

ATHIS. II^e Atelier Informatique et Histoire

« L'historien, le texte et l'ordinateur »

Ecole Normale LSH de Lyon, 27-28 novembre 2006

Table ronde sur les problèmes de la lemmatisation

Les problèmes de la lemmatisation : un corpus de textes hagiographiques

Caroline Philippart de Foy

Pour cette table ronde portant sur les problèmes de la lemmatisation, je voudrais vous parler des difficultés que j'ai personnellement rencontrées lors de la lemmatisation d'un corpus de textes du moyen âge.

Je travaille sur des textes hagiographiques et plus précisément des vies et des passions latines traduites du grec. Le corpus que j'ai défini est composé de 36 textes, rédigés en Italie, principalement à Rome et à Naples, à des époques différentes (allant du V^e au XI^e siècle). Le choix des textes a surtout dépendu des éditions existantes, mais, pour permettre différents types de comparaisons, j'ai veillé à sélectionner plusieurs textes de même région, de même époque, de même auteur, dans les cas où il est connu, et, chaque fois que c'était possible, différentes traductions d'un même texte.

Mon idée est d'étudier ce qui caractérise ces textes, de voir ce qu'ils ont en commun et de savoir si l'on peut établir des critères de traduction propres à une époque, une région, une école ou un traducteur. C'est dans ce but que j'ai choisi d'utiliser les méthodes quantitatives de la statistique, qui permettent de prendre en considération un grand nombre de textes et de relever ce qui les caractérise.

La première étape du travail fut donc la préparation des textes à ces traitements quantitatifs : leur numérisation et leur lemmatisation selon les méthodes du LASLA, dont on vient de parler.

La lemmatisation exige une lecture attentive et minutieuse des textes. S'il est vrai que le travail est facilité par le système d'analyse semi-automatique, il faut tout de même vérifier tous les mots et corriger ou compléter ce qui n'a pas été analysé correctement ou automatiquement. C'est ainsi que le travail de lemmatisation met déjà à la lumière toute une série de points, observations ou problèmes, qui à la simple lecture peuvent passer inaperçus.

A côté des questions générales soulevées par toute lemmatisation, j'ai rencontré d'autres problèmes plus spécifiques, qui résultent de la nature du corpus.

Tout d'abord, celui-ci, comme je l'ai dit, est formé d'une trentaine de textes rédigés à différentes époques et par différents auteurs : donc, bien qu'homogène par le sujet, il présente une grande diversité dans la langue. Or, ce n'est pas la même chose de lemmatiser l'œuvre d'un seul auteur et celle de plusieurs auteurs dispersés dans le temps et dans l'espace. Il est primordial de se fixer des normes et de maintenir une cohérence tout au long de l'analyse, ce qui dans le cas présent n'a pas manqué de poser problème, car, chaque fois que j'essayais de me fixer une ligne de conduite, elle était remise en question par l'analyse d'un nouveau texte.

Ensuite, il s'agit de textes du moyen âge, écrits dans un latin différent du latin classique habituellement étudié au LASLA. Cela n'est pas un problème en soi, puisque des textes du moyen âge y ont déjà été lemmatisés et que, pour ce faire, des modifications ont été apportées en vue de faciliter la reconnaissance automatique des mots, notamment des *ae* ou *oe* écrits en *e* et inversement. La difficulté vient ici en réalité de la « qualité » des textes. Ce sont

des oeuvres hagiographiques, des vies et des passions de saints, des textes qui souvent accordent plus d'importance à leur contenu qu'à leur forme, des textes qui sont traduits du grec et pas toujours dans un latin très correct. On sait qu'il existe au moyen âge des confusions dans l'emploi des déclinaisons et des conjugaisons, qu'il y a un relâchement dans l'utilisation des cas, des modes, des temps, des voix, etc. J'ai rencontré toutes ces difficultés à des degrés divers dans le corpus. Les textes qui présentent le plus de problèmes sont évidemment ceux qui ont été traduits de manière très littérale, car la traduction au mot à mot a entraîné des erreurs et des constructions étranges, calquées sur le grec, qui rendent le texte peu clair, voire dans certains cas incompréhensible, et par conséquent très difficile à lemmatiser.

J'ai donc été confrontée à des problèmes d'ordre morphologique aussi bien que syntaxique. Comment procéder en effet quand on se trouve face à des formes comme *ferrus* (au lieu de *ferrum*), *saxos* (au lieu de *saxa*) ou *dulciam* (au lieu de *dulcem*) ? Comment traiter un verbe qui change de conjugaison à l'intérieur d'une même œuvre ? Comment analyser un *dico quia* ou un accusatif qui dépend de la préposition *cum* ?...

Mon propos étant d'étudier ce qui caractérise ces traductions, j'ai recherché la manière de lemmatiser qui rendrait compte au mieux de ces particularités linguistiques. C'est ainsi que, dans certains cas, j'ai choisi de m'écarter de la manière habituelle de procéder du LASLA.

Prenons en premier lieu la morphologie. Les difficultés résultent principalement du flottement qui existe dans les déclinaisons et les conjugaisons. (**Exemples 1**)

Face à des formes inhabituelles du type de *saxos*, j'avais deux possibilités :

- ou bien créer de nouveaux lemmes pour respecter la déclinaison utilisée (sans oublier que ce n'est pas toujours la même tout au long d'un texte)
- ou bien classer ces formes sous le lemme normal, en notant la particularité de la terminaison dans le code numérique.

Plus concrètement : pour la forme *saxos*, j'avais le choix entre créer un lemme *saxus* qui rendrait compte de sa forme de masculin, bien qu'elle ne soit à ma connaissance pas attestée dans les dictionnaires, ou la ranger sous le lemme normal *saxum*, tout en gardant la forme *saxos*. Ici le problème n'est pas encore très compliqué, car, comme on reste dans une deuxième déclinaison et qu'il n'y a pas dans le code d'indication du genre pour les substantifs, la codification reste toujours la même (12).

Prenons en revanche le cas de *tonitrua*. Les dictionnaires donnent *tonitrus, us* (4^{ème} déclinaison) et *tonitruum, i* (2^{ème} déclinaison). Dans le corpus, on trouve deux occurrences de *tonitruum* (une au singulier, l'autre au pluriel), mais il y a aussi une forme *tonitrua* pour un nominatif singulier. (**Exemples 2**)

Le mot est décliné comme s'il appartenait à la première déclinaison. Une confusion entre ces deux déclinaisons est assez facile à comprendre, par un glissement de la forme neutre pluriel *tonitrua* vers la forme du féminin singulier *tonitrua*. Reste à trouver une solution pour lemmatiser. Si je crée un nouveau lemme (*tonitrua, ae*), je crée un doublet dans le corpus, puisque ailleurs j'ai lemmatisé *tonitruum, i*. Si en revanche je considère qu'il s'agit plutôt d'une erreur de déclinaison et que je décide de ranger cette forme sous le lemme *tonitruum*, alors je mêle sous un même lemme des formes de première et de deuxième déclinaison, donc deux codes différents (11 et 12). Si cela ne gêne pas le décompte des occurrences des catégories grammaticales (ici substantifs) ou des sous-catégories (substantifs de la première déclinaison ou de la deuxième), cela complique par contre le décompte des lemmes par sous-catégories. Des cas comme *tonitrua* posent donc problème, quelle que soit la

solution adoptée : ou bien je multiplie le nombre de lemmes, ou bien je complique leur décompte.

En tenant compte du caractère spécifique du corpus et des flottements que l'on connaît dans l'emploi des déclinaisons au moyen âge, il ne m'a pas paru opportun de créer de nouveaux lemmes pour classer ces formes étranges. Il me semble en effet que des formes comme *saxos*, *ferrus* ou *tonitrua* s'expliquent plus par une confusion de genres et de déclinaisons que par l'existence d'un terme *saxus*, *ferrus* ou *tonitrua*. Par conséquent, j'ai choisi de ranger sous un seul et même lemme (celui du dictionnaire) les formes régulières et les formes irrégulières, en gardant toutefois trace du changement de déclinaison dans le code numérique.

J'ai procédé de la même manière pour les adjectifs. Ainsi, j'ai classé sous le lemme *dulcis* les formes régulières de l'adjectif de la deuxième classe, avec le code 24 (adj. de la 2^{ème} classe en *-is*), mais aussi la forme *dulciam*, à laquelle j'ai donné le code d'un adjectif de la première classe (21). De même, j'ai rangé la forme *cognatibus* sous l'adjectif *cognatus*, la forme *lugubram* sous l'adjectif *lugubris*, etc. **(Exemples 3)**

Cela m'a paru être le meilleur choix à faire, en tout cas celui qui permettait de garder les particularités morphologiques sans multiplier inutilement les lemmes. Mais procédant de cette manière, j'en suis venue à m'interroger sur les distinctions faites dans les dictionnaires pour des mots comme *caelum* (au neutre singulier) et *caelus* (au masculin singulier) ou *inferni* (au masculin pluriel) et *inferna* (au neutre pluriel), ou encore *cingula* (au féminin), *cingulus* (au masculin) et *cingulum* (au neutre), etc. On sait en effet que certains mots sont déjà utilisés sous différentes déclinaisons en latin classique et, dans ce cas, le LASLA suit les dictionnaires et donne deux ou trois lemmes différents. Mais dans la mesure où l'emploi des déclinaisons est devenu plus lâche au moyen âge et notamment dans ce corpus, je me suis interrogée sur la pertinence de reproduire ici ces distinctions. Et vu que je mêlais déjà des codes différents sous un même lemme, j'ai choisi d'appliquer le même traitement à ces mots.

Toutefois la mise en œuvre de ce principe n'a pas toujours été évidente. Quand on trouve des formes bizarres telles que *saxos*, on réalise tout de suite qu'il y a un problème. Il en va tout autrement pour des mots du type *inferni/inferna*. On lemmatisera en effet naturellement *inferni* dans un texte qui emploie toujours ce mot au masculin pluriel et *inferna* dans le texte qui l'utilise tout le temps au neutre pluriel. Mais, en analysant de cette manière, on aura en finale, sans vraiment s'en être rendu compte, deux lemmes pour le corpus : *inferni* et *inferna*. Dans la pratique, on constate le problème quand on relit les index ou quand on trouve dans un même texte des formes des deux genres utilisées indifféremment.

En ce qui concerne les verbes, un élément supplémentaire vient s'ajouter à la confusion occasionnelle des conjugaisons, c'est l'emploi des voix. Il y a là aussi un flottement par rapport à l'usage classique. On voit au moyen âge des verbes déponents devenir actifs ou passifs et inversement.

Tout comme pour les substantifs et les adjectifs, j'ai choisi de ranger sous un seul lemme des formes de différentes conjugaisons comme *tondeo* et *tondo*, avec un code 52 pour les formes de deuxième conjugaison et 53 pour les formes de troisième conjugaison. Mais j'ai aussi décidé de regrouper sous un même lemme les formes déponentes, actives et passives d'un verbe. Par exemple, j'ai regroupé sous le lemme *largior* les formes normales déponentes avec le code 5M (verbe déponent de la quatrième conjugaison), mais aussi les formes actives et passives sous les codes 54 et 5D. Cela m'a à nouveau paru plus logique de procéder de cette manière que de lemmatiser *largio* à côté de *largior*. **(Exemples 4)**

Ces questions de morphologie illustrent bien la problématique de la lemmatisation d'un texte et de celle d'un corpus, qui doit être la plus homogène possible. Même si tous ces cas ne concernent finalement qu'un petit nombre de mots et qu'ils n'auront probablement pas beaucoup de poids dans les calculs statistiques, il m'a tout de même paru important de les traiter de manière logique et cohérente.

En outre, si je me suis permis de m'écarter des principes du LASLA et de procéder à ces regroupements, auxquels on ne penserait probablement pas quand on lemmatise des textes classiques, c'est aussi parce que je sais que je travaille en vase clos et que je n'ai pas l'intention de comparer mon corpus à la base de données du LASLA. Il y aurait en effet peu d'intérêt à comparer ces Vies de saints aux œuvres de Cicéron ou de César, par exemple. Ce n'est ni la même époque, ni le même genre littéraire. Or on ne peut comparer que des choses qui sont comparables. Mon idée est d'étudier les textes du corpus les uns par rapport aux autres, de voir ce qui les caractérise et ce qu'ils ont en commun. Mon approche est donc interne : commencer par les textes signés et bien situés dans le temps et dans l'espace, rechercher leurs caractéristiques, voir si l'on peut établir des critères de traduction, des traditions localisées, etc., et ensuite y comparer les textes anonymes ou sur lesquels on ne possède que peu d'informations pour voir s'ils présentent des traits communs et si l'on peut les rattacher à l'un ou l'autre groupe défini. Les modifications que j'ai apportées à la lemmatisation ne présentent donc pas d'inconvénient, car l'important est que tous les textes pris en compte soient lemmatisés de la même manière.

D'un point de vue syntaxique, il y a dans le corpus des constructions atypiques et peu conformes à la grammaire classique. Par exemple, on trouve des impératifs dans des subordinées, relatives et autres, des interrogations indirectes construites à l'indicatif, voire à l'infinitif, des propositions infinitives là où on ne les attend pas et des subjonctifs là où on attendrait des infinitifs, des ablatifs absolus apposés au sujet du verbe principal, etc. ; la concordance des temps n'est pas non plus toujours respectée. Ces cas *a priori* dérangeants ne posent en réalité pas de véritable problème pour la lemmatisation, mais donnent simplement des résultats inhabituels. **(Exemples 5)**

Dans le dernier exemple (5 f), la proposition complétive est introduite par un subordonnant exprimé dans le texte, mais le verbe à l'intérieur de cette complétive est à la première personne. L'éditeur a manifestement été gêné par cette rupture de construction et a décidé de changer la ponctuation pour en faire une sorte de discours direct. Je pense toutefois que, puisque le subordonnant est bien exprimé dans le texte, il faut analyser la proposition comme une subordinée. Cette rupture de construction se rencontre d'ailleurs déjà dans le modèle grec du texte latin.

La ponctuation des textes peut du reste influencer plus profondément la manière de lemmatiser. Je pense en particulier au pronom relatif.

Quand on lemmatise un pronom relatif, on donne son analyse morphologique ainsi que, comme pour tout subordonnant, le mode et le temps du verbe qu'il introduit et on ajoute au verbe subordiné le code du relatif.

Mais que faut-il faire quand on se trouve face à un pronom relatif placé après une ponctuation forte ? quand il est mis en tête d'une phrase qui est présentée comme une nouvelle phrase et non comme une subordinée dépendant de la précédente ? Faut-il systématiquement l'analyser comme un relatif de liaison, c'est-à-dire comme l'équivalent d'un démonstratif accompagné d'une conjonction de coordination ? et par conséquent analyser le verbe de la phrase comme un verbe principal ?

La théorie des relatifs de liaison a longtemps été discutée. L'usage est reconnu, mais il est mis en question par certains philologues qui considèrent que l'on recourt un peu vite à cette explication quand on se trouve face à des phrases trop longues et difficiles à traduire dans une langue moderne.

Dans la pratique, quand on lemmatise, je pense que notre perception des relatifs de liaison, et donc leur analyse, est influencée par la ponctuation du texte. Or celle-ci dépend pour beaucoup de l'éditeur. Je me suis rendu compte de cette difficulté quand j'ai trouvé dans des textes différents des formules parallèles, mais ponctuées différemment : certains relatifs étant placés en tête de phrase et d'autres présentés comme de vrais relatifs subordonnants. Le corpus est constitué de textes édités de différentes manières, par des éditeurs différents, dans des pays et à des époques différents. Les normes d'édition ne sont donc pas les mêmes dans tout le corpus et le système de ponctuation varie d'un texte à l'autre. Par conséquent, il m'est arrivé de m'écarter de la ponctuation de l'édition d'un texte pour maintenir une analyse cohérente dans tout le corpus. **(Exemples 6)**

Par ailleurs, en tenant compte de la possible évolution de l'emploi du relatif au moyen âge, j'ai choisi de lemmatiser les relatifs comme des relatifs de liaison dans quelques cas précis :- quand le texte est une traduction très littérale du grec et que le pronom relatif latin traduit un pronom démonstratif grec.

- quand il y a clairement une rupture de construction et que l'on ne voit pas comment expliquer le relatif autrement que dans une nouvelle phrase (ablatif absolu du type *quo facto* ou *quo audito*, expression figée, *cum* historique, dialogue, ...).

Le relatif est par ailleurs employé dans certains textes du corpus pour traduire l'article grec. Dans une version de *La vie d'Anastase le Perse* qui traduit si littéralement son modèle grec qu'elle en respecte l'ordre des mots, les relatifs sont même utilisés dans ce cas sous des formes figées, et cela à côté d'emplois normaux. *Qui* est employé pour l'article masculin, *quae* pour le féminin et le neutre, qu'il s'agisse de singulier ou de pluriel, et très souvent aussi sans introduire de verbe. J'ai choisi de traiter ces formes figées avec les autres relatifs plutôt que de créer un lemme à part. **(Exemples 7)**

Pour terminer, examinons le cas des prépositions. A côté des prépositions courantes, on rencontre des adverbess devenus prépositions (*foris, foras, ..*). **(Exemples 8)** Par ailleurs, on trouve aussi pour les prépositions classiques des constructions inhabituelles, comme un accusatif régi par un *cum* ou un *de*, un génitif par un *de* ou un *ex*, un ablatif par un *ad*, Pour ces cas, j'ai à nouveau adapté l'analyse du LASLA : je les ai rangés sous leur lemme respectif *cum*, *de*, *ex* ou *ad*, en changeant le cas régi dans le code numérique : à côté de l'habituel *de* et ablatif 70600, on trouve donc *de* et accusatif 70300, *de* et génitif 70400. **(Exemples 9)**

Le dernier exemple (9 f) illustre encore une fois la difficulté de lemmatiser certaines phrases. Faut-il en effet considérer *fallacia* et *blandimenta* comme des accusatifs neutres pluriels dépendant de la préposition *cum* (le substantif *blandimentum* et l'adjectif *fallax* substantivé ou le substantif *fallacium* signalé dans Du Cange) ? Ou bien faut-il considérer que *blandimenta* est un ablatif singulier, mais qu'il a été décliné comme s'il appartenait à la première déclinaison par erreur, erreur qui serait par ailleurs facilement explicable par la présence du *cum fallacia* (alors substantif de la première déclinaison) et aussi du *cum fiducia* un peu plus haut ?

Un cas comme celui-ci ne peut être résolu qu'en replongeant dans le texte pour voir comment les prépositions y sont utilisées, s'il y a d'autres erreurs de déclinaison, ... Donc, même si l'on essaie de maintenir une analyse homogène pour tout le corpus, il ne faut pas oublier que les choix restent faits au cas par cas, texte par texte. On rappellera aussi que, quand on lemmatise, on doit faire entrer le texte et chacun de ses mots dans un code. Et là où

un éditeur pourrait donner deux explications dans une note, on ne peut en choisir qu'une. Il est dès lors très important lors de la lemmatisation de toujours noter les cas qui posent problème et les choix que l'on a faits, pour pouvoir y retourner par la suite.

La lemmatisation est donc un travail important. Elle oblige à examiner les textes de très près. A elle seule, elle soulève de nombreuses questions et met en évidence des points qui mériteront d'être étudiés plus amplement. Si ceux-ci relèvent davantage de la philologie que de l'histoire, ils pourront cependant avoir des implications historiques, en mettant à jour par exemple les rapports des textes, l'existence d'écoles ou de traditions localisées.

Chaque corpus peut apporter ses problèmes spécifiques et la personne qui lemmatise doit trouver les solutions qui conviennent le mieux au corpus et à l'étude envisagée. En ce qui concerne mon propre corpus, j'ai essayé de préserver au mieux les particularités linguistiques de chaque texte, tout en maintenant une analyse homogène tout au long du travail.

Les problèmes que j'ai rencontrés proviennent donc d'une part de la diversité des textes (éditions différentes, auteurs différents, époques différentes, ...) et de la difficulté de maintenir cette cohérence sur l'ensemble du corpus, et d'autre part de la problématique de la langue (latin médiéval, confusion dans les déclinaisons et les conjugaisons, ...) et de l'application d'une méthode d'analyse de textes classiques à des textes médiévaux traduits du grec qui s'écartent souvent de la norme classique.

Exemples

Exemples 1

a) Non solum autem hoc, sed et **saxos** magnos quos quattuor ex eis baiulantes imponebant ei ... (*Passio Anastasii Persae*, § 20, l. 267)

b) Videns uero iudex quia nihil plus exigit - sed per huiusmodi magnis doloribus sicut **ferrus** frigidi aquae limabitur diuina gratia corroboratus - iussit eum mitti in carcere donec iterum intimaret imperatori de eo. (*Passio Anastasii Persae*, § 35, l. 480)

Exemples 2

a) Sic uero intonuit uox illa quasi **tonitruum** fuisset, ita ut omnes conterriti erigerent se, etiam et ipse imperator. (*Vita Amphilochii*, ms. Würzburg, § 5, l. 28)

b) Cumque custodes secundum praeceptum imperatoris obseruarent ad locum, subito factae sunt tenebrae magnae, **tonitrua**, fulgora et inundationes aquarum sicut diluuium ingens, ita ut perterriti ministri diaboli effugarentur. (*Passio Marini pueri*, § 19, l. 31)

c) Compleuerunt orationem episcopi et omnes respondentes amen accesserunt ad gloschomum ; facta est **tonitrua** in aera, ita ut omnes caderent de timore eius. (*Passio Febroniae*, § 41, phr. 1)

Exemples 3

a) «Quare uis sicut unus horum male perire ? Sed magis adquiesce imperatori et uiue et accipe honores ab eo et esto in palatio sicut unus ex nostris et ne derelinquas **dulciam** uitam hanc.» (*Passio Anastasii Persae*, § 38, l. 521)

b) Sequenti autem die diffamatum est quod accidit per totam ciuitatem ; concurrentibus ad eam undique **cognatibus** et amicis et qui iuxta eorum domum habitabant, et mirabantur in dictis, et in stupore Deum laudabant in gloriosa liberatione puelle. (*Passio Abbibi*, 2, § 34, l. 11)

c) Surgens autem festinanter comprehendit uenerabile templum et inuenit filiam adsistentem mansionario et **lugubram** uestem circumindutam et nil distans a peregrina ;... (*Passio Abbibi*, 2, § 33, l. 9)

Exemples 4

a) Deinde consuetas super eum orationes dicens et quaeque monachorum lex praecipit faciens, tunc **tonduntur** iuuenis crines et exuens etiam omnem saecularem actum induitur monachicum indumentum, annumeraturque cum fratribus. (*Vita Iohannis Calybitae*, § 7, l. 16)

b) Tu itaque, dilectissime, utriusque lingue studium exercens, si quid perperam dixerim, stilo fraterno corrigere studeas ; si quid bene, redemptori laudes referamus, qui nobilis gratiam interpretandi secundum apostolum **largire** dignatus est. (*Passio Abbibi*, 3, l. 11)

c) Et sic fabricata est in eodem loco basilica : in qua orationes eius florent et beneficia **largiuntur** omnibus contingentibus sepulchrum eius per christum dominum nostrum ... (*Passio Luciae*, p. 109, l. 29)

d) Stupebat autem Zosimas, audiens quoniam et canonem monasterii inscia quasi quae nosset dicebat ; nihilque aliud clamabat nisi gloriam Dei, qui maiora quam petitur diligentibus se **largitur**. (*Vita Mariae Aegyptiacae*, § 20, l. 36)

Exemples 5

a) Beatus uir Dei dixit : «Deus, **qui cognosce** hominem antequam natus sit, tu et pro ista ancilla cogita ut saluetur, ut tibi placeat, et habeat partem cum sanctis tuis.» (*Vita Euphrosynae*, § 4, l. 21)

b) Tunc precepit Licinius imperator Ausonio principi **ut** : «Qui resisterit contra preceptum nostrum, igni eum **trade**». (*Passio Abbibi*, 1, § 5, l. 1)

c) Eunomius dixit : «cras coram imperatore ac senatu disputemus». Et erat tota ciuitas suspensa in audientiam Eunomii et Amphiloerii et dicebant ciues ad inuicem : «putas **quis uincet** ? Iste an ille ?». (*Vita Amphiloerii*, ms. Mantoue, § 4, l. 35)

d) Frater diacone somnium uidi et ualde turbatus sum quia nescio **quomodo discernere** eum. (*Passio Pelagiae*, § 14, l. 81)

e) ... et inuicem se amplexantes, multas permanserunt horas lacrimantes et minime ualentes aliquid loqui, **existentibus** in ululatu, ut etiam qui aduenerant mirabantur uidentes eas in tanto dolore lacrimantes. (*Passio Abbibi*, 2, § 33, l. 11)

f) Unus qui stabat ibi miles adnunciauit Seleno dicens **quia** «Iuuenculam **inuenimus** in monasterio illo nimis speciosam ; de ea loquitur Primus comes ut sit domino meo Lisimacho in coniugio». (*Passio Febroniae*, § 17, phr. 9)

Εἷς δὲ τῶν κακίστων στρατιωτῶν δραμὼν ἀπήγγειλεν τῷ Σελίνῳ λέγων ὅτι· νεάνιδά τινα εὐρήκαμεν ἐν τῷ μοναστηρίῳ σφόδρα περικαλλῆ καὶ ...

Exemples 6

a) Haec autem bonum certamen certauit, cursum consummauit, fidem seruauit, Christum manifestauit et accepit coronam gloriae a Christo Iesu domino nostro, **cui** est honor et gloria et potestas et imperium per infinita saecula saeculorum. (*Passio Anastasiae*, § 9, l. 25)

b) Et sic fabricata est in eodem loco basilica : in qua orationes eius florent et beneficia largiuntur omnibus contingentibus sepulchrum eius per christum dominum nostrum : **cui** est honor et gloria uirtus et potestas in saecula saeculorum. (*Passio Luciae*, p. 109, l. 29 ; cfr exemple 4 c)

c) Et postquam surrexit, rogauit eum, ut conbureret illam nefandissimam cartulam, **quod** et factum est. Et uidentes populi execrabile chirographum et negatoriam cautionem combustam igni, ceperunt cum multitudine lacrimarum diutissime clamare : «Kyrieleison». (*Poenitentia Theophili*, § 39, l. 2)

d) Precepit autem magister militum [ut] coram omni exercitu et populo, presente et illo Gotho, ut legeret<ur> predicta doctrina. **Quod** et factum est. Et cum audiret omnis multitudo quod in doctrina haberet, interrogant Gothum : «Sic se ueritas habet ?». (*Passio Abbibi*, 2, § 43, l. 8)

e) Talibus ergo documentis instructus gratias egit omnipotenti deo, eo quod dignus esset celesti uisitatione muniri ; sicque in sanctae religionis amore solidatus tantam constantiam animus illius ex sanctae trinitatis symbolo accepit, ut contra omnes hostes sanctae Ecclesiae solus paratus esset procedere ad dimicandum ; **quod** et fecit. (*Vita Gregorii Thaumaturgi*, II, 58, l. 5)

Exemples 7

a) Vnigenitus filius ... **qui** coequalis atque consubstantialis ... (*Passio Anastasii Persae*, § 1, l. 4)

Ὁ μονογενὴς Υἱὸς ... ὁ συναίδιος καὶ ὁμοούσιος ...

b) Sed qui humanae uitae a principio insidiator et fascinator atque dolosus et piorum inimicus non habens qualiter induceret in tantam bonorum **quae** per Christi gratia in hominibus magnale donum, ...(*Passio Anastasii Persae*, § 3, l. 29)

... τῆς διὰ Χριστοῦ χάριτος ...

c) ... et rursum proprium **quae** pietatis est ornatum Christi recipiebat aecclesia,... (*Passio Anastasii Persae*, § 3, l. 41)

... πάλιν τὸν οἰκεῖον τῆς εὐσεβείας κόσμον ...

Exemples 8

a) Spiculator uero perduxit eam **foras** ciuitatem ... (*Passio Anastasiae*, § 8, l. 10)

b) amputauit capita eorum **foris** muros Sebastiae ciuitatis super quendam lapidem ... (*Passio Blasii*, p. 153, l. 25)

Exemples 9

a) Et iussit abbas foras de monasterio ...manere patrem **cum** puellam. (*Vita Euphrosynae*, § 4, l. 6)

b) ...aliqui uero **de** gentiles nunciauerunt de monasterio illo Seleno, ... (*Passio Febroniae*, § 16, phr. 2)

c) Et hoc facto, gaudium magnum factum est uniuersis fidelibus et consolatio multa **de** frequentium eorum tribulationum, ... (*Passio Anastasii Persae*, § 29, l. 391)

d) Habens uero conuinctum, ut dictum est, iuuenem quendam **ex** seruorum pro quadam causa tentum et nolens contristare eum, ... (*Passio Anastasii Persae*, § 25, l. 327)

e) Abiit in pretorio **ad** Lisimacho. (*Passio Febroniae*, § 33, phr. 7)

f) Asclepiades, uerba mea cum fiducia audi, quoniam nos homines de confinio mortis eruimus ; nam uos **cum** fallacia et **blandimenta** eis mortem inpingitis. (*Passio Romani*, § 2, l. 6)

Editions

Passio Abbibi (BHL 8b, 7477b) : P. CHIESA, *Il dossier agiografico latino dei santi Gurias, Samonas e Abibos*, dans *Aevum*, 65 (1991), p. 238-253.

Vita Amphilochii (ms. Mantoue) : P. CHIESA, *Una traduzione inedita di Anastasio Bibliotecario ? Le «vitae» latine di sant'Anfilochio*, dans *Studi Medievali*, 28 (1987), p. 894-897.

Vita Amphilochii (ms. Würzburg) : P. CHIESA, *Una traduzione inedita di Anastasio Bibliotecario ? Le «vitae» latine di sant'Anfilochio*, dans *Studi Medievali*, 28 (1987), p. 898-903.

Passio Anastasii Persae (BHL 410b) : C. VIRCILLO FRANKLIN, *The Latin Dossier of Anastasius the Persian. Hagiographic Translations and Transformations*, Toronto-Ontario, 2004, p. 272-298.

Passio Anastasiae (BHL 404) : H. DELEHAYE, *La Passion de sainte Anastasie la Romaine*, dans *Studi dedicati alla memoria di Paolo Ubaldi*, Milan, 1937, p. 17-26, réédité dans Id., *Mélanges d'hagiographie grecque et latine*, Bruxelles, Société des Bollandistes, 1966 (*Subsidia hagiographica*, 42), p. 396-402.

Passio Blasii (BHL 1370) : MOMBRIUS, *Sanctuarium seu Vitae Sanctorum*, Paris, A. Fontemoing, 1910², I, p.150-153.

Vita Euphrosynae (BHL 2722) : A. BOUCHERIE, *La Vie de sainte Euphrosyne, texte romano-latin du VIII^e-IX^e siècle*, dans *Revue des langues romanes*, II (1871), p. 26-40.

Passio Febroniae (BHL 2844) : P. CHIESA, *Le versioni latine della Passio Sanctae Febroniae. Storia, metodo, modelli di due traduzioni agiografiche altomedievali*, Spolète, CISAM, 1990, p. 155-191.

Vita Gregorii Thaumaturgi (BHL 3677m, 3678, 3678d) : E. D'ANGELO, *Pietro Subdiacono napoletano. L'opera agiographica*, Florence, Sismel, 2002 (*Edizione nazionale dei testi mediolatini*, 7, serie I, 4), p. 151-182.

Vita Iohannis Calybitae : P. CHIESA, *Vita e morte di Giovanni Calibita e Giovanni l'Elemosiniere. Due testi 'amalfitani' inediti*, 1995, p. 29-45.

Passio Luciae (BHL 4992) : MOMBRIUS, *Sanctuarium seu Vitae Sanctorum*, Paris, A. Fontemoing, 1910², II, p.107-109.

Vita Mariae Aegyptiacae (BHL 5415) : ROSWEYDE, *Vitae patrum* ; P. L., 73, col. 671-690.

Passio Marini pueri (BHL 5538) : *Catalogus codicum hagiographicorum bibliothecae regiae Bruxellensis*, II, Bruxelles, 1886 (*Subsidia Hagiographica*, 2), p. 184-191.

Passio Pelagiae (traduction *sanctissimus* ou C) : F. GUILLAUMONT, dans P. PETIMENGIN (éd.), *Pélagie la pénitente. Métamorphoses d'une légende*, I. *Les textes et leur histoire*, Paris, Etudes augustinienes, 1981, p. 258-272.

Passio SS. Romani et sociorum (BHL 7299b) : H. DELEHAYE, *Saint Romain, martyr d'Antioche*, dans *Analecta Bollandiana*, 50 (1932), p. 261-270.

Poenitentia Theophili (BHL 8121) : G. G. MEERSSEMAN, *Kritische glossen op de griekse Theophilus-legende (7^e eeuw) en haar latijnse vertaling (9^e eeuw)*, dans *Mededelingen van de Koninklijke Vlaamse Academie voor Wetenschappen, Letteren en Schone Kunsten van België*, 25, 4 (1963), p. 17-32.