



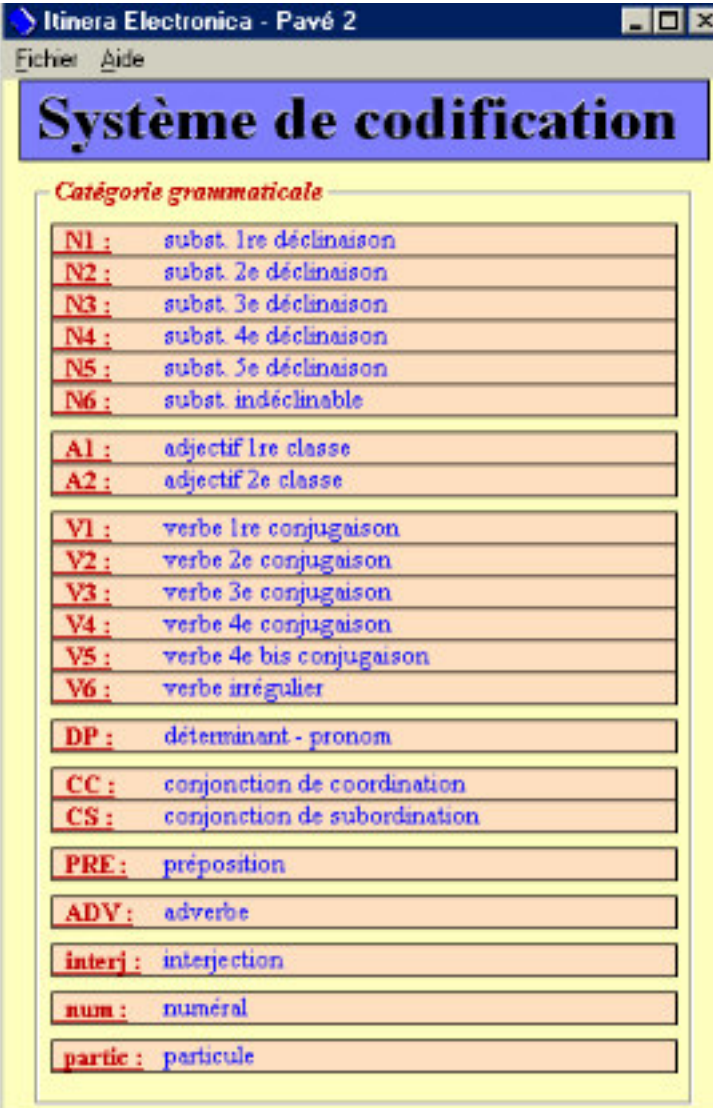
Módulo NooJ para Latim

*Workshop on Language Resources for
Teaching and Research*

José Paulo da Costa Tavares

Porto, 23 de Abril de 2008

Módulo NooJ para Latim



The screenshot shows a software window titled 'Itinera Electronica - Pavé 2' with a menu bar containing 'Fichier' and 'Aide'. The main content area is titled 'Système de codification' and lists various grammatical categories with their corresponding codes. The categories are grouped under the heading 'Catégorie grammaticale'.

Catégorie grammaticale	
N1 :	subst. 1re déclinaison
N2 :	subst. 2e déclinaison
N3 :	subst. 3e déclinaison
N4 :	subst. 4e déclinaison
N5 :	subst. 5e déclinaison
N6 :	subst. indéclinable
A1 :	adjectif 1re classe
A2 :	adjectif 2e classe
V1 :	verbe 1re conjugaison
V2 :	verbe 2e conjugaison
V3 :	verbe 3e conjugaison
V4 :	verbe 4e conjugaison
V5 :	verbe 4e bis conjugaison
V6 :	verbe irrégulier
DP :	déterminant - pronom
CC :	conjonction de coordination
CS :	conjonction de subordination
PRE :	préposition
ADV :	adverbe
interj :	interjection
num :	numéral
partic :	particule

Módulo NooJ para Latim

Itinera Electronica - Analyse morphologique

Fichier Edition Thesaurus Dictionnaire Aide

Analyse morphologique

Formes	Lemme	Traduction	Cat. gramm.	Cas, genre,	Mode,	Indic.
Gallia	-	-	-	-	-	*
est	esse, sum, fui	être; exister	V6	3 sg.	ind. prés. act.	
omnis	omnis, e	tout	A2	nom. m. sg.	-	
diuisa	diuido	dividre	V3	nom. f. 3 sg.	ind. parf.	*
in	in	en; dans; sur;	PRE	-	-	
partes	pars, partis	part; partie	N3	acc. f. pl.	-	
tres	tres, trium	trois	num	nom. f. pl.	-	
quarum	qui, quae,	qui, que, dont	DP	gén. f. pl.	-	
unam	unus, a, um	un seul	DP	acc. f. sg.	-	
incolunt	-	-	-	-	-	*
Belgae	-	-	-	-	-	*
aliam	-	-	-	-	-	*
Aquitani	-	-	-	-	-	*
tertiam	tertius, a, um	troisième	numéral	acc. f. sg.	-	
qui	qui, quae,	qui, que, dont	DP	nom. m. sg.	-	
qui	qui, quae,	qui, que, dont	DP	nom. m. pl.	-	
ipsorum	ipse, ipsa,	même, de	DP	gén. m. pl.	-	
lingua	lingua, ae (f)	langue;	N1	abl. f. sg.	-	
Celtae	-	-	-	-	-	*
nostra	noster,	notre, nos; le	DP	nom. nt. pl.	-	
nostra	noster,	notre, nos; le	DP	acc. nt. pl.	-	
Galli	-	-	-	-	-	*

Forme :
diuisa

Lemme :
diuido

Traduction :
dividre

Cat. gramm. :
V3

Cas, genre, nombre :
nom. f. 3 sg.

Mode, temps, voix :
ind. parf. pas.

P2 P3

Valider

Enregistrer la forme et les analyses dans le Thesaurus.

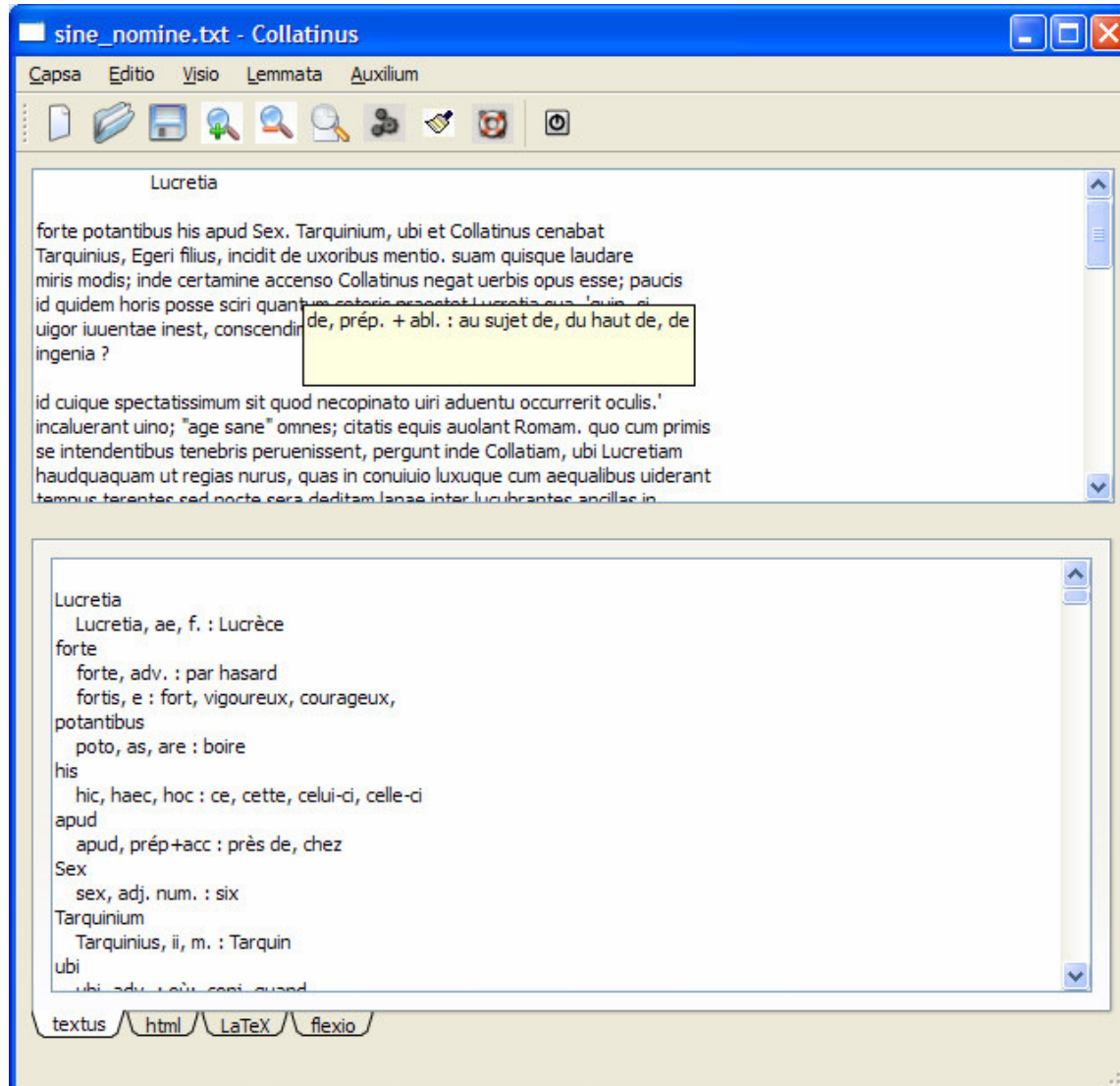


Módulo NooJ para Latim

- *Clavis Prophetarum, liber III*
 - 43153 itens
 - 18827 desconhecidos
 - 51433 anotações (24326 itens)

- **taxa de cobertura: 56%**

Módulo NooJ para Latim



The screenshot displays the NooJ software interface for a Latin text file named "sine_nomine.txt - Collatinus". The interface includes a menu bar with options: Capsa, Editio, Visio, Lemmata, and Auxilium. Below the menu is a toolbar with various icons for file operations and search. The main text area contains the following Latin text:

Lucretia

forte potantibus his apud Sex. Tarquinius, ubi et Collatinus cenabat Tarquinius, Egeri filius, incidit de uxoris mentio. suam quisque laudare miris modis; inde certamine accenso Collatinus negat uerbis opus esse; paucis id quidem horis posse sciri quantum ceteris praestet Lucretia sua. 'quis ei uigor iuuentae inest, conscendit de, prép. + abl. : au sujet de, du haut de, de ingenia ?

id cuique spectatissimum sit quod necopinato uiri aduentu occurrerit oculis.' incaluerant uino; "age sane" omnes; citatis equis auolant Romam. quo cum primis se intendentibus tenebris peruenissent, pergunt inde Collatiam, ubi Lucretiam haudquaquam ut regias nurus, quas in conuiuio luxuque cum aequalibus uiderant tempus terentes sed nocte sera deditam lapas inter lucubrantem ancillas in

The bottom section of the interface shows a list of lemmata (word forms) and their grammatical information:

- Lucretia
- Lucretia, ae, f. : Lucrèce
- forte
- forte, adv. : par hasard
- fortis, e : fort, vigoureux, courageux,
- potantibus
- poto, as, are : boire
- his
- hic, haec, hoc : ce, cette, celui-ci, celle-ci
- apud
- apud, prép+acc : près de, chez
- Sex
- sex, adj. num. : six
- Tarquinius
- Tarquinius, ii, m. : Tarquin
- ubi
- ubi, adv. : où, en, quand

At the bottom of the window, there are tabs for different output formats: textus, html, LaTeX, and flexio.



Módulo NooJ para Latim

- **Lexical Analysis,**
 - - *_properties.def* (organiza as anotações em categorias)
 - - *Clitics.nom* (permite o reconhecimento automático de formas acompanhadas de clíticos)
 - - *Latin.02.nod* (dicionário electrónico)
- **Projects**
 - - *De bello gallico.nop* (obra de César anotada)
- **Syntactic Analysis**
 - - *NoloComp.nog* (reconhece as formas compostas de *nolo*)
 - - *PassiveComp.nog* (consruída com o intuito de reconhecer formas compostas de verbos na voz passiva)



Módulo NooJ para Latim

<i>Etiqueta</i>	<i>Categoria</i>	<i>Exemplo</i>
A	Adjectivo	aeneo,aeneus, A +FLX=Aeneus+pos+ab+s+m
ADV	Advérbio	altius,alte, ADV +FLX=Alte+comp
CONJ	Conjunção	et, CONJ
INT	Interjeição	o,o, INT
N	Nome	rosarum,rosa, N +FLX=Rosa+gen+p
PREP	Preposição	ad, PREP
PRO	Pronome	me,ego, PRO +pes+ac+s+m
V	Verbo	amare,amo, V +FLX=Amo+INF+Prés+Act
NUM	Numeral	tribus,tres, NUM +card+FLX=Tres+ab+s+n



Módulo NooJ para Latim

- *uniter, ADV*
- *cum, PREP*
- *ehem, INT*
- *ergo, CONJ*

- *suave, ADV+FLX=Alte*
- *ditius, ADV+comp*
- *melius, bene, ADV+comp*



Módulo NooJ para Latim

○ Nomes

- Forma canónica
 - nominativo do singular
- *basiliscus, N+FLX=Dominus+m*
- *sapinus, N+FLX=Dominus+f*

- *confusio, N+FLX=Labor+f*
- *confusionis, confusio, N+FLX=Laboris+f*

- *Abdera, N+pr+FLX=Africa+f*



Módulo NooJ para Latim

○ Numerais

- *centum, NUM+card*
- *decem, NUM+card*
- *duodecim, NUM+card*

- *unus, NUM+card+FLX=Unus*



Módulo NooJ para Latim

○ Adjectivos

- Nominativo singular masculino
- *arboreus, A+FLX=Aeneus*
- *audax, A+FLX=Felix*
- *audacis, audax, A+FLX=Felicis*



Módulo NooJ para Latim

○ Verbos

- *sum, V+FLX=Sum*
- *censeo, V+FLX=Moneo+t*
- *bellor, V+FLX=Hortor+dep*
- *rudo, V+FLX=Decedo+x*

- *lego, V+FLX=Lego+t*
- *legi, lego, V+FLX=Amavi*
- *lectum, lego, V+FLX=Amatum*



Módulo NooJ para Latim

- *adeo, V+FLX=Eo*
- *aditum, adeo, V+FLX=Amatum*
- *adii, adeo, V+FLX=Amavi*
- *adivi, adeo, V+FLX=Amavi*

- *adplico, V+FLX=Amo+t*
- *adplicavi, adplico, V+FLX=Amavi+t*
- *adplicitum, adplico, V+FLX=Amatum+t*
- *adplicui, adplico, V+FLX=Amavi+t*
- *adplicatum, adplico, V+FLX=Amatum+t*

- *occido, V+FLX=Lego+t*
- *occido, V+FLX=Decedo+x*



Módulo NooJ para Latim

- Dicionário:
 - 68694 entradas
 - 49852 lemas
 - 104 regras flexionais
 - 2912863 formas flexionadas



Módulo NooJ para Latim

- *Clavis Prophetarum*

- 11936 formas diferentes
- 1127 desconhecidas

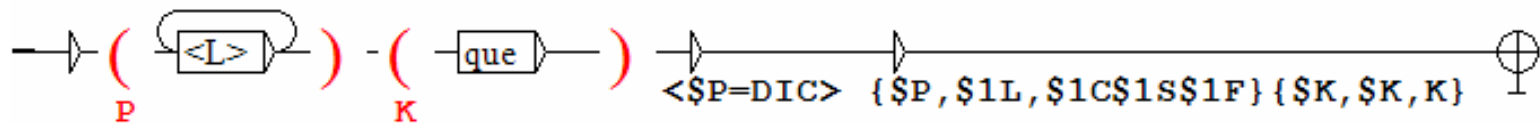
90,56%.

- *De Bello Gallico*

- 1582 formas diferentes
- 61 desconhecidas

96,2%.

Módulo NooJ para Latim





Módulo NooJ para Latim

Endereços web

- NooJ
 - www.nooj4nlp.net
- VISL
 - visl.hum.sdu.dk/visl/la/
- Analyse Morphologique
 - pot-pourri.fltr.ucl.ac.be/itinera/Hors_ligne/Analyse/
- Collatinus
 - www.collatinus.org
- Jspell
 - linguateca.di.uminho.pt/jspell