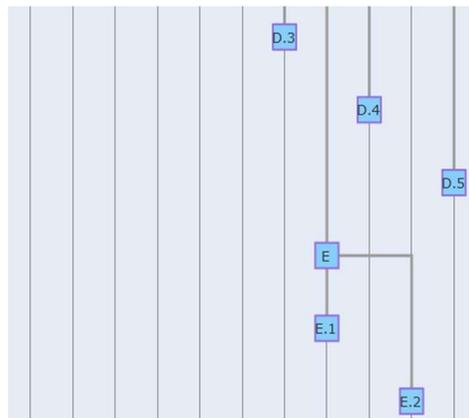
Ainsi B.3 comporte l'indication « typogr. », d'où l'on peut déduire que le corps d'un caractère est au domaine de la typographie ce que le corps de la maison est au domaine de l'architecture. Le même raisonnement vaut pour C.3 comportant l'indication « chim. » : ici un « corps estrange » (C.3) est au domaine de la chimie ce qu'un « petit corps » (C.2) est à celui de la physique. Et E.2 porte la mention rhétorique « fig. », d'où il s'ensuit que E.2 est un emploi figuré (abstrait) de E.1 dans le domaine des teintures : le styliste donne du corps à une idée, comme le teinturier donne du corps à la teinte d'un tissu.



Niveaux	Siècle	Forme/colloc	Histoire
A			« Partie matérielle des êtres animés »
A.1.a	9		ca 881 corps « l'organisme humain (p. oppos. à l'âme, à l'esprit) » (Eulalie, 2 ds Henry Chrestomathie, p. 3); 1remoitié du xiies. cors
A.1.b	11	cors	ca 1050 cors « cadavre » (Alexis, éd. Ch. Storey, 583);
A.1.c	12	le cors Nostre Seinur	ca 1160 le cors Nostre Seinur « l'Eucharistie » (Wace, Rou, éd. A. J. Holden, t. 1, p. 188, 753); 1524 corps glorieux (Gringore, Le
A.2.a	11		ca 1050 corps « personne, individu » (Alexis, éd. Ch. Storey, 399) _ ca 1465, Maistre Pierre Pathelin, éd. R. T. Holbrook, 186;
A.2.b	12	sor lor cors deffendant	fin xiies. sor lor cors deffendant « en se défendant » (Conquête de Jérusalem, éd. C. Hippeau, 141); ca 1220 seur son cors deffendant «
A.2.c	13		1283 prise de (leur) cors (Ph. de Beaumanoir, Coutumes de Beauvaisis, éd. A. Salmon, t. 2, p. 269, 1522);
A.2.d	16	à corps perdu	1580 à corps perdu (R. Garnier, Antigone, 1151 ds Tragédies, éd. W. Foerster, t. 3, p. 41);
A.2.e	16	garde du corps	1549 garde du corps « ensemble de personnes chargées de la garde d'un souverain » (Est.) _ 1671, Pomey ds FEW t. 17, p. 516 b; 1567
A.2.f	17	drôle de corps	1672 drôle de corps (Montfleury, La Fille capitaine, V, 3 ds Livet Molière, t. 2, s.v. drôle);



A.3	12	cors	début xiies. cors « tronc du corps » (Roland, éd. J. Bédier, 1586);
A.4	12	cors	ca 1170 cors « partie d'un vêtement qui couvre le buste » (Wace, Rou, éd. A. J. Holden, t. 2, 8120); 1575 corps de robbe (Inventaire
B			« Partie principale »
B.1	13	cors de la cité	xiiies. [ms.] cors de la cité « partie principale de la ville » (G. de Villehardouin, Conquête de Constantinople, éd. E. Faral, 387
B.2	13	cors de la maison	xiiies. cors de la maison [lat. : ædem] (Macchabées, éd. P. Meyer ds T.-L.); 1590 cors de logis (Comptes manoir Rouen, 500 ds IGLF);
B.3	16	corps d'une lettre	1528 typogr. corps d'une lettre (Inventaire de Louis Royer ds Mém. de la Soc. de l'hist. de Paris et de l'Île-de-France, t. 21, 1894,
B.4	18	corps de délit	1754 corps de délit (Encyclop. t. 4, p. 267 a); 1824 corps du délit (Balzac, Annette, t. 4, p. 65).
C			« Objet matériel »
C.1	13	cors celestre	ca 1270 cors celestre (J. de Meung, Rose, éd. E. Langlois, 17094);
C.2	16	petitz corps	1552 petitz corps « atomes » (Ronsard, Amours, éd. P. Laumonier, t. 4, p. 40);
C.3	16	corps estrange	1561 anat. corps estrange (A. Paré, Œuvres, éd. J.-F. Malgaigne, t. 2, p. 76 a); 1680 corps étranger (Rich., s.v. étranger);
C.4	16		1580 corps « tout objet matériel » (B. Palissy, Discours admirables, p. 430 ds IGLF);
C.5	16	corps simple	1585 chim. corps simple (N. Du Fail, Contes et discours d'Eutrapel, éd. J. Assézat, t. 1, p. 274).
D			« Groupe (de personnes, de choses) »
D.1	13	corps de lois	fin xiiies. corps de lois (Couronnement Renart, éd. Méon, 2528 ds T.-L.), attest. isolée; 1671 (Pomey);
D.2.a	14	cors de ville	1304 cors de ville (Roisin, éd. Brun-Lavainne, p. 345);
D.2.b	15		1434 corps « ensemble organisé de personnes (en gén.) » (Journal de Clément de Fauquembergue, éd. A. Tuetey, t. 3, p. 142); 1585 corps



D.3	15		1469 corps « groupe de soldats » (Lettre de Louis XI ds Bartzsch); av. 1662 corps d'armée (Pascal, Proph. 26 ds Littré); 1579 corps de
D.4	17		1607 corps des artisans, etc. (Hulsius); 1771 esprit de corps (Turgot, Œuvres, éd. G. Schelle, t. 3, p. 521);
D.5	19		1835 corps de ballet (Ac.).
E			« Consistance »
E.1	16		1580 d'une teinture (B. Palissy, Discours admirables, éd. A. France, p. 460 ds IGLF : les teintures sont toutes diaphanes,
E.2	18	donner du corps	av. 1715 fig. donner du corps (à une idée, etc.) (Fénelon ds Guérin)

Remarque 1 sur le tableau de conversion ► Dans l'état actuel de vérification du contenu de la colonne « Forme de mot / collocation », certaines corrections sont en attente : Dans les lignes A et A.2.a, il manque *corps*, en A2c il manque *prise de (leur corps)* et dans les lignes D.3/4/5 il manquemanque respectivement *corps* ; *corps des artisans, etc.* et *corps de ballet/*

Remarque 2 sur l'interface graphique ► Le mode élémentaire de visualisation des branches de l'arbre adopté pour les vocables étudiés ici entraîne fréquemment des interrogations sur la superposition

de deux ou plusieurs segments d'arêtes. À titre d'exemple, le nœud D donne lieu à six arêtes en aval, au 13^e s. vers D.1, au 14^e vers D.2 puis D.2.a, au 15^e s. vers D.3, au 16^e s. vers E, puis E.1, au 17^e s. vers D.4 et au 19^e s. vers D.5. Mais la zone entourée d'un rectangle rouge discontinu (page précédente) est peu lisible, car les deux arêtes D-D.3 et D.2-D.2.b ont un segment superposé. Il faut comprendre que l'arête D-D.3 vient de plus haut et se poursuit plus bas que D.2.b. **NOUS TESTONS ACTUELLEMENT LA POSSIBILITÉ DE COLORIER DIFFÉREMMENT LES NŒUDS DES DIFFÉRENTES BRANCHES AFIN DE LES DISTINGUER VISUELLEMENT.**

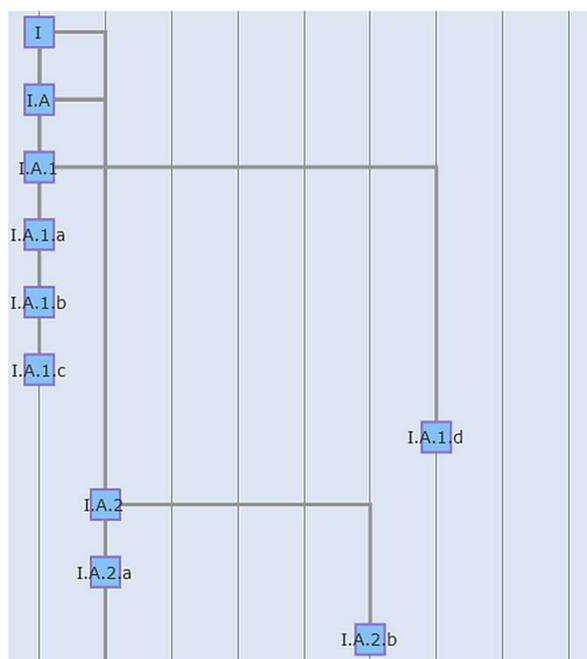
3.3. La polysémie cumulative de l'adj. et adv. *même*

La polysémie évolutive de *même* se décline sur 46 lignes dans le tableau de conversion graphique de cet adjectif et adverbe et l'interface graphique compte le même nombre de nœuds. Cela signifie que tous les nœuds sont substantiels et qu'aucun nœud auxiliaire n'a été ajouté. Cette configuration assez exceptionnelle tient au fait que tous les nœuds superordonnés (de premier et de second rang) sont renseignés.

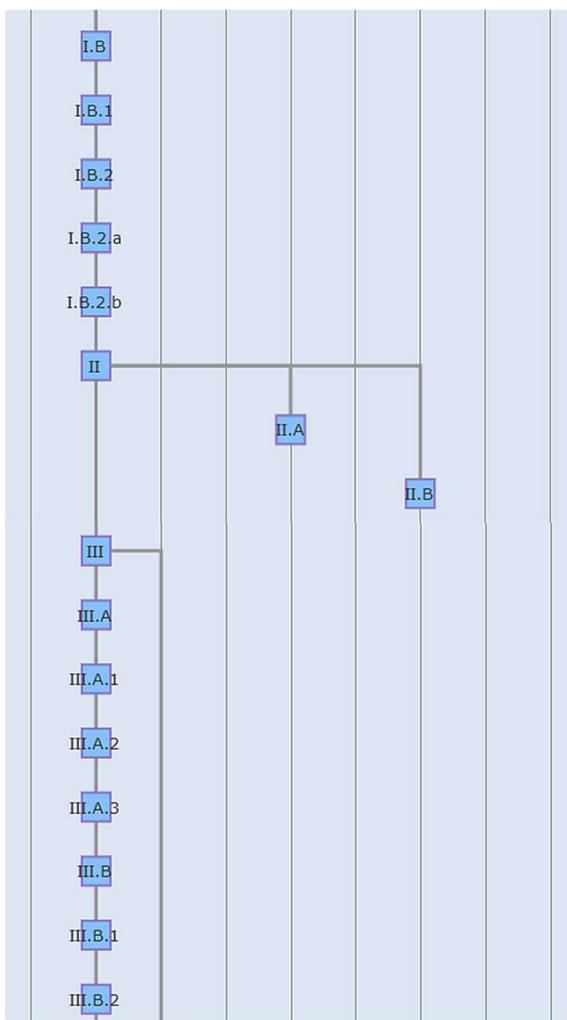
Par exemple la rubrique de 1^{er} rang I est motivée morphologiquement « Adj. d.fr. *meïsmes*, cas régime *meïsme*, la sous-rubrique de 2^e rang I.A. est motivée sémantiquement : « Marque l'insistance, indiquant qu'il

s'agit exactement de la personne ou de la chose en question, sens du lat. *ipse* ». Et la rubrique de 3^e rang I.A.1 est motivée sur le plan syntagmatique : « placé derrière » (vs I.A.2 : « placé devant »).

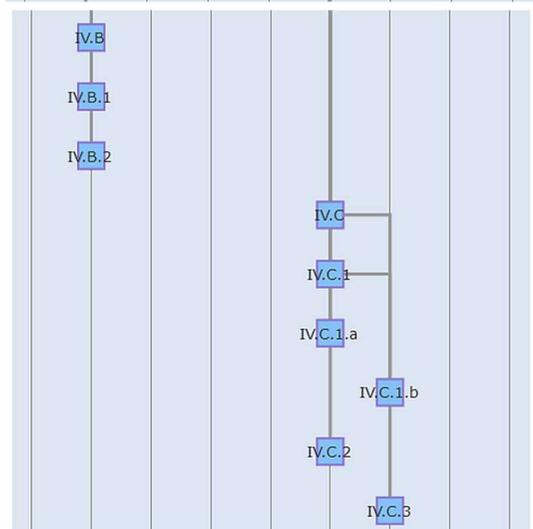
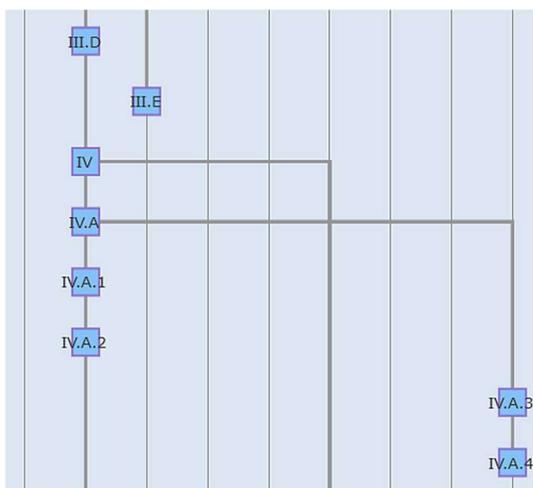
La rubrique II a également une motivation morphologique : « Pronom indéfini indiquant l'identité, la ressemblance : le même ». La rubrique III est motivée généalogiquement par équivalence sémantique aux pronoms latins *ipse* (« lui-même », déjà invoqué en I.A) et *id* (« celui-ci », III.1 vs III.B). Enfin la rubrique IV, consacré aux emplois en locution, se subdivise entre locutions adverbiales (IV.A) et prépositionnelles (IV.B).



Niveaux	Siècle	Forme/colloc	Histoire
I			Adj. a. fr. <i>meïsmes</i> , cas régime <i>meïsme</i> .
I.A			Marque l'insistance, indiquant qu'il s'agit exactement de la personne ou de la chose en question; sens du lat. <i>ipse</i>
I.A.1			placé derrière
I.A.1.a	11		ca 1050 un pron. pers. (St Alexis, éd. Chr. Storey, 118: A lui medisme; 284: de sei medisme); ca 1100 (Roland, éd. J. Bédier,
I.A.1.b	11		ca 1050 un subst. désignant une pers. (St Alexis, 539: Co lur est vis que tengent Deu medisme); ca 1100 (Roland, 400: L'emperere
I.A.1.c	11		ca 1050 un pron. poss. (Alexis, 432: A grant duel met la sue carn medisme);
I.A.1.d	17		1651 un subst. mis en appos. exprimant une qualité, indique que le déterminé possède cette qualité au plus haut point (Scarron, Roman
I.A.2			placé devant
I.A.2.a	12		1121-34 un subst. désignant une pers. (Philippe de Thaon, Bestiaire, 2526 ds T.-L.: Saint Esperit Qu'en meïsme Dé vit);
I.A.2.b	16		1578 un subst. mis en appos. et exprimant une qualité, indique que le déterminé possède cette qualité au plus haut point (Ronsard,



I.B			Marque l'identité absolue; sens du lat. idem.
I.B.1	12		placé derrière le subst. ca 1100 (Roland, 204: Nuncerent voz cez paroles meïsmes); ca 1170 (Marie de France, Lais, éd. J. Rychner,
I.B.2			placé devant le subst.
I.B.2.a	12		ca 1165 précède le déterm. (Benoît de Ste-Maure, Troie, éd. L. Constans, 19799: Enz en meïsmes la semaine); ca 1170 (Rois, éd. E.
I.B.2.b	12		ca 1165 placé entre le déterm. et le subst. (Guillaume d'Angleterre, éd. M. Wilmotte, 2186: tote une meïsmes cose).
II			Pron. indéf. indiquant l'identité, la ressemblance: le même.
II.A	15	faire le mesme	Empl. comme subst. av. 1473 faire le mesme «faire la même chose» (Juv. des Ursins, Charles VI, 1382 ds Littré); 1541 dire le mesme
II.B	17		1643 en fonction d'attribut (Corneille, Polyeucte, I, 1).
III		meïsmes, -s, -s	Adv. meïsmes [pourvu de l'-s adv., ne se distingue morphologiquement de l'adj. que dans les cas où ce dernier est
III.A		ipse	Avec les valeurs du lat. ipse «même, précisément, justement; [mon, ton, son] propre». Placé derrière
III.A.1	12		ca 1100 le subst. lui-même précédé d'un poss. que l'adv. renforce (Roland, 2551: Sun cors meïsmes i asalt e requert); ca 1165
III.A.2	12		ca 1170 le pron. pers. (Marie de France, Lais, Fresne, 72: A sei meïsmes se desmente; 470: Vers mei meïsmes meserrai);
III.A.3	12		ca 1170 un adv. de lieu, marque la précision (Chrétien de Troyes, Erec, éd. M. Roques, 4188: Illuec meïsmes an la place); 1176-80
III.B			Avec la valeur du lat. id.
III.B.1	12		placé derrière le subst. ca 1170 (Marie de France, Lais, Fresne, 66: En l'an meïsmes); 1174-76 (Guernes de Pont-Ste-Maxence, St
III.B.2	12		placé devant le subst. précédé de l'art. ca 1170 (Marie de France, Lais, Lanval, 219: meïsmes l'an Après la feste Seint Johan



III.D	12		Marque un renchérissement, une gradation, une mise en relief par rapport à un ensemble ca 1170 (Bérout, Tristan, éd. E.Muret,
III.E	13		Au sens de «particulièrement, surtout» 2emoitié XIIIes. (Lég. de G. de Roussillon, 39 ds T.-L.).
IV			Loc.
IV.A			Loc. adv.
IV.A.1	12	de meïsame	ca 1100 de meïsame «aussi, de la même manière» (Roland, 592);
IV.A.2	12	a meïsmes	ca 1160 a meïsmes «à proximité, dans le voisinage» (Eneas, éd. J. Salverda de Grave, 2423);
IV.A.3	19	quand même	1839 quand même «cependant, néanmoins» (Stendhal, Chartreuse de Parme ds Romans et nouvelles, éd. H. Martineau, t. 2, p.407: Si je
IV.A.4	19	tout de même	1831 tout de même «en dépit de tout, néanmoins» (Sue, loc. cit.).
IV.B			Loc. prép.
IV.B.1	12	a meïsmes de	ca 1160 a meïsmes de «près de» (Eneas, 928: A meïsmes de la cité); ca 1195 a meimes + subst. (Ambroise, Guerre sainte, éd. G. Paris,
IV.B.2	12	estre a meïsmes de	1176-81 estre a meïsmes de + inf., pron. ou adv. pronom. (Chrétien de Troyes, Charrette, éd. M. Roques, 6727); 1306 (Guillaume
IV.C			Loc. conj.
IV.C.1.a	16	de mesmes a	av. 1514 de mesmes a «de même que, comme» (Lemaire de Belges, Couronne Margaritique, IV, 63 ds Hug.);
IV.C.1.b	17	tout de mesme que	1606 tout de mesme que (Nicot);
IV.C.2	16	mesmes que	1583-90 mesmes que «et même, bien plus» (Brantôme, Cap. estr., Dom Pedro de Tolledo, II, 21 ds Hug.), devenu pop. en fr. mod.;
IV.C.3	17	quand même	1671 quand même introduisant une prop. concessive au cond. (Pomey: Quand même je le voudrois, je ne le pourroie pas)

4. CONCLUSION GÉNÉRALE

Nous avons formulé précédemment nos conclusions sur l'élaboration de la version de démonstration du projet MGPE (► section 2.5, p. 20).

Compte tenu de la **disponibilité** dans un format provisoire (en cours de vérification) de la quasi-totalité **des tableaux de conversion et des interfaces graphiques des ±20 000 entrées H-É** du TLFi, nous pouvons conclure – au-delà de la version de démonstration – à la robustesse de nos résultats d'un double point de vue :

- a) celui de l'élaboration d'une **théorie** de la polysémie évolutive (ou cumulative, voir la précision terminologique p. 4)
- b) celui du retraitement de **données sous-exploitées** (les entrées H-É des articles du TLFi, dotées d'une annotation seulement formelle, vis-à-vis de l'annotation fonctionnelle des entrées lexicographiques), afin de permettre à terme des recherches transversales tout comme pour les entrées lexicographiques ; il est à noter que tous les tableaux de conversion sont téléchargeables (cf. p. 10), ex.

◆ COPIE D'ÉCRAN DU TABLEAU DE CONVERSION DU N.F. **ardeur**

↓ pour téléchargement

Export			
Niveaux	Siècle	Forme/colloc	Histoire
1	12		1120-50 « chaleur brûlante » (Gd mal fit Adam, éd. H. Suchier, I, 105 ds
2			au fig.
2.a	12		1170-91 « désir charnel » (Béroul, Tristan, éd. E. Muret, 1195, <i>ibid.</i> :
2.b	13		1203 « passion amoureuse » (Chastelain de Coucy, éd. Crapelet,
2.c	14		mil. xives. « désir violent » (Li Batard de Bouillon, éd. A. Scheler,

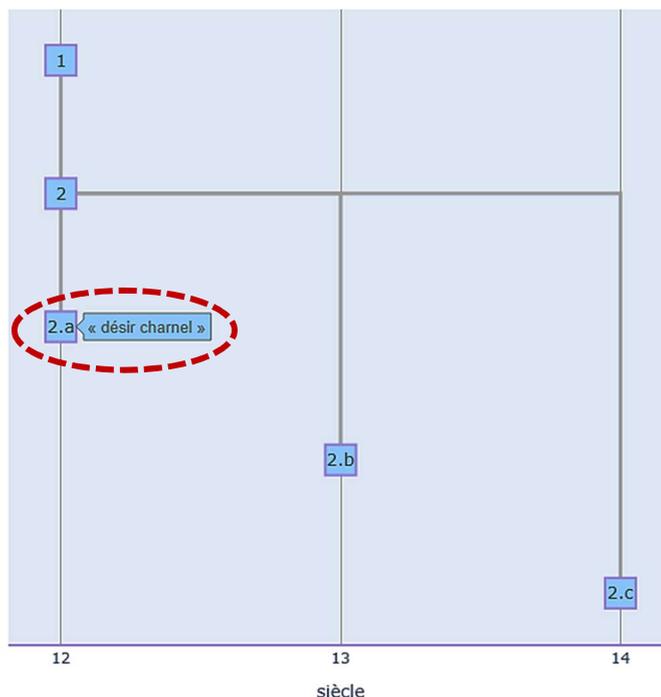
◆ VERSION TÉLÉCHARGÉE DU MÊME TABLEAU DE CONVERSION DANS UN FICHIER EXCEL⁸

	A	B	C	D
1	Niveaux	Siècle	Forme/colloc	Histoire
2	1	12		1120-50 « chaleur brûlante » (Gd mal fit Adam, éd. H. Suchier, I, 105 ds T.-L. : Senz repos ardor aux enfers);
3	2			au fig.
4	2.a	12		1170-91 « désir charnel » (Béroul, Tristan, éd. E. Muret, 1195, <i>ibid.</i> : Sire, en nos les lépreux a si grant ardor, Soz ciel n'a dame qui un jor Pëust sofrir nostre convers);
5	2.b	13		1203 « passion amoureuse » (Chastelain de Coucy, éd. Crapelet, 239, <i>ibid.</i> : Moulit souvent souspiroit d'ardour);
6	2.c	14		mil. xives. « désir violent » (Li Batard de Bouillon, éd. A. Scheler, Bruxelles 1877, 170, <i>ibid.</i> : de combattre as Turs avoient grant ardour). Rem. Nombreuses attest. au masc., au xvies. (Hug.).

Il est finalement à noter que nos **interfaces graphiques** sont conçues de manière à faire apparaître la définition de toute rubrique dotée d'une définition en faisant glisser la souris sur le nœud correspondant, ex. pour la rubrique 2.a de l'entrée **ardeur** :

⁸ Les segments en rouge dans le tableau sont lisibles sur l'écran en faisant glisser la flèche de la souris sur chaque cas au contenu a priori incomplet.

ardeur



Ces résultats, dans leur état d'avancement actuel, ouvrent déjà de vastes perspectives

- sur le **classement des entrées par profils graphiques** (en fonction de la localisation et dissémination des rubriques),
- et sur la possibilité de mener une enquête générale sur les extensions sémantiques de chaque siècle (cf. François, à paraître) avec **recherche de leurs motivations possibles**

Ils permettront notamment de reprendre à nouveaux frais le questionnement d'Antoine Meillet dans « Comment les mots changent de sens » (1906). Meillet se demandait il y a plus d'un siècle comment et pourquoi les locuteurs s'entendent implicitement sur le recyclage d'un mot ancien pour désigner une réalité devenue suffisamment saillante pour réclamer une désignation lexicale. On peut évoquer

- un **changement de perception du monde**, notamment l'entrée en scène d'un nouveau phénomène ou outil scientifique ou technologique (par ex. le recyclage au 17^e s. par Descartes de l'ancien vocable *lentille* pour désigner un verre circulaire poli de manière à produire une image grossie, ou celui du n.m. *arbre* cf. *arbre mécanique*⁹)
- ou un **changement sociétal** : c'est ce qu'a montré par ex. Jost Trier dans sa thèse de 1931 sur l'évolution du champ conceptuel de l'entendement (*Sinnbezirk des Verstands*) durant le Moyen-Âge allemand entre une première époque dominée par la féodalité, l'Église et la paysannerie, et une seconde marquée par l'émergence d'une forme primitive de bourgeoisie avec la place croissante de l'artisanat et du commerce au détriment de la paysannerie, dont le reflet lexical a été l'extension de la valeur péjorative de *die List* (le mode de pensée de la

⁹ Cf. Wikipedia <[https://fr.wikipedia.org/wiki/Arbre_\(m%C3%A9canique\)](https://fr.wikipedia.org/wiki/Arbre_(m%C3%A9canique))> : « Un arbre est un organe mécanique transmettant une force ou une puissance entre deux autres organes mécaniques (...) LA FORME GÉNÉRALEMENT CYLINDRIQUE DE CET ORGANE EST À L'ORIGINE DE SON NOM » et l'entrée H-É du n.m. *arbre* : 2. a) 1100-50 « pièce horizontale qui transmet le mouvement dans une machine motrice » (*Pèlerinage Charlemagne*, éd. E. Koschwitz-G. Thureau, 372, *ibid.* : Altresil fait torner com ARBRE DE MOLIN)

paysannerie), vocable qui finit au 13^e siècle par signifier la ruse et la perfidie. L'évolution de la signification du n.m. et adj. *vilain* a connu un cheminement similaire, comme le montre son entrée H-É :

1. *Ca* 1119 subst. « paysan libre » (Philippe de Thaon, *Comput*, éd. E. Mall, 132) – xiv^es., repris par les historiens du dr. au xvii^es., v. K. Baldinger ds *R. Ling. rom.* t. 26, p. 313;
1169-78 « roturier » (Jean de Meun, *Rose*, éd. F. Lecoy, 18587);
ca 1135 adj. « qui a les caractères du paysan » (*Couronnement Louis*, éd. Y. G. Lepage, 5), pour l'évol. sém., v. aussi Hollyman, pp. 162-164; d'où
- 2.a) 1135 « bas » (*ibid.*, 1886); *ca* 1155 « laid (moralement) » (Wace, *Brut*, éd. I. Arnold, 1478);
- 2.b) *ca* 1200 « laid (physiquement), qui déplaît à la vue » (*Roman de Guillaume de Dole*, éd. G. Servois, 268)

Au terme de la vérification, aussi complète que possible, de l'algorithme de construction des tableaux de conversion et des interfaces graphiques, laquelle occupera sans doute les membres du projet durant plusieurs mois, notre ambition sera double :

- a) assurer une large diffusion à ces résultats, notamment en demandant à l'**ATILF** son **aval scientifique** dans le cadre du dépôt d'un dossier d'ouverture d'un poste d'ingénieur(e) à l'appui de l'achèvement du projet,
- b) et publier notamment un article de synthèse en anglais dans une revue appropriée, par exemple le *Journal of Historical Linguistics*.

Précision indispensable : la constitution de ce mini-groupe de recherches n'a aucune base institutionnelle et notre collaboration est fondée uniquement sur la curiosité scientifique et le souci de l'intérêt des internautes concernés à titre professionnel ou simplement culturel par l'histoire de la langue française, en l'absence de tout financement.

RÉFÉRENCES BIBLIOGRAPHIQUES

◆ PUBLICATIONS DES MEMBRES DU PROJET

- CHARDON Laurette / REYNAUD Justine / FRANÇOIS Jacques (2023), « La polysémie évolutive du lexique français (12^e-20^e siècle), *Cahier CRISCO 37*, Université de Caen-Normandie <hal-04227851v1>
- FRANÇOIS Jacques (2007), *Pour une cartographie de la polysémie verbale*. Louvain : Peeters.
- FRANÇOIS Jacques (2008), *Une approche diachronique quantitative de la polysémie verbale*. Cahier du CRISCO 24. Université de Caen-Normandie. <hal-01870490v1>
- FRANÇOIS Jacques (2009), « Perte de prédicativité et auxiliarisation en français. Examen intégratif de deux ressources lexicales ». In A. Ibrahim (dir.), *Prédicats, prédication et structures prédicatives*, Cellule de Recherche en Linguistique, 147-161.
- FRANÇOIS Jacques (2010), « L'étude de la polysémie verbale entre dérivation et invariance ». *Actes du 2^e Congrès Mondial de Linguistique Française*. La Nouvelle-Orléans, juillet 2010. (https://www.linguistiquefrancaise.org/articles/cmlf/abs/2010/01/cmlf2010_000268/cmlf2010_000268.html)
- FRANÇOIS Jacques (2012), « Pour une lexicologie multidimensionnelle : comment penser simultanément les dimensions onomasiologique, sémasiologique et diachronique ? In F. Lautel-Ribstein (dir.) *Formes sémantiques, langages et interprétations – hommage à Pierre Cadiot*. Tribune internationale des langues vivantes, 123-132.
- FRANÇOIS Jacques (2020), « Pour un retraitement informatisé et dynamique des notices historiques du TLFi ». *Cahiers de Lexicologie* 117, 55-92 <hal-04443224v1>
- FRANÇOIS Jacques (2021a), « Comment visualiser l'évolution historique des polysémies lexicales : l'itinéraire sémantique de *terre* et *monde* » *Zeitschrift für romanische Philologie* 137(3), 625-665 <hal-04443255v1_>
- FRANÇOIS Jacques (2021b), « Les fluctuations historiques de la polysémie lexicale'. *Travaux de linguistique* n° 81(2), 57-98 < hal-04443255v1 >
- FRANÇOIS Jacques (à paraître), « Les extensions de sens », E. Buchi (dir.), *Manuel d'étymologie lexicale romane*, chap. 35. Berlin : De Gruyter.
- FRANÇOIS Jacques / CHARDON Laurette / REYNAUD Justine (2023), « La polysémie évolutive des vocables du français contemporain visualisée sur la base des entrées historiques du TLFi ». *Bulletin de la Société de Linguistique de Paris* 118(1) : pp. 75-124. <hal-04598504v1>

◆ PUBLICATIONS DE RÉFÉRENCE SUR LA SÉMANTIQUE HISTORIQUE ET LA POLYSÉMIE ÉVOLUTIVE

- BLANK A. (1997), *Prinzipien des lexikalischen Bedeutungswandels am Beispiel der romanischen Sprachen*. Berlin : De Gruyter (Beihefte zur Zeitschrift für romanische Philologie), [Principes de l'évolution sémantique du lexique illustré par les langues romanes].
- BLANK Andreas (2000), « Pour une approche cognitive du changement sémantique : aspect sémasiologique ». In J. François (dir.), *Théories contemporaines du changement sémantique*. Mémoire IX de la Société de Linguistique de Paris, p.59-74.
- BRÉAL Michel (1897), *Essai de sémantique* (Paris : Hachette) .

- DARMESTER Arsène (1887), *La vie des mots étudiée dans leurs significations*. Paris : Delagrave.
- GEERAERTS Dirk (1997), *Diachronic prototype semantics – A contribution to historical lexicology*. (Oxford : Clarendon).
- GEERAERTS Dirk (2010), *Theories of lexical semantics* (Oxford : Oxford University Press).
- KOCH Peter (2001) „Bedeutungswandel und Bezeichnungswandel. Von der kognitiven Semasiologie zur kognitiven Onomasiologie“, in : *Zeitschrift für Literaturwissenschaft und Linguistik* 121, 7–36 [Évolution des significations et des désignations. De la sémasiologie cognitive à l’onomasiologie cognitive]
- MEILLET Antoine (1906), « Comment les mots changent de sens », *Année sociologique*, 9: 1-38 [rééd. en 1921 dans *Linguistique générale et linguistique historique*]
- NYROP Krystoffer (1913) , *Grammaire historique de la langue française*. vol.4 : *Sémantique*. Copenhague-Leipzig-Paris.
- PAUL Hermann (1880sq), *Prinzipien der Sprachgeschichte* (Halle: Max Niemeyer) [Principes de linguistique historique]
- SPERBER Hans (1923), *Einführung in die Bedeutungslehre* (Bonn: Schroeder) [Introduction à la sémantique].
- STERN Gustav (1931), *Meaning and Change of Meaning* (Goteborg: Elanders)
- TRIER Jost (1931), *Der deutsche Wortschatz im Sinnbezirk des Verstandes : die Geschichte eines sprachlichen Feldes*. Thèse d’habilitation, université de Münster. [Le vocabulaire allemande dans l’espace sémantique de l’entendement : l’histoire d’un champ linguistique]
- ULLMANN Stephen (1951), *The Principles of Semantics* (Oxford : Basil Blackwell; Glasgow: Jackson).
- ULLMANN Stephen (1959), *Précis de sémantique française* (Berne : Francke).
- ULLMANN Stephen (1962), *Semantics: An Introduction to the Science of Meaning* (Oxford : Basil Blackwell).
- ULLMANN Stephen (1974), *Words and their meanings*. Australian National University, Lectures 1974. <https://openresearch-repository.anu.edu.au/>
- VON WARTBURG Walter (1929sq), *Französisches Etymologisches Wörterbuch – eine Darstellung des galloromanischen Sprachschatzes*. (2003), Édition française par Eva Buchi (Paris : Champion) [Dictionnaire étymologique du français – Un tableau du vocabulaire gallo-roman]

INDEX DES 12 VOCABLES INSPECTÉS DE PRÈS

<i>aimer</i>	v.t.	18
<i>avantage</i>	n.m.	17
<i>ardeur</i>	n.f.	8-10, 12, 14-15, 38
<i>corps</i>	n.m.	30-34
<i>défendre</i>	v.t.	19
<i>fondement</i>	n.m.	19
<i>honneur</i>	n.m.	18
<i>même</i>	adj./adv.	35-37
<i>méthode</i>	n.f.	14
<i>performance</i>	n.f.	18
<i>posséder</i>	v.t.	18
<i>vilain</i>	n.m./adj.	41

GLOSSAIRE

notion à expliciter	notion super-ordonnée	Explication
amplitude	entrée H-É	L' amplitude d'une ► entrée H-É se définit comme le nombre de siècles qu'elle couvre.
composante dérivée	rubrique substantielle	Les composantes dérivées d'une ► rubrique substantielle concernent les sens dérivés de celui de la ► composante primaire ; elles sont dotées d'une datation de 1 ^{ère} attestation, d'une référence, et fréquemment d'une métadonnée de domaine thématique ou de relation rhétorique (ex. <i>d'où</i> ; <i>au fig.</i>) ; elles sont ignorées dans le tableau de conversion afin de ne pas surcharger l'interface graphique.
composante primaire	rubrique substantielle	Une ► rubrique substantielle peut comporter une ou plusieurs composantes séparées par le signe typographique <;> ; seule la composante primaire est prise en compte dans le tableau de conversion graphique
dissémination	rubrique	La dissémination d'une ► entrée H-É se définit statistiquement comme l'écart-type entre la localisation temporelle de chacune de ses rubriques
entrée H-É	article du TLFi	Tous les articles du TLFi comportent une entrée H-É à la suite de l'entrée lexicographique, sauf en cas de dégroupement de l'article en dépit d'une même étymologie :
interface graphique (arborescente)	entrée H-É	L'interface graphique d'une ► entrée H-É est produite par un algorithme graphique à partir des données fournies par le ► tableau de conversion graphique ; elle construit un arbre de dérivation des rubriques superposé à un cadre orthogonal avec les rubriques en lignes et les siècles en colonnes.
localisation séculaire	rubrique	La localisation séculaire d'une ► rubrique est établie à partir de la datation de sa 1 ^{ère} attestation fournie par l'entrée H-É et harmonisée au niveau du siècle
localisation séculaire	entrée H-É	La localisation séculaire d'une ► entrée H-É est fournie par sa ► moyenne séculaire
moyenne séculaire	entrée H-É	La moyenne séculaire d'une ► entrées H-É est calculée à partir du ► nombre de ses rubriques divisé par son ► amplitude.

nombre de rubriques	entrée H-É	Le nombre de rubriques d'une ► entrée H-É est à la base le nombre de ses parties introduites par un identifiant (en gras). Ce nombre fournit le nombre de lignes du ► tableau de conversion graphique de l'entrée
polysémie évolutive	vocable	La polysémie évolutive (ou cumulative , voir p. X) d'un ► vocable se définit comme le profil historique de sa polysémie, considérée à partir de son développement final au 20 ^e siècle -en fonction des données les plus récentes des entrées H-É)
public	format d'une entrée H-É	Toute entrée H-É se présente sous deux ► formats : le format public fourni dans le portail lexical du CNRTL-ORTOLANG, et le format annoté en langage ► XML développé à l'origine par l'INaLF (format SGML) converti ultérieurement en HTML puis en XML
rubrique auxiliaire	entrée H-É	Une rubrique auxiliaire se définit comme une rubrique d'une entrée H-É dénuée d'identification, mais nécessaire pour construire l' ► interface graphique arborescente de la polysémie évolutive de l'entrée.
rubrique substantielle	entrée H-É	Une rubrique substantielle se définit comme une rubrique d'une ► entrée H-É dotée d'un identifiant. Fréquemment, le nombre de nœuds de l'interface graphique arborescente d'une entrée H-É est supérieur au nombre de rubriques substantielles inventorié dans le tableau de conversion graphique.
tableau de conversion graphique	entrée H-É	Le tableau de conversion graphique d'une ► entrée H-É est construit à partir de la segmentation (dans l'un ou l'autre format public ou annoté) de ses ► rubriques substantielles ; qui identifie quatre composantes de l'entrée : son <u>identifiant</u> , la <u>datation</u> de la 1 ^{ère} attestation de son contenu, la <u>forme de mot</u> ou la <u>collocation</u> à laquelle elle s'applique, et le <u>reste</u> de la (1 ^{ère} composante de) la rubrique, notamment la définition entre guillemets.