

LE LEL DE GALLARATE

EXPOSE PAR R. BUSA S. J.

Lorsque je parlai à Strasbourg en 1963, au troisième Congrès de "Vita Latina", j'avais dit que l'Index Thomisticus Total serait fini avant la fin de l'année 1965. Mais il n'en fut rien. Je pense maintenant que nous ne l'achèverons pas avant 1969. La raison principale de ce retard fut la suivante: la préparation du lexique électronique de langue latine a nécessité plus de temps que je ne pensais, lorsque je disais au Congrès, que j'étais entraîné par l'automation dans le gouffre de l'homographie.

Je peux aujourd'hui vous informer que notre lexique latin électronique, complet et systématique est achevé; il contient 170.000 formes de mots latins et l'expérience en a déjà démontré la validité, puisque nous l'avons déjà utilisé pour 10.500.000 mots.

J'exposerai donc 3 points: ce qu'est le lexique électronique latin et son but; son utilisation possible pour d'autres textes latins, et enfin sa composition.

NATURE ET BUT D'UN LEXIQUE LATIN ELECTRONIQUE

1. Le lexique latin électronique est une série de mots latins inscrits sur bande magnétique et codifiés avec des codes adéquats, afin que la machine puisse avec eux définir, distinguer et codifier automatiquement l'un après l'autre les mots d'un texte.

L'ordinateur a, en effet, été instruit pour comparer tous les mots de n'importe quel texte latin avec cette liste-là, dont les mots sont déjà garantis et codifiés. Cette comparaison a été faite jusque là dans 2 intentions, dont la deuxième est comme l'on verra facilement, de beaucoup la plus importante.

2. La première tâche est de rechercher les erreurs typographiques, qu'elles se trouvent dans le texte lui-même ou qu'elles proviennent de la perforation, de les séparer des vrais mots, bien que, hélas, on ne puisse pas ainsi trouver toutes les erreurs.

Une fois la comparaison faite, l'ordinateur fait une liste des mots, dont il n'a pas trouvé le correspondant dans le LEL. Parmi ces mots, certains seront des mots vrais, qu'il faut ajouter au LEL, d'autres, par contre, seront erronés. Par exemple, parmi les mots de la Somme Théologique qui ne se trouvaient pas dans le LEL, il y avait *istola et urando*: le contexte a montré que ces mots avaient été découpés de façon fautive et qu'il fallait lire *epistola et iurando*. Par contre, dans la Summa c. Gentiles le mot *co* ne se trouvait pas non plus dans LEL: mais le contexte montra qu'il n'y avait pas d'erreur, car St. Thomas y parle des mots qui commencent par les 2 lettres *co*, comme *cognosco* ou d'autres. Cette méthode ne permet pas de détecter toutes les erreurs, mais seulement les mots qui sont impossibles, ou du moins tout à fait improbables dans le texte latin. Par exemple, si le mot *quod* a été perforé *quot*, la comparaison avec le LEL ne dévoile pas l'erreur, car ces deux mots ont un sens en latin.

3. La seconde fonction du LEL est la lemmatisation automatique, dont nous parlerons plus largement.

Nous entendons par concordance l'index qui présente dans un certain ordre tous les contextes de chaque mot d'un texte, c'est-à-dire que sous la forme de chaque mot il présente toutes les phrases qui renferment cette forme.

C'est pourquoi les types de concordances peuvent s'étendre entre les deux extrêmes suivants. Mais auparavant, je dois donner notre définition de deux mots: forme et lemme.

La forme signifie, pour nous, le mot pris comme unité graphique, comme type d'une série de symboles graphiques.

Le lemme est le titre que l'on donne au paradigme d'un mot, dont la flexion présente plusieurs formes; c'est-à-dire le titre d'un mot pris comme unité à la fois graphique et sémantique.

Par exemple, lorsque je dis: "*sum, es, fui*", ces mots sont des flexions du verbe "*sum*"; "*sum, es, fui*" sont des formes, tandis que "*sum*" est un lemme, c'est-à-dire le titre du paradigme, qui représente donc ses nombreuses formes.

Ainsi le lemme *rapax, rapacis* a des formes comme *rapax, rapacis, rapa-*

ci, rapacem, rapace etc. . . .

4. On peut donc décrire ainsi les types extrêmes de concordances: dans le premier les contextes se présentent simplement dans l'ordre des formes; dans le second les contextes se présentent dans l'ordre des formes groupées lexicalement sous chaque lemme.

Selon le premier type, les formes *sum, sunt* se trouvent parmi les mots qui commencent par la lettre *s*; *es* et *est* parmi ceux qui commencent par *e*. Ainsi les formes *facies* seront toutes mélangées, qu'elles appartiennent au lemme *facies, faciei*, ou au lemme *facio, facere*.

Au contraire, dans le second type de concordances, les emplois du mot homographe *facies* se trouvent divisés en deux lemmes, et sous le lemme *sum* nous aurons même des formes qui commencent par *e* ou *f*. En outre dans le même lemme, on rassemble l'un après l'autre les différents emplois logiques possibles que chaque lemme peut avoir dans chaque phrase: par exemple avant tout tous les endroits dans lesquels la conjonction *ut* est finale, après les endroits où elle est consécutive etc.

En d'autres termes, les concordances de ce type sont non seulement de simples concordances, mais même un véritable lexique.

En effet, c'est après avoir examiné un par un les contextes d'une concordance que l'auteur imprimera chaque contexte enrichi du code de son emploi logique et classé selon ce critère.

5. Voyons donc maintenant la lemmatisation qu'opère notre LEL: l'ordinateur rassemble chaque mot sous son lemme, mais sans tenir compte de son emploi logique dans les différentes propositions.

6. Notre lemmatisation en effet a d'abord pris comme règle, une liste des lemmes déjà faite du *Lexicon Totius Latinitatis* de Forcellini qui comprend environ 90.000 lemmes. Puis à propos de chaque forme des mots nous nous sommes efforcés de répondre à cette question: De quel lemme ce mot est-il la forme? De celui-ci ou de celui-là? D'un seul ou de plusieurs? Nous y avons répondu en déterminant et codifiant tout ce qui est vrai du mot isolé, c'est-à-dire pris dans l'absolu, c'est-à-dire en dehors de toute proposition. Et nous avons confié tout ça aux cartes perforées, en distinguant les formes univoques qui appartiennent à un seul lemme, de celles, homographes, qui appartiennent à plusieurs lemmes.

Après avoir classé automatiquement par ordre alphabétique les mots de la *Summa c. Gentiles* et de la *Somme Théologique*, nous avons d'abord fait

ce travail sur leurs 50.000 formes environ. Puis nous avons donné à la machine ce premier LEL, imprimé sur bande magnétique, en même temps que le texte des *Commentaria in 4 LL. Sententiarum*.

A partir de là l'ordinateur a accompli trois opérations: en premier lieu il a imprimé les formes, dont nous n'avons pas trouvé de correspondance dans le LEL, ainsi que leurs concordances respectives; parmi les formes, dont il a trouvé la correspondance dans le LEL, les unes étaient codifiées par des codes comme univoques, les autres comme homographes: par conséquent en second lieu il a transporté automatiquement les codes des formes univoques, de la bande magnétique du LEL à tous les mots respectifs sur la bande magnétique du texte; et en troisième lieu il a imprimé la concordance des homographes.

Nous avons dû faire ensuite deux opérations:

— La première consista en ceci: dans la concordance à tous les contextes de chaque mot homographe nous avons ajouté à la main le code de la valeur qu'il possède dans chaque passage. Cette notation a été reportée d'abord sur cartes perforées, puis sur bande magnétique; l'ordinateur peut ainsi trier chaque forme homographe et la partager entre ses lemmes.

— La seconde opération consista à lemmatiser les formes qui ne se trouvaient pas dans le premier LEL et à les y ajouter.

Ensuite ce LEL, enrichi des formes relevées dans le "*Commentaria in 4 LL. Sent.*" a été appliqué à la "*Quaestio Disputata de Veritate*". Et de nouveau l'ordinateur effectua les trois opérations, et de nouveau nous aussi les deux suivantes; puis de la même façon notre LEL, enrichi encore de nouvelles formes, a été appliqué à la "*Quaestio De Malo*" et ainsi de suite pour les autres oeuvres.

7. Parmi les mots de la *Summa c. Gentiles*, au nombre de 325.000, nous avons relevé 20.000 formes de mots. Nous y ajoutâmes les 30.000 formes tirées des 1.600.000 mots de la *Summa Theologiae*; puis 12.000 formes tirées des 1.500.000 mots du "*Commentaria in Sentent.*", 4.600 formes tirées des 1.100.000 mots des "*Quaestiones Disp. et Quodl.*". Bref, 4.500.000 mots ont donné 66.000 formes, auxquelles il faut ajouter, à cause de l'homographie, 8.000 autres formes.

8. Puisque notre lemmatisation est seulement de caractère morphologique, elle prend en considération des données beaucoup moins nombreuses que la lemmatisation, appelons-la lexicale, qui recherche les catégories logiques que les mots assument dans le contexte.

Or, si nous avons également exclus de notre travail les deux types extrêmes de concordances, c'est à cause de son caractère massif. Les oeuvres complètes de St. Thomas renfermant environ 10.500.000 mots d'une part nous ne pouvions pas disposer du temps nécessaire pour effectuer les opérations de la lemmatisation lexicale. D'autre part la seule concordance des formes est à coup sûr utile au philologue, s'il s'agit d'oeuvres peu étendues; mais nous pensons que l'énorme quantité des contextes thomistes, par leur étendue même entraverait leur emploi raisonnable; et, par conséquent notre tâche devait être de classer, organiser et sélectionner les matériaux au service des usagers autant qu'il est possible en tenant compte des critères morphologiques.

9. Puisque l'utilisation de machines électroniques permet d'employer, dans l'analyse des textes, la méthode inductive qui semble devoir être jugée la méthode parfaite, la façon d'opérer suivante se révélait, aussi, possible et souhaitable: dans les concordances des formes, par l'examen de chaque forme dans tous ses passages, chercher avant tout à quel lemme elle appartient en chaque contexte, et conclure après si elle appartient à un ou à plusieurs lemmes et lesquels.

En d'autres termes, il s'agissait de déterminer par l'induction, le nombre et la nature des lemmes thomistes.

Mais à cause encore de l'ampleur de ce travail, nous avons été obligés d'abandonner cette méthode et d'adopter comme règle une liste déjà établie des lemmes communs à la latinité.

10. Si nous n'avons pas recouru à un lexique de latin médiéval, c'est parce que pour les oeuvres de St. Thomas aucun ne nous a semblé suffisant et adéquat. Les lexiques que nous avons entre les mains sont, ou trop brefs, ou incomplets. Le "Du Cange", par ailleurs, ne comporte pas généralement les mots courants, dont se sert St. Thomas, et la plupart de ceux qu'il comporte, ne se rencontrent pas dans St. Thomas. Si nous avons choisi le Forcellini de préférence au Thesaurus Linguae Latinae, c'est parce que dernier est encore incomplet.

Mais il n'est pas exclus, si Dieu nous le permet, qu'un jour nous puissions ajourner notre LEL en le comparant aux lemmes du TLL.

11. Le succès a démontré que nous avons eu raison d'adopter un lexique de la latinité classique comme règle pour classer les mots d'un auteur médiéval. En effet si nous considérons seulement la morphologie, il n'y a pas

une forme, dans les oeuvres authentiques de St. Thomas qui ne soit pleinement cohérente avec les normes grammaticales du latin classique. Les lemmes de St. Thomas qui ne se trouvent pas dans Forcellini sont environ 2.000: mais parmi ceux-ci un petit nombre tout au plus sont dérivés d'une racine non latine ou grecque, comme *ballivus* et *gulielmus*.

12. Loin de regretter, nous devons nous réjouir de ce que l'utilisation de l'ordinateur nous ait contraints à maintenir rigoureusement notre recherche de classification dans les limites de la morphologie. L'expérience en effet nous a confirmé que ceci offre aux philologues deux avantages effectifs: tout d'abord l'effort systématique est contraint de concentrer la méthode inductive en premier lieu et davantage sur les formes des mots, que sur les lemmes des mots; ensuite on est forcé à définir avec plus de soin et plus clairement la limite entre le système morphologique des mots et leur système logique ou lexical.

UTILISATION POSSIBLE DE NOTRE LEL POUR D'AUTRES OEUVRES LATINES

13. Après ce que je viens de dire, il est facile de comprendre la raison qui permet d'employer utilement notre LEL également pour l'analyse automatique de textes d'auteurs, autres que St. Thomas, et même classiques, à condition de tenir compte de deux considérations.

Tout d'abord il ne faut pas oublier que chaque texte fournira un nombre plus ou moins grand de formes qui ne se trouvent pas encore dans notre LEL, et qui devront ensuite y être ajoutées; ce nombre sera certainement plus grand pour les textes de Plaute ou de Pétrone que pour les textes de Sénèque ou de Cicéron. Par ailleurs le traitement de l'homographie, dont nous donnerons plus loin un résumé des points principaux, devra peut-être être modifié à partir de trois considérations: premièrement à partir de la comparaison avec les lemmes que d'autres lexiques ont peut-être recensés et non le Forcellini; deuxièmement: une homographie, que nous avons traitée comme négligeable chez St. Thomas, est peut-être possible chez un autre auteur troisièmement: l'administration de l'homographie onomastique doit être développée selon des critères différents pour St. Thomas et pour les autres auteurs.

Il faut remarquer aussi que modifier automatiquement la graphie qui ne distingue pas les lettres *u* et *v* en une graphie qui les distingue, ne requiert presque pas de travail.

14. Notre LEL sera donc un instrument toujours plus universel pour lemmatiser analytiquement les divers textes de langue latine, au fur et à mesure qu'il sera continuellement enrichi des formes recensées.

Mais on peut envisager aussi une autre façon pour réaliser un LEL qui soit vraiment universel:

- En premier lieu, en flexionnant automatiquement chacun des lemmes recensés dans le Forcellini et le TLL, l'ordinateur pourrait retrouver toutes leurs formes possibles.
- En second lieu, qu'on trouve le moyen de dresser la liste de toutes les formes qui s'écartent des règles morphologiques qu'on considère normales et des paradigmes qu'on considère communs.
- En troisième lieu, qu'on dresse aussi les listes de toutes les variations possibles purement graphiques et en même temps les listes de toutes les combinaisons de mots possibles, comme *intelligo* et *intellego*, *verumst*, *videsis*.

Afin de rendre possible au moins le premier point nous avons formulé nos lemmes thomistes de façon à ce que leur flexion intégrale soit possible automatiquement pour les formes régulières. Si nous pouvons passer à sa réalisation, à notre avis, beaucoup d'avantages en découleront, qui faciliteront notamment l'individuation des homographes. Enfin l'extension de cette formulation aux autres lemmes de la langue latine apportera une aide et posera des problèmes de grand intérêt à ceux qui voudront approfondir et élargir leur méthode systématique de connaissance de la langue latine.

COMPOSITION DU LEXIQUE LATIN ELECTRONIQUE

15. Comme je l'ai dit précédemment, nous avons recensé tous les lemmes du Forcellini, qui sont au nombre de 90.000, y compris les lemmes de l'onomastique et les lemmes qui renvoient à d'autres lemmes. Nous avons retenu comme lemmes seulement les mots qui sont imprimés en caractères gras au commencement de chaque article et non ceux qui se trouvent seulement dans le cours des articles.

Ces lemmes ont d'abord été perforés sur des cartes, puis organisés dans une liste alphabétique générale et pourvus d'un code numérique.

16. Il faut encore remarquer que l'ordinateur analyse chaque mot individuel, c'est-à-dire chaque série de symboles graphiques, isolée des autres mots voisins par des intervalles ou des signes de ponctuation: par exemple

Julius Caesar sera coupé par l'ordinateur en 2 mots distincts qui seront classés sous 2 lemmes différents. Il en sera de même pour les mots *dictum est* et d'autres semblables.

Par conséquent si le typographe a imprimé le mot *jusjurandum* tantôt en un mot, tantôt en 2 mots, l'ordinateur classera dans le premier cas le mot sous le lemme *jusjurandum*, mais dans le second sous les deux lemmes *jus* et *juro*. L'ordinateur opérera de la même façon, quand il rencontrera *suiipsius* et *sui ipsius*, et ainsi de suite.

Mais ceci ne doit pas être attribué à une insuffisance de l'ordinateur, car le relevé des syntagmes ne doit pas se faire dans les premières phases de l'analyse automatique, mais dans les phases postérieures: on ne peut en effet rassembler raisonnablement plusieurs mots en syntagmes selon une méthode systématique adéquate que si l'on a, auparavant, fait le recensement statistique des corrélations réciproques des mots individuels et si l'on a découvert pour chaque texte, les critères spécifiques, qui permettent de reconnaître entre les séries de plusieurs mots lesquels sont les véritables syntagmes, c'est-à-dire les unités lexicales élémentaires composées de plus d'un mot.

17. Après ces considérations, nous devons traiter quatre arguments: la polysémie, la polymorphie, quelques mots spéciaux, enfin les erreurs.

Le problème de la polysémie nous a été imposé immédiatement par cette question: quand les formes doivent-elles être partagées entre plusieurs lemmes, c'est-à-dire entre paradigmes différents? Cette question aboutissait à cette autre question: quels sont les critères pour délimiter les unités morphologiques et sémantiques? Nous avons accepté la distinction des lemmes que Forcellini présente, même lorsque nous ne comprenons pas pour quelle raison nous devons nous abstenir de lui attribuer des incohérences. Mais nous avons regroupé les emplois adjectivés et substantivés des flexions participiales sous le respectif lemme verbal et toutes les flexions adverbiales sous le respectif lemme nominal.

Par conséquent, tandis que nous avons conservé *creatura, ae* comme lemme distinct, nous avons inclus *peccatum, i* et *votum, i* sous *pecco, are* et *voveo, ere*; de la même façon nous avons assigné les adverbes *facile, faciliter, facilius* au lemme *facilis, e*.

18. Nous avons été obligés de consacrer davantage de temps et de considérations à l'homographie. Et tout d'abord, pour rechercher les limites entre la polysémie qui doit être traitée comme homographie, et la polysémie qui

ne doit pas l'être.

19. Nous avons ainsi défini 4 niveaux d'homographie: le premier, lorsque les mots latins, tels qu'il sont écrits, peuvent être aussi les mots d'une autre langue; le second, lorsque les significations des mots homographes sont radicalement différentes: par exemple, *facies* qui est visage, et *facies* du verbe faire; le troisième, lorsque les mots homographes viennent de la même racine, mais doivent être inclus dans des paradigmes morphologiques différents, exemple: *intellectus, us* et *intellectus, a, um*; le quatrième, lorsque des flexions différentes du même paradigme du même lemme sont graphiquement identiques: ainsi, *veni* ind. parf. et *veni* impératif; *albo* datif et *albo* ablatif.

20. Nous avons négligé le premier niveau et nous n'avons pas même introduit le quatrième niveau dans notre LEL, car l'énorme quantité de cette homographie nous en empêchait. Par conséquent, chaque forme reçoit les codes qui lui reviennent en premier dans son paradigme: par exemple le mot *amatis* est marqué comme seconde pers. plur. ind. prés. act. et non comme part. pass. dat. plur.; *bonis* est marqué comme dat. plur. 1° decl. et non 2de. Mais comme recherche collatérale, nous avons perforé toutes les désinences -régulières, grecques et contractées- de la grammaire latine pour en tirer, plus tard, une liste des homographies des désinences. Cette recherche deviendrait plus complexe, quand les désinences considérées en elles-mêmes, c'est-à-dire indépendamment de tout thème, seraient comparées avec l'index inverse de nos mots.

21. Nous avons donc élaboré seulement l'homographie des deuxième et troisième niveaux, comme le montre le spécimen. Nous avons trouvé que pour 100 formes, douze sont homographes au moins selon un de ces deux niveaux.

22. Par conséquent nous avons classé cette multitude d'homographes en trois parties. En premier lieu, nous avons mis à part sous le titre "homographa quae non sunt selecta", un petit nombre de formes, mais chacune relevant avec une très grande fréquence: ce sont: *eas, eo, licet, minus, modo, quam, quod, rei, solum, sui, vero, secundum*.

En second lieu, nous avons marqué plusieurs formes avec un code qui dit que nous avons négligé cette homographie: nous avons indiqué à quels lemmes cette forme peut être attribuée, mais nous en avons choisi un, auquel l'ordinateur assigne automatiquement toutes les occurrences de la dite forme. Nous avons jugé très peu probable que dans notre texte, par ex, *creator* signi-

fie non pas le créateur, mais l'impér. futur passif du verbe *creo, are*; ou que *ibis* ne soit pas le futur du verbe *eo, ire*, mais l'oiseau égyptien *ibis, ibidis*.

En troisième lieu nous trouvons les formes homographes que nous acceptons de sélectionner avant de conclure notre index.

La machine nous imprime la concordance de ces mots, dont nous sélectionnons à la main les emplois qui appartiennent à l'un ou l'autre lemme, comme j'ai dit plus haut (n. 6).

Les limites entre l'homographie que nous jugeons négligeable et celle que nous sélectionnons peuvent être modifiées sans peine par quiconque voudrait appliquer notre LEL à d'autres auteurs.

23. Notre désir de contenir dans des limites raisonnables l'homographie à sélectionner, nous a conduits à prendre d'autres décisions: aussi bien pour l'homographie onomastique que pour l'homographie entre quelques enclitiques et quelques désinences. En réalité nous n'avons pas développé l'homographie de tous les noms propres, mais seulement de ceux que nous avons jugés probables dans nos oeuvres. Aux autres nous avons ajouté un code qui dit: "Sachez que cette forme peut être assignée à des lemmes -plus précisément à un ou à plusieurs lemmes- que Forcellini reporte dans son Onomastique". Nous avons assigné aussi des codes particuliers aux mots comme *actio-ne, divisive, tumet*, qui peuvent être aussi *actio-ne, divisi-ve, tu-met*: si nous enlevons la dernière partie de ces mots, il reste cependant des mots latins entiers et possibles. Nous n'avons cependant pas distingué si le mot restant appartient au même lemme ou à un lemme différent de celui du mot complet, ou s'il est univoque, ou une seconde fois homographe.

24. Nous avons rencontré des polymorphies de différents types; nous avons trouvé en effet que des mots de même sens appartiennent à des lemmes différents, ou seulement selon la morphologie (différentes déclinaison ou conjugaison, ou bien différentes désinences à l'intérieur de la même déclinaison ou de la même conjugaison), ou seulement selon la graphie, par exemple *herus* et *erus*, *condicio* et *conditio*, *ius* et *jus*.

Nous avons regroupé toutes ces formes sous un même lemme; mais on a réuni plusieurs autres lemmes sous le même code numérique de ce premier lemme, que nous avons appelé par suite "unificateur". Dans à peu près 25.000 lemmes qui correspondent aux 170.000 formes trouvées dans St. Thomas, les lemmes unifiés sont plusieurs milliers. Si nous ne l'avions pas fait, le nombre de mots homographes aurait encore augmenté.

25. Nous avons aussi des listes où sont entassés des lemmes spéciaux: les numéraux; mots qui ne sont pas des mots, mais des lettres ou des symboles; les mots translittérés du grec ou de l'hébreu; les mots abrégés; les mots incomplets, qui ne se trouvent pas entiers dans l'autographe ayant été effacés par l'auteur, et les mots dont l'homographie n'a pas été sélectionnée. Tous les mots qui sont affectés par une enclitique univoque (voir n.23) sont répétés sous le lemme de l'enclitique. Seulement sept enclitiques finales (-ce, -cum, met, -ne, -que, -tenus, -ve) ont été considérées.

26. Pour ceux qui effectuent un travail de ce type, l'erreur surtout de l'homme est une véritable terreur. En effet, codifier 170.000 formes revient au même que d'apposer 40 symboles ou lettres à chaque forme. Cela fait, en tout, sept millions. Dans une telle quantité, il est impossible d'éviter totalement que les impitoyables lois du hasard et de la chance -hélas! quelle malchance! - n'exercent leur tyrannie. L'expérience nous enseigne que les opérations de correction et de vérification exigent trois fois plus de temps que la première production. En d'autres termes, le quart du temps doit être attribué à la première élaboration, les autres trois quarts à la recherche et à la correction des erreurs.

Permettez-moi de faire une prophétie facile: lorsque quelqu'un ouvrira pour la première fois le volume de notre LEL, "statim et illico" une erreur dans la première page blessera ses yeux, à coup sûr; et cette erreur, nous échappa toujours saine et sauve, à nous qui avons relu cette page maintes fois et avec grande attention! Je prie ce lecteur de ne pas vitupérer contre nous, mais de faire deux considérations: avant tout le nombre d'erreurs qui nous ont échappé, demeure-t-il dans la limite d'une quantité négligeable? Par ailleurs, quel résultat aurait-il pu atteindre, s'il avait été à notre place?

Enfin qui nous signalera les erreurs qu'il aura découvertes, que celui-ci sache qu'il aura accompli une bonne action.

SPECIMINA

27. Specimina quae sequuntur, excerpta ex 3.800 paginis, praebent uoces quae in nostro LEL initio litterarum A et E inueniuntur: et primo quidem ordine alphabetico digestas, secundo autem easdem sub suo quaque lemme dispositas; quibus harum conspectus lemmatum succedit. Symbolorum autem quae in harum paginarum regionibus siue columnis reperiri possunt significationem subiungemus.

28. = = col. 1 = numerus qui est formae tessera.
 = = col. 2 = forma.
 = = col. 3 = formae typus:
 - 0 = huius formae tesseras automaticè machina transferet in singula textus huic respondentia uerba.
 - 1,2 ecc. = uerborum huic respondentium, utpote homographorum, machina concordatiam imprimat.
 - uacuum = de hac, utpote quae est homographa neglecta, machina non curat.
29. = = col. 4 = tesserae ad concordandum:
 - P = uerbum "plenum".
 - V = uerbum "uacuum".
 - W = uerbum "nullius momenti".
 - 1 = huius dabitur index syntagmatum.
 - 2 = huius dabitur concordantia proprie dicta.
 - uacuum = uerbum "homographum neglectum".
 = = col. 5 = numerus qui est lemmatis tessera.
30. = = col. 6 = tesserae morphologicae:
 primo loco:
 - 1 = flexio nominalis positiui gradus.
 - 2 = flexio nominalis comparatiui gradus.
 - 3 = flexio nominalis superlatiui gradus.
 - uacuum = indeclinabile vel flexio uerbalis.
- secundo loco: hinc deinceps quae de his tesserae dicuntur ualent de omnibus lemmatibus, exceptis quibus tesserae 9L000, 99993, 99995T, 99997B.
- O = aduerb. indecl.
 - G = nomen indecl.
 - S = (etiam) praepositio.
 - A,B,C,D,E = 1, 2, 3, 4, 5 decl.
 - F = decl. irreg.
 - J,K,L,M = 1, 2, 3, 4 con.
 - N = con. irreg.
- tertio loco:
 - A, J = indic. act., pass.

- B, K	= conjunct.	act., pass.
- C, L	= imperat.	act., pass.
- D, M	= partic.	act., pass.
- E, N	= gerundium uerbi	act., pass.
- O	= gerundium.	
- P	= supinum.	
- H, Q	= infin.	act., pass.

quarto loco:

- 1	= praesens.
- 2	= imperfectum.
- 3	= futurum.
- 4	= praeteritum.
- 5	= plusquampraet.
- 6	= fut. ant.

quinto loco:

- A, J	= nom.	sing., plur.
- B, K	= gen.	sing., plur.
- C, L	= dat.	sing., plur.
- D, M	= acc.	sing., plur.
- E, N	= voc.	sing., plur.
- F, O	= abl.	sing., plur.
- G	= aduerb.	
- H	= et acc. et abl.(saltem): u.g. <i>hoc, se.</i>	

sexto loco:

- 1	= masc. gen.	
- 2	= fem. gen.	
- 3	= neutri gen.	
- 4, 7	= prima pers.	sing., plur.
- 5, 8	= secunda pers.	sing., plur.
- 6, 9	= tertia pers.	sing., plur.

septimo loco: tessera formae affectae enclitica aut var. graph.

31.	= col. 7	= lemma
	= col. 8	= typus homographiae:
	- 1	= uerbum uniuocum.

- K = uerbum homographum de quo machina ut de uniuoco operatur.
 - 2 = uerbum homographum siue seligendum siue neglectum.
 - 3 = uerbum (forte) homographum, non selectum.
 - W = 1 cum encl. hom.
 - X = 2 cum encl. hom.
 - S = K cum encl. hom.
32. = = col. 9 = quot huiusmodi formae homographae.
 = = col. 10 = tessera onomastica:
 - 1 = haec forma potest pertinere etiam ad unum uel plura lemmata Onomastici Forc.
 - 2 = nomen proprium.

33. In tertio autem specimine symbola inueniri possunt:

- = = col. 2 = tessera unificationis:
 - 1 = lemma cui aliud aut alia unificantur.
 - 2, 3 . . . = lemma alii unificatum.
- = = col. 3 = tessera lemmatis:

primo loco: tesserae eadem sunt ac quae secundo loco col. 6 n. 31 declarantur.

secundo loco:

- G = nomen indecl. communis aut incerti generis.
 - M = nomen quod flectitur tantum sec. g. masc.
 - F = nomen quod flectitur tantum sec. g. fem.
 - N = nomen quod flectitur tantum sec. g. neut.
 - X = nomen habens plurales formas gen. diuersi a sing. aut inter se.
 - 1 = nomen unius nom., quod al. casum habet genere diuersum.
 - 2 = nomen duorum nom.
 - 3 = nomen trium nom.
 - A = actiuum.
 - D = deponens.
 - U = semideponens.
- = = col. 4, 6, 8 tesserae quibus machina automaticè uerbum in om-

nes suas formas flectere instruitur.

= = col. 5, 7, 9 = lemmatis thematum uoces:

in 7 et 9 quaedam signa machinae indicant quanta pars uocis ex col. 5 sit repetenda, ommissa quia mensuram secus totum excederet:

// = repetitur integra uox.

/-/ = repetitur tota uocis pars inuariata, scil. usque ad lineolam.

/07/ = repetuntur uocis totidem, u.g. septem, primae litterae.