

Notice de l'édition électronique du texte médiéval à l'épreuve des agrégations de lettres et de grammaire en 2021

Villon, *Lais* (L.), *Testament* (T.), *Poésies diverses* (P.)

<http://www.atilf.fr/dmf/VillonAgregation/>

agregation@atilf.fr

Cette édition électronique a été réalisée d'avril à septembre 2020, au sein de l'équipe Linguistique historique française et romane du laboratoire ATILF (CNRS|UL), par Sylvie Bazin-Tacchella, Jean-Michel Jézéquel et Gilles Souvay.

Elle a été faite à partir du texte établi par l'édition critique de Jean-Claude Mühlethaler, [François Villon, *Lais*, *Testament*, *Poésies diverses*, 2004, Paris, Honoré Champion](#).

Malgré la relecture des lemmes et étiquettes d'un très grand nombre de mots, quelques erreurs peuvent encore être présentes ; il est possible de nous les signaler à agregation@atilf.fr.

La lemmatisation et l'étiquetage morphosyntaxique ont été réalisés au moyen de la plate-forme informatique LGeRM, développée par Gilles Souvay, ingénieur de recherche ATILF (CNRS|UL).

Nous citer :

François VILLON, *Lais*, *Testament*, *Poésies diverses*. Édition électronique par Sylvie Bazin-Tacchella, Jean-Michel Jézéquel et Gilles Souvay, aide à la préparation au concours des agrégations de lettres et de grammaire 2021. <http://www.atilf.fr/dmf/VillonAgregation/>, ATILF - CNRS & Université de Lorraine. Juillet 2020.

SOMMAIRE

I. Comment cette édition électronique des textes de Villon a-t-elle été élaborée ?	3
II. Comment utiliser cette ressource ?	4
A. Lecture en continu	4
B. Requêtes simples	5
1. Par forme	5
2. Par lemme	6
• Traitement des parenthèses () et crochets []	7
• Traitement des mots dans les titres	7
• Comment déterminer un lemme ?	7
• Remarque sur les étiquettes morphosyntaxiques des formes verbales	8
C. Requêtes avancées	9
1. Forme/lemme qui commence par ou se termine par ...	9
2. Concordancier : mot précédent / mot pivot / mot suivant	10
a) Jouer avec les lemmes et les formes	10
b) ... avec les étiquettes morphosyntaxiques	11
• Corrélation entre lemme et étiquette.	12
• Le cas des formes contractées	12
• Quelques étiquettes morphosyntaxiques utiles	13
c) ... avec les codes grammaticaux	14
• Code grammatical (lexicologie) ≠ étiquette morphosyntaxique (édition électr.)	14
• Recourir au code grammatical dans une requête	15
3. Aller encore plus loin	15
III. Autres informations utiles	16
A. Ressources du DMF pour l'agrégation 2021	16
1. Villon cité dans les articles du DMF	16
2. Synthèses grammaticales du DMF	16
B. Les notes grammaticales ou philologiques propres à cette édition électronique (en exposant dans le texte)	16
IV. Index	17

I. Comment cette édition électronique des textes de Villon a-t-elle été élaborée ?

Phase 1 : le texte formel.

L'édition papier au programme a été traitée par un logiciel de reconnaissance de caractères (dit OCR). Le résultat a ensuite été relu humainement à trois reprises et, au besoin, amendé manuellement. Le risque de coquilles est mince. (information technique : le texte a été saisi au format XML-TEI).

Phase 2 : la lemmatisation & l'étiquetage morphosyntaxique.

Le texte formel ainsi obtenu a été traité par la plate-forme de lemmatisation LGeRM : à chaque mot ont été affectés automatiquement un lemme et une étiquette morphosyntaxique. Par exemple, sont affectées du lemme [AVOIR1](#) toutes les formes du verbe *avoir* : l'infinitif *avoir*, le participe passé *eu*, les formes conjuguées *ai*, *ay* (ind. pst P1) ou *aura*, *arra* (ind. fut.I P3) ou encore *ot*, *eust* (ind. ps. P3). Ainsi, les variations flexionnelles et les variations graphiques sont neutralisées et subsumées sous leur lemme commun.

La lemmatisation automatique a ensuite été reprise humainement et manuellement. Repasser sur l'ensemble des mots des textes de Villon pour vérifier le lemme et l'étiquette de chacun serait un travail titanesque. Cette étape de "post-correction" humaine a été conduite de manière rationalisée : la plate-forme LGeRM indique les mots qui lui ont posé problème d'une part, et certaines formes sont connues pour requérir relecture humaine du traitement automatique d'autre part. À titre d'exemple, les 550 formes *ne / n'* ont été systématiquement relues de façon à discriminer l'adverbe de négation [NE1](#) de la conjonction de coordination [NE2](#).

Les lemmes sont ceux du [DMF](#) et les étiquettes morphosyntaxiques sont celles du projet [CATTEX2009](#). Ainsi, la forme *escollier* a pour lemme [ÉCOLIER](#) et pour étiquette NOMcom. Ces informations se retrouvent aussi bien dans le tableau des résultats d'une recherche (colonnes de droite et de gauche) que dans le mini-menu qui s'affiche spontanément lors du passage de la souris sur n'importe quel mot des textes de Villon.

1	ÉCOLIER , subst. masc.	I (En) l'an (mil) quatre cens cinquante six, Je, François Villon, escollier , Considerant, de sens r	mot 19 NOMcom p.44 v.2
2	ÉCOLIER , subst. masc.	Donne tenir publique Ou l' escollier le maist Lieu n'est ou ce march	<ul style="list-style-type: none"> • voir dans le glossaire • voir dans le DMF
3	ÉCOLIER , subst. masc.	Qu'Amours occist de se Ung povre petit escoll Qui fut nommé Franç	<ul style="list-style-type: none"> • NOMcom <p>Toutes les occurrences Concordancier</p>

Remarque : la mention *subst. masc.*, dans la colonne de gauche, est le "code grammatical" que le DMF associe à chaque lemme, indépendamment de l'attestation dans le texte. Cf. *infra* II.C.2.c) pour plus d'information.

Autre exemple : la forme *ne*, si elle est adverbe de négation aura le lemme [NE1](#) et l'étiquette ADVneg, et si elle est conjonction de coordination ce sera [NE2](#) CONcoo.

Les lemmes sont toujours cliquables et renvoient directement à l'article du [DMF](#) correspondant. Les étiquettes morphosyntaxiques sont une précision grammaticale supplémentaire, souvent très utile.

Malgré la relecture des lemmes et étiquettes d'un très grand nombre de mots, quelques erreurs peuvent encore être présentes ; il est possible de nous les signaler à agregation@atilf.fr.

II. Comment utiliser cette ressource ?

Se connecter à <http://www.atilf.fr/dmf/VillonAgregation/>, cliquer sur *Entrer dans le texte*.

La navigation se fait au moyen du bandeau des onglets suivant :



A. Lecture en continu

Onglet noir TEXTE.

Lecture du texte en continu

Accueil Texte Formes Lemmes Cré

Texte Droits S'identifier

Début << 100 >> page

284 « Ho***, ne te do***** ta**
 Et ne de***** te* do***** !
 Se tu n'a[** ta** qu*eu** Ja*** Cu***,
 Mi*** va*** vi*** so*** gr** bu***
 Po*** qu*av*** es** se*****
 288 Et po***** so*** ri*** tu*****. »

XXXVII
 Qu*av*** es** se***** ... Qu* di* ?
 Se*****, la*** ! ne l'es* il ma** ?
 292 Se*** ce qu* Da*** en di**,
 So* li** ne co***** ja***
 Et du se***** je me de****.
 Il n'ap***** a mo*, pe****
 Au* th***** le re*****
 296 Ca* c'es* of**** de pr*****

XXXVIII
 Sy ne su** - bi** le co*****
 Fi** d'an***** po***** de*****
 D'es***** ne d'au*** si****.
 300 Mo* de** es* mo**. Di** en ai* l'am* !

MAIS1, adv. d'intensité et conj.

- voir dans le glossaire
- voir dans le DMF

■ ADVgen

Toutes les occurrences
Concordancier

Le texte est consultable en continu page par page ou strophe par strophe *via* le menu déroulant de droite au-dessus du texte. Pour des raisons de droit d'édition, deux tiers des lettres de chaque mot ne sont pas visibles en lecture continue. La fonctionnalité reste cependant utile en miroir avec l'édition papier. Ainsi, à chaque mot sur lequel l'on achoppe lors de la lecture de l'édition papier, il est possible, sans faire de requête, d'en afficher immédiatement le lemme dans l'édition électronique en passant le pointeur de la souris sur cette forme.

Menu déroulant de droite : L —> *Lais* (ou *Petit Testament*) ; T —> *Testament* ; P —> *Poésies diverses*.

À tout moment, lorsque du texte de Villon est affiché, que ce soit en lecture continue dans cet onglet noir TEXTE, ou dans un tableau de résultats de requête où apparaît le contexte des occurrences (voir II.B. *infra*), il suffit de **passer le pointeur de la souris au-dessus d'un mot pour voir son lemme et son étiquette**. En cliquant sur *voir dans le DMF*, l'on est directement redirigé dans une nouvelle fenêtre **vers l'article correspondant du DMF**. En cliquant sur *voir dans le glossaire*, l'on accède à toutes les attestations du lemme (cf. *infra* II.B.2. Requête par lemme).

B. Requetes simples

1. Par forme

Onglet noir FORMES, sous-onglet jaune RECHERCHE.

Taper *estre*, sélectionner *Texte exact*, bouton *Appliquer*, bouton *Attestations*.

■ Synthèse sur la forme			
estre		20 attestations	
■ Détail des attestations			
1	ÊTRE ₁ , verbe	Sy fut mis devant ce cadés Pour estre jugiez a mourir.	mot 3138 VERinf p.90 v.136
	estre	XVIII L'empereur si l'araisonna :	
2	ÊTRE ₁ , verbe	Et lors, qui m'eust veu condescendre A mal, estre ars et mis en cendre Jugié me feusse de ma voys :	mot 3333 VERinf p.92 v.165
	estre	Car enffant n'a, frere ne seur, Qui lors vouldist estre son pleige.	mot 4449 VERinf p.102 v.320
	estre	XLI La mort le fait fremir, pallir,	
4	ÊTRE ₁ , verbe	Ce confort prent povre viellart Lequel d' estre plaisant raillart Ot le bruyt, lors que jeune estoit,	mot 5211 VERinf p.108 v.425
	estre	« Or y pensez, belle Gaultiere, Qui escolliere souliez estre , Et vous, Blanche la Savetiere :	mot 6026 VERinf p.114 v.534
6	ÊTRE ₂ , subst. masc.	N'espargniez homme, je vous prie, Car vieilles n'ont ne cours nè estre Ne que monnoye qu'on descreye.	mot 6064 NOMcom p.114 v.539
	estre		

La recherche par forme simple s'en tient à l'exacte forme du texte. Elle peut permettre d'afficher ensemble des mots affectés de lemmes différents. Ainsi, pour un travail sur la forme *estre*, s'affichent aussi bien les formes de l'infinitif du verbe [ÊTRE₁](#) que les formes de l'infinitif substantivé lexicalisé, rangé sous le lemme [ÊTRE₂](#).

En cliquant sur les références de la page et du vers, l'on est renvoyé à la page correspondante en lecture continue, de façon à avoir un contexte plus large et une présentation du texte similaire à l'édition papier qui fait référence.

2. Par lemme

Onglet noir LEMMES, sous-onglet jaune RECHERCHE.

Taper *prendre*, sélectionner *Texte exact*, bouton *Appliquer*, bouton *Attestations*. Ainsi s'affichent l'ensemble des attestations du lemme **PRENDRE**, toutes variations graphiques et flexionnelles comprises.

Vue Recherche par lemme

The screenshot shows the 'Recherche' view for the lemma 'PRENDRE, verbe'. It includes a navigation bar with tabs for 'Accueil', 'Texte', 'Formes', 'Lemmes', and 'Crédits et contact'. Below this, there are sub-tabs for 'Recherche', 'Nomenclature', 'Cooccurrences', 'Concordancier', and 'Index lemmatisé'. The main content area displays a table of attestations for the lemma 'PRENDRE', with columns for the form, the text, and the source (mot, VER, p.).

Forme	Texte	Mot	VER	p.	v.
preigne 1 att.					
1	Bons comme fut saint Marcial. Ainsi en preigne au feu daulphin ! Je ne lui soubzhaicte autre mal,	mot 2698	VERcjcj	p.88	v.70
pren 1 att.					
1	Que l'en ne me viengne espier ! Pren ancre tost, plume, pappier ! Ce que nomme escriptz vistement,	mot 7841	VERcjcj	p.128	v.789
prenant 1 att.					
1	Jhesus regnant qui n'a ne fin ne cesse. Le Tout Puissant, prenant nostre foiblesse, Laissa les cieulx et nous vint secourir,	mot 8706	VERppa	p.136	v.905
prendre 7 att.					
1	Ce que je pourray acquerir : L'en ne doit [trop] prendre des siens Ne trop ses amis surquerir.	mot 935	VERinf	p.50	v.135

Les attestations (ou occurrences) sont classées par forme graphique. En haut à droite, sont comptabilisés le nombre de formes graphiques différentes ainsi que le nombre total d'attestations du lemme.

Vue Export

Afin d'avoir une liste intégrale des occurrences classées par ordre d'apparition, il convient de cliquer sur le bouton blanc *Export*, situé entre la barre des onglets et le cartouche du lemme.

Pour une liste intégrale des occurrences rangées par forme — puis par ordre d'apparition au sein de chaque liste d'une forme —, il faut alors cliquer sur le bouton blanc *Concordancier*.

The screenshot shows the 'Concordancier' view for the lemma 'PRENDRE, verbe'. It includes a navigation bar with tabs for 'Recherche', 'Nomenclature', 'Cooccurrences', 'Concordancier', and 'Index lemmatisé'. The main content area displays a table of attestations for the lemma 'PRENDRE', with columns for the form, the text, and the source (mot, VER, p.).

Mot	VER	p.	v.	Forme	Texte
				prins	grief d' amours allegence . [Et se j' ay en ma faveur Ses doux regards et beaulx semblans
				prins	ung autre coing . Le regard de celle m' a Qui m' a esté felonnie et dure . Sans
				prins	; Aussi a Jehan Raguier la somme De cent frans sur tous mes biens . Mais quoy ? Je
				prendre	que je pourray acquerir : L' en ne doit [trop] des siens Ne trop ses amis surquerir . Item
				prendre	veult vendre , Et La Vache qu' on ne peult . Le villain qui la trousse au col ,
				Prins	et a Cholet Je laisse a la foyz ung canart sur les murs , comme on souloit , Envers

Nomenclature des lemmes

Accueil	Texte	Formes	Lemmes	Crédits et contact
Recherche	Nomenclature	Cooccurrences	Concordancier	Index lemmatisé
mots lexicaux	noms propres	mots étrangers		mots avec lemme réservé
<< A B C D E F G H I J L M N O P Q R S T U V Y >>				
PRÊCHER, verbe	4	1	prescher	
PRÊCHEUR, subst. masc.	1	1	prescheurs	
PRÉCIEUX, adj.	3	1	precieulx	
PRÉCIPITER, verbe	1	1	precepikte	
PRÉDIRE, verbe	1	1	predit	
PREMIER₁, adj. et adv.	4	3	premier premiere premiers	
PREMIÈREMENT, adv.	2	1	premierement	
PRENDRE, verbe	47	15	preigne pren prenant prendre prendroie prenez prenoïe prenons prens prent prindrent prins prinse pris prise	

Afin d'avoir une liste synoptique des différentes formes graphiques d'un même lemme, il convient de cliquer sur le bouton blanc *Nomenclature*, au-dessus du cartouche du lemme (à partir de la première capture d'écran p.6).

Le bouton *Nomenclature* redirige vers la liste alphabétique exhaustive des lemmes utilisés dans cette édition électronique. Pour chaque lemme sont présentés respectivement le nombre d'occurrences, le nombre de formes graphiques différentes et ces formes. Les boutons blancs (lettres et flèches) au-dessus et sous le tableau permettent de circuler dans la totalité de la liste des lemmes de cette édition.

AVOIR₁, verbe	315	47	a a[s] ai aient aiés aiez a(i)rait ait aray araz aront arra as aura auroit auront avés avez avions avoir avoit avons avoye ay aye ayent ayés ayez ayons ayt é eu eues eurent eusse eusse[n]t eusse[s] eussent eusse(s) eussiez eus(t) eust eüst euz on ont ot
---------------------------------	-----	----	---

Ci-contre, le cas du lemme [AVOIR₁](#) qui recense des formes graphiques avec parenthèses et crochets.

- Traitement des parenthèses () et crochets []

Le texte versifié de Villon, établi par J.-C. Mühlethaler chez Champion, contient, outre les lettres alphabétiques, des crochets et des parenthèses. Ces signes d'édition philologique viennent corriger la leçon du manuscrit C : les crochets introduisent des éléments absents du ms. C mais requis, quand les parenthèses retranchent des éléments présents dans le ms. C mais indésirables.

Les parenthèses et crochets ont été informatiquement traités dans cette ressource comme des symboles composant la forme d'un mot. Une requête par lemme aura pour avantage de ne pas manquer les mots "déformés" par les signes correctifs d'édition. Ainsi, une requête par lemme de [HOMME](#) inclura bien l'occurrence *homme(s)*, qu'une requête par forme de *homme* aurait manquée.

- Traitement des mots dans les titres

Une recherche par lemme permet aussi d'éviter des occurrences contenus dans les titres. Par exemple, une recherche par forme présentera 25 occurrences de *ballade*, quand une recherche par lemme retiendra seulement les trois occurrences de *ballade* au sein du texte versifié de Villon.

- Comment déterminer un lemme ?

Comme le montre la dernière capture d'écran, la graphie d'un lemme s'aligne, autant que faire se peut, sur l'orthographe moderne. Pour retrouver la graphie d'un lemme, il est possible de faire une requête de forme (cf. II.B.1. *supra*) et de voir le lemme qui lui a été affecté. Une façon encore plus

efficace est d'interroger le [DMF](#) à partir de l'occurrence médiévale (même en soumettant une forme verbale fléchie, ou une forme élidée) : les différents lemmes possibles pour cette forme graphique seront proposés. Parfois le code grammatical (*adj.*, *verbe*, *etc.*) qui suit le lemme suffira à choisir, parfois il faudra aller consulter rapidement les articles DMF des différents lemmes proposés pour s'assurer de retenir le bon.

Vue d'une recherche sur [DMF](#)

Mot ou forme

Mot ou forme: voir | Filtre sur les entrées | Initiale des entrées | Nomenclature

voir | Rechercher | Effacer | + options

Saisir un mot ou une forme sans se préoccuper des entrées du DMF : des propositions s'afficheront.
 La recherche porte sur les variantes graphiques connues du lemmatiseur.

Résultat de la recherche

La forme *voir* est connue du lemmatiseur avec l'analyse suivante :

VOIR2, adj., subst. masc. et adv. | famille | structure | sans exemple | complet | textes | proverbes
 [T-L : *voir* ; GD : *voir2* ; DÉCT : *voir* ; FEW XIV, 329b *verus*]

VOIR1, verbe | famille | structure | sans exemple | complet | textes | proverbes
 [T-L : *vëoir* ; GD : *veoir* ; GDC : *veoir* ; DÉCT : *vëoir* ; FEW XIV, 421a *videre* ; TLF : *voir*]

VAIR, adj. et subst. masc. | famille | structure | sans exemple | complet | textes | proverbes
 [T-L : *vair2* ; GD : *vair1* ; DÉCT : *vair1* ; FEW XIV, 184b *varius* ; TLF : *vair*]

Pour la forme *voir*, trois lemmes sont proposés : [VOIR1](#) pour le verbe, [VOIR2](#) pour la forme nominale issue du latin *verus, a, um* et [VAIR](#).

- Remarque sur les étiquettes morphosyntaxiques des formes verbales

Onglet LEMMES, sous-onglet RECHERCHE, taper *voir* ou *voir1*, choisir alors *VOIR1* dans la liste des résultats, cliquer sur *Attestations*.

Accueil | Texte | Formes | Lemmes | Crédits et contact

Recherche | Nomenclature | Cooccurrences | Concordancier | Index lemmatisé

<< VOIR1, subst. masc. | Nomenclature | >> | Concordancier | Export

VOIR1, verbe | Voir DMF | 18 formes | 40 attestations

tri par forme | 5 occurrences par forme

veu 3 att.		
1	Qui m'eust fait en bon cuer entrer, Et lors, qui m'eust veu condescendre A mal, estre ars et mis en cendre	mot 3328 VERppe p.92 v.164
2	Et du Ladre de dessus ly. LXXXIII Se du Ladre eust veu le doyz ardre, Ja n'en eust requis reffrigere	mot 8030 VERppe p.130 v.817
3	Comme homme viel qui chancelle et trespigne L'ay veu souvent, quant il s'alloit coucher, Et une foiz il se fist une bigne	mot 11346 VERppe p.156 v.1255
vist 1 att.		
1	Jehan de Calaiz, honorable homme, Qui ne me vist des ans a trente Et ne scet comment on me nomme,	mot 15860 VERcjg p.190 v.1846
vit 1 att.		
1	Rancontré soit des bestes feu gectans Que Jason vit , querant la thoison d'or, Ou transmüé d'omme en beste sept ans	mot 21852 VERcjg p.330 v.41

L'on peut constater, dans la colonne de droite, que la granularité des étiquettes morphosyntaxiques ([CATTEX2009](#)) des formes verbales s'arrête à la discrimination entre forme conjuguée et formes non conjuguées : VERcjg, VERinf, VERppe (pour part.passé), VERppa (pour les formes verbales en -ant).

C. Requêtes avancées

1. Forme/lemme qui *commence par* ou *se termine par* ...

Onglet noir FORMES ou LEMMES, sous-onglet jaune RECHERCHE. Choisir l'option *texte en fin* ou *texte en début* peut permettre, par exemple, de travailler sur la morphologie.

Exemple avec *texte en fin*

Chercher les adverbes et les substantifs suffixés en *-ment*. En onglet noir FORMES et RECHERCHE, si l'on tape *ment* et sélectionne l'option *texte en fin*, l'on obtient toutes les attestations dans le texte se terminant par *-ment*, dont les verbes fléchis *desment* ([DÉMENTIR](#)) et *dorment*. En onglet noir LEMMES et sous onglet jaune RECHERCHE, cette même requête écartera ces dernières occurrences verbales car la requête se fera sur la graphie canonique des lemmes. L'on a alors seulement les adverbes et substantifs se terminant en *-ment*, et encore un peu de bruit avec l'adjectif [CLÉMENT](#).

Accueil Texte **Formes**

Recherche Nomenclature

Filtre de sélection

ment Appliquer Effacer

texte exact
 texte en début
 texte à l'intérieur
 texte en fin
 expression régulière

Liste des éléments

aiment 1	Attestations
aucunement 3	Concordances
autrement 2	Tri A Z 1 9
ayment 3	
bonnement 1	
charitablement 1	
clement 1	
commencement 3	
commandement 1	
commencement 1	
comment 7	
communement 2	
consentement 1	

Accueil Texte Formes **Lemmes**

Recherche Nomenclature Cooccurrences Concord

Filtre de sélection

ment Appliquer Effacer

texte exact
 texte en début
 texte à l'intérieur
 texte en fin
 expression régulière

sensible à la cas
 sensible aux dia

Liste des éléments

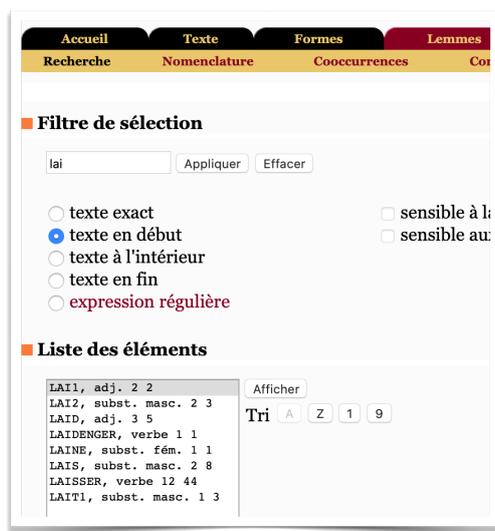
AUCUNEMENT, adv. 1 3	Afficher
AUTREMENT, adv. 1 2	Tri A Z 1 9
BONNEMENT, adv. 1 1	
CHARITABLEMENT, adv. 1 1	
CHEMINEMENT, subst. masc. 1 1	
CLÉMENT, adj. 1 1	
COMMANDEMENT, subst. masc. 1 1	
COMMENCEMENT, subst. masc. 2 4	
COMMUNÈMENT, adv. 1 2	
CONSETEMENT, subst. masc. 1 1	
DOUCEMENT, adv. 1 1	
DUREMENT, adv. 1 1	
ENFANTEMENT, subst. masc. 1 1	
ENGENDREMENT, subst. masc. 1 1	
ENSEIGNEMENT, subst. masc. 1 1	

Remarque : il n'est pas (encore) possible de sélectionner différents résultats (formes ou lemmes) de la requête afin d'afficher ensemble, au sein d'une même liste, les occurrences des différents résultats sélectionnés.

Exemple avec *texte en début*

Par exemple, le texte de Villon présente différentes formes graphiques *lay lays lais layz let*, dont certaines sont homonymiques entre elles. Ces formes se répartissent sur cinq lemmes [LAI1](#) [LAI2](#) [LAID](#) [LAIS](#) [LAIT](#). Une requête *lai* (onglet noir LEMME, RECHERCHE, option *texte en début*) peut permettre un travail sur ces homonymes. Il faudra afficher les résultats un par un pour avoir les occurrences en contexte.

Vue Recherche lemme commence par lai-



Filtre de sélection

lai [Appliquer] [Effacer]

texte exact
 texte en début
 texte à l'intérieur
 texte en fin
 expression régulière

sensible à l: []
 sensible au: []

Liste des éléments

LAI1, adj. 2 2	Afficher
LAI2, subst. masc. 2 3	
LAI3, adj. 3 5	
LAI4, subst. masc. 1 1	
LAI5, subst. masc. 2 8	
LAI6, subst. masc. 2 8	
LAI7, subst. masc. 1 3	

Tri A Z 1 9

Vue Nomenclature lai-



Recherche	Nomenclature	Cooccurrences	Concordancier	Index lemmatisé
mots lexicaux	noms propres	mots étrangers	mots avec lemme réservé	
<< [A B C D E F G H I J L M N O P Q R S T U V Y] >>				
LAI1, adj.	2	2	lay lays	
LAI2, subst. masc.	3	2	laiz lay	
LAI3, adj.	5	3	laide lays let	
LAI4, subst. masc.	1	1	laidanger	
LAI5, subst. masc.	1	1	layne	
LAI6, subst. masc.	8	2	lais laiz	
LAI7, verbe	44	12	laisa laissas laissay laisse laissè laisser laissié laissons lessay lesse lesserez lessié	
LAI8, subst. masc.	3	1	let	

De surcroît, dans ce cas très précis d'une recherche des lemmes commençant par certaines lettres, la consultation du sous-onglet jaune NOMENCLATURE (onglet noir LEMMES) sera efficace pour avoir une vue synoptique de la répartition des formes graphiques selon les différents lemmes. Cliquer sur bouton *L* dans la liste haute des boutons alphabétiques, puis choisir en bas dans le menu déroulant la lettre *I* de façon à faire débiter l'affichage avec les lemmes concernés.

L'option *sensible à la casse* : en la cochant, l'on demande au logiciel de distinguer les majuscules des minuscules.

L'option *sensible aux diacritiques* : en la cochant, l'on demande au logiciel de distinguer les lettres sans diacritique des lettres diacritées. Si l'option est décochée, en cherchant *ne* (onglet noir FORMES puis RECHERCHE), le logiciel proposera les trois formes *ně, ne, né* ; cochée, il n'y a plus qu'un résultat trouvé, *ne*.

2. Concordancier : mot précédent / mot pivot / mot suivant

Onglet noir FORMES ou LEMMES, sous-onglet jaune CONCORDANCIER.

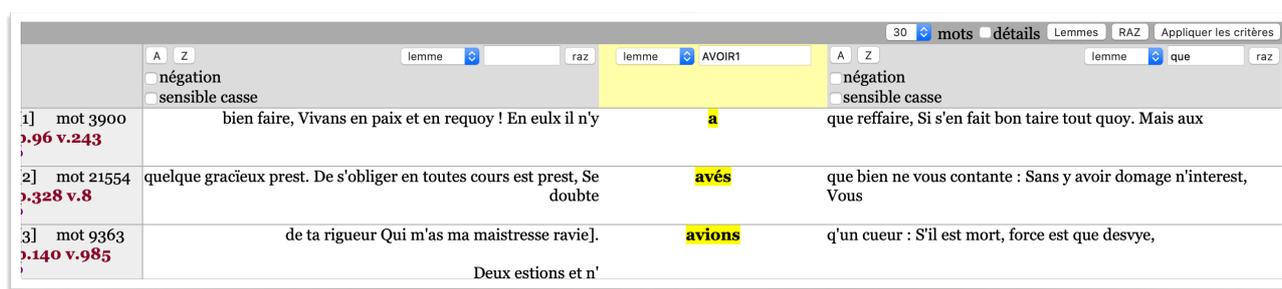
La fonctionnalité Concordancier permet de conduire des requêtes plus complexes en posant des critères aussi sur le mot qui précède et/ou le mot qui suit.

Le mot pivot (colonne jaune) doit être une sélection de lemme ou de forme.

Le mot précédent (colonne de gauche) et le mot suivant (colonne de droite) peuvent être la sélection d'une forme ou d'un lemme, mais aussi d'une étiquette morphosyntaxique ou d'un code grammatical.

a) Jouer avec les lemmes et les formes

Un exemple peut être : [AVOIR1] suivi de [QUE]



	A Z	lemme	raz	lemme	AVOIR1	A Z	lemme	que	raz
		<input type="checkbox"/> négation					<input type="checkbox"/> négation		
		<input type="checkbox"/> sensible casse					<input type="checkbox"/> sensible casse		
1]	mot 3900 v.96 v.243	bien faire, Vivans en paix et en requoy ! En eulx il n'y		a		que reffaire, Si s'en fait bon taire tout quoy. Mais aux			
2]	mot 21554 v.328 v.8	quelque gracieux prest. De s'obliger en toutes cours est prest, Se doubte		avés		que bien ne vous contante : Sans y avoir damage n'interest, Vous			
3]	mot 9363 v.140 v.985	de ta rigneur Qui m'as ma maïstresse ravie]. Deux estions et n'		avions		q'un cuer : S'il est mort, force est que desvye,			

De ce premier cas, l'on peut vouloir ajouter un nouveau critère : absence de la négation grammaticale **NE1**. Pour ce faire il convient de cocher, dans la première colonne, l'option *négation*. L'on a alors : non [NE1] + [AVOIR1] + [QUE].

	A	Z	lemme	NE1	raz	lemme	AVOIR1	A	Z	lemme	QUE	raz
[1]	mot 3900			<input checked="" type="checkbox"/>	négation							
	p.96 v.243			<input type="checkbox"/>	sensible casse							
							a					
[2]	mot 21554						avés					
	p.328 v.8											
[3]	mot 6535						avoit					
	p.118 v.606											

On peut modifier le taille des contextes droit et gauche en faisant varier le nombre de mots dans un menu déroulant de la partie supérieure du concordancier.

b) ... avec les étiquettes morphosyntaxiques

Un autre exemple pourrait être de rechercher toutes les occurrences du verbe *avoir* précédées d'un pronom. Il convient alors de sélectionner *étiquette* dans la colonne de gauche, et de taper PRO (bien respecter les majuscules). S'affichent toutes les occurrences précédées d'un pronom, qu'il soit démonstratif (PROdem), personnel (PROper), adverbial (PROadv), relatif (PROrel) ou interrogatif (PROint).

	A	Z	étiquette	PRO	raz	lemme	AVOIR1	A	Z	forme	raz
[1]	mot 227			<input type="checkbox"/>	négation						
	p.46 v.33			<input type="checkbox"/>	sensible casse						
							a				
[2]	mot 231						a				
	p.46 v.34										
[3]	mot 437						a				
	p.46 v.62										

Il est aussi possible de renseigner une étiquette intégrale (format AAAaaa). Un autre exemple avec le verbe *avoir* : (ADVgen) + [AVOIR1].

	A	Z	étiquette	ADVgen	raz	lemme	AVOIR1	A	Z	forme	raz
[1]	mot 14924			<input type="checkbox"/>	négation						
	p.182 v.1712			<input type="checkbox"/>	sensible casse						
							araz				
[2]	mot 10236						aura				
	p.146 v.1101										
[3]	mot 10575						aura				
	p.150 v.1148										

L'ensemble des étiquettes morphosyntaxiques utilisées sont le fruit du projet **CATTEX2009**. Le format des étiquettes pour cette édition électronique est le suivant : AAAaaa. Trois lettres capitales qui correspondent à la première colonne *Catégorie* de **CATTEX2009** et trois lettres minuscules qui

correspondent à la deuxième colonne *Type* de [CATTEX2009](#). Nous avons retenu le jeu minimal des étiquettes [CATTEX2009](#), les colonnes suivantes ne sont donc pas utilisées dans notre étiquetage morphosyntaxique. Une conséquence en est le niveau de granularité des étiquettes des formes verbales qui s'arrête à la distinction entre formes conjuguées et non conjuguées (VERcjc, VERinf, VERppe, VERppa), cf. capture d'écran *supra* II.B.2. point n°4.

- Corrélation entre lemme et étiquette.

Souvent, il n'y a qu'une seule étiquette possible pour un lemme ; par exemple, pour l'occurrence *escollier* l'on aura le lemme [ÉCOLIER](#) et l'étiquette NOMcom, pour l'occurrence *en* l'on aura [EN1](#) PRE ou bien [EN2](#) PROadv.

Cependant, pour une attestation *que*, si le lemme sera toujours [QUE](#), l'étiquette, en revanche, pourra varier selon trois types : CONsub, PROrel & PROint. Pour le si riche morphème *que*, notre étiquetage ne raffine pas plus en détail pour chacune de ces trois catégories d'analyse grammaticale.

La possibilité de variation de l'étiquette pour un même lemme tient au niveau d'analyse auquel chacun ressortit : le lemme est le résultat d'une analyse lexicologique (c'est bien le propre d'un dictionnaire, nos lemmes étant ceux du [DMF](#)), quand l'étiquette morphosyntaxique est le fruit d'un début d'analyse grammaticale en parties du discours.

- Le cas des formes contractées

Exemple : la forme médiévale *ou* peut être appareillée de quatre paires lemme-étiquette différentes : [OÙ](#) PROrel, [OÛ](#) PROint, [OU1](#) CONcoo, ou encore [EN1](#) & [LE](#) PRE.DETdef. Dans cette dernière hypothèse, *ou* est la forme médiévale contractée de la préposition *en* et de l'article défini *le*, le format de l'étiquette grossit alors pour s'adapter et est du type AAA.AAAaaa.

En faisant une recherche par lemme *via* LEMMES RECHERCHE ou *via* CONCORDANCIER, les formes contractées de deux lemmes (dites multi-lemmes) apparaissent aussi, bien qu'on ne recherche qu'un seul des deux lemmes.

Nomenclature des lemmes

Accueil Texte Formes Lemmes Crédits et contact

Recherche Nomenclature Cooccurrences Concordancier Index lemmatisé

mots lexicaux noms propres mots étrangers mots avec lemme réservé

<< A B C D E F G H I J L M N O P Q R S T U V Y >>

EN1, prép. 311 5 [e]n [en] en es ou

Ci-contre, dans l'onglet noir LEMMES et l'onglet jaune NOMENCLATURE, les formes contractées *es* et *ou* sont bien référencées comme attestations de la préposition [EN1](#).

Cependant, dans le CONCORDANCIER, en faisant une recherche par étiquette morphosyntaxique avec DETdef ou seulement DET, la forme contractée PRE.DETdef n'apparaîtra pas dans les résultats. Cette limite actuelle sera corrigée à l'avenir.

• Quelques étiquettes morphosyntaxiques utiles

- NOM : NOMcom NOMpro
- ADJ : ADJqua ADJndf
- DET : DETpos DETdem DETdef DETndf
- PRO : PROdem PROper PROadv PRrel PROint
- VER : VERcvg VERinf VERppe VERppa
- ADV : ADVgen ADVneg
- PRE
- CONsub CONcoo

forme	lemme (articles DMF)	étiquette (jeu minimal CATTEX2009)	
<i>Pourras</i>	μPORT-ROYAL <i>(ce n'est pas un article du DMF bien sûr, mais un lemme ad hoc est quand même créé pour l'édition électronique)</i>	NOMpro	
<i>escollier</i>	ÉCOLIER	NOMcom	
<i>griefve</i>	GRIEF2	ADJqua	
<i>ce</i>	CE1	DETdem	
	CE2	PROdem	
<i>peu</i>	POUVOIR	VERppe	
	PEU	ADVgen	
<i>y</i>	Y2	PROadv	
<i>que</i>	QUE	PROrel	
		PROint	
		CONsub	
<i>ne</i>	NE1	ADVneg	
	NE2	CONcoo	
<i>dont</i>	DONT	PROrel	
		PROint	
	DONC	ADVgen	
formes contractées			
<i>ou</i>	« en+le »	EN1 & LE	PRE.DETdef
<i>es</i>	« en+les »	EN1 & LE	PRE.DETdef
<i>au</i> <i>aux</i>	« à+le » « à+les »	À & LE	PRE.DETdef

c) ... avec les codes grammaticaux

- Code grammatical (lexicologie) ≠ étiquette morphosyntaxique (édition électr.)

Le recours aux codes grammaticaux est dans une large part redondant avec le paramètre des étiquettes morphosyntaxiques, car ces deux paramètres ressortissent à une analyse en partie de discours. C'est dans un souci d'exhaustivité que la possibilité de requête par le paramètre du code grammatical est ici présentée.

Les codes grammaticaux ont été créés par l'équipe du DMF pour le DMF, de façon à renseigner la partie de discours pour chaque vedette d'article (lemme) du DMF. Ainsi, à chaque lemme du DMF est associé un et un seul code grammatical. En revanche, l'étiquetage morphosyntaxique que nous réalisons est une couche d'information grammaticale propre à ce travail d'édition électronique. Les étiquettes morphosyntaxiques sont donc indépendantes des lemmes du DMF. Les cas, ci-dessous, d'une forme contractée et du morphème *que* illustrent bien la subtile différence entre le code grammatical et l'étiquette morphosyntaxique :

L'on peut, par exemple, faire la simple requête de *aux* dans l'onglet noir FORMES. Les formes *aux* et *au* sont la contraction de la préposition et de l'article. À cet égard, les deux lemmes **À** et **LE** sont attribués conjointement à ces formes. Le lemme **LE** du DMF est toujours complété par le code grammatical *art. déf.*, et le lemme **À** du DMF est toujours complété par le code grammatical *prép.*, car les codes grammaticaux sont inhérents aux articles lexicographiques du DMF. En revanche, l'étiquette morphosyntaxique est propre au travail d'édition électronique conduit sur un texte. Dans le cas des formes contractées, il a été arrêté un étiquetage qui en rend compte sous le format AAA.AAAaaa ; les formes *aux* et *au* ont donc reçu l'étiquette PRE.DETdef.

1	LE , art. aux déf. À , prép.	Considerant, de sens rassis, Le frain aux dens, franc au collier, Qu'on doit ses euvres conseiller,	mot 29 PRE.DETdef p.44 v.4
---	--	--	---

L'on peut, par exemple, faire la simple requête de *que* dans l'onglet noir FORMES. Le morphème *que*, quoique très riche et varié dans son fonctionnement linguistique, constitue une seule entrée lexicologique sous le lemme **QUE** du DMF. Le code grammatical associé au lemme **QUE** se doit donc d'être pluriel, assavoir *conj.*, *rel. interr.* Cependant, l'étiquetage morphosyntaxique propre à ce travail d'édition électronique discrimine la partie du discours associée, en l'espèce ci-dessous l'étiquette est CONsub.

8	QUE , conj., rel. interr. que	Du nombre des amoureux sains. VII Combien que le depart me soit Dur, si fault il que je l'eslongne ;	mot 342 CONsub p.46 v.49
---	--	--	---

adj.
adj. et adv.
adj. et pron. indéf.
adj. et subst.
adj. et subst. masc.
adj. indéf.
adj. num.
adj. poss.
art. contr.
art. déf.
art. indéf.
adv.
adv. et prép.
conj.
dém.
dét. dém.
interj.
loc. adv.
loc. conj.
nom propre
nom. indéf.
prép.
pron.
pron. adv.
pron. indéf.
pron. déf.
pron. dém.
pron. pers.
subst.
subst. fém.
subst. masc.
verbe

- Recourir au code grammatical dans une requête

Ci-contre une liste quasi-exhaustive des codes grammaticaux utilisés dans les articles du DMF. Il manque *rel. interr.*

Le recours au paramètre du code grammatical est possible pour le mot précédent ou le mot suivant dans le Concordancier. Il est possible de conduire une requête par l'exact code grammatical =*code* ou en précisant seulement un segment du code grammatical \subset code. La recherche par inclusion est bien plus ouverte que celle par égalité.

Deux remarques importantes :

- le code grammatical renseigne toutes les parties de discours possibles d'un lemme du DMF, et non la partie de discours pertinente de l'occurrence. Cf. le cas expliqué du morphème *que* dans le sous point précédent.
- pour s'assurer du code grammatical d'un lemme, se référer à l'article du DMF : son code grammatical associé suit le lemme en vedette écrit en lettres capitales.
- l'on constatera que la terminologie grammaticale peut diverger entre le code grammatical du DMF et l'étiquette morphosyntaxique de l'édition électronique. Soit le lemme [ÊTRE2](#) attribué aux attestations de l'infinitif substantivé : le code grammatical du DMF en est *subst. masc.*, quand l'étiquette morphosyntaxique associée aux attestations est NOMcom.

3. Aller encore plus loin

Les utilisateurs initiés aux expressions rationnelles/régulières (regex) verront que le recours à ce langage est possible dans le sous-onglet jaune RECHERCHE des onglets noirs FORMES et LEMMES. Le niveau d'application de la regex sera seulement dans l'unité fondamentale, assavoir un mot, une ponctuation.

Pour une requête avec des critères sur un segment supérieur à trois mots contigus, il est possible d'interroger la base Frantext. Le texte mis à disposition par Frantext pour les besoins de l'agrégation est le texte même de cette édition électronique. Cependant, sur Frantext, la lemmatisation et l'étiquetage morphosyntaxique n'ont été ni relus, ni corrigés humainement.

III. Autres informations utiles

A. Ressources du DMF pour l'agrégation 2021

1. Villon cité dans les articles du [DMF](#)

Le texte de Villon a fait partie du corpus à partir duquel les rédacteurs du DMF, sous la direction de Robert Martin, ont construit les articles du DMF dans les années 1990-2000. Ils s'étaient appuyés sur l'édition Droz de 1970 ; l'édition Champion 2004 n'en diverge que de peu. Le texte de Villon apparaît donc dans de nombreux articles du DMF. Les citations de Villon au sein des articles du DMF peuvent être recherchées et retrouvées rapidement au moyen de la commande clavier CTRL+F *Villon* — sans oublier au préalable de cliquer sur l'onglet COMPLET, en haut de l'article, de façon à afficher aussi les citations.

2. Synthèses grammaticales du DMF

Les articles de mots grammaticaux, rédigés par Robert Martin, sont des synthèses grammaticales précises, qui s'appuient sur les acquis récents de la linguistique. Elles rendent compte de l'état de langue du moyen français (MF 1350-1500). Ces articles peuvent être d'une grande utilité à l'agrégatif pour élaborer ses fiches de synthèse. À titre d'exemple, les articles de synthèse [NE1](#), [NE2](#), [QUE](#), [SI3](#), [SI4](#).

B. Les notes grammaticales ou philologiques propres à cette édition électronique (en exposant dans le texte)

T.746 p.126	On s'attendrait au coordonnant NE dans ce contexte négatif. Sans doute est-ce pour « prévenir la lecture d'un OU négatif ». cf. MARTIN & WILMET (1980) Syntaxe du Moyen Français, SOBODI, p.275 haut.
T.1514 p.170	« TROP équivaut à MAIS dans l'expression TROP BIEN ». cf. MARTIN & WILMET (1980) Syntaxe du Moyen Français, SOBODI, p.276 haut.

IV. Index

- lemme (DMF)
 - ▶ qu'est-ce qu'un lemme ? : I.
 - ▶ comment connaître la graphie conventionnelle qui caractérise un lemme ? : II.B.2. point n°3
- étiquette morphosyntaxique ([CATTEX2009](#)) : I. | II.B.2. point n°4 | II.C.2.b)
- code grammatical (DMF) : II.C.2.c)
- *Nomenclature* (fonctionnalité) : II.B.2. p.7 | II.C.1.
- *Concordancier* (fonctionnalité) : II.B.2. p.6 | II.C.2.
- *Recherche par forme* (fonctionnalité) : II.B.1.
- *Recherche par lemme* (fonctionnalité) : II.B.2.
- Liste
 - ▶ de l'ensemble des attestations d'un lemme classées par forme : II.B.2. p.6
 - ▶ de l'ensemble des attestations d'un lemme classées par ordre d'apparition : II.B.2. p.6
 - ▶ des différentes formes graphiques d'un lemme : II.B.2. p.7 (cf. *Nomenclature*)

Pour toute question ou suggestion relative à cette notice d'utilisation, écrire à jeanmichel.jezequel@atilf.fr.