

Omtrent frequentieonderzoek op antieke teksten, inzonderheid Griekse*

door

ERIK GOERLANDT

Men kan zich niet erg vergissen als men poneert dat de humane wetenschappen in het algemeen zich bezighouden met de studie van allerhande variabelen. Die stelling geldt ook voor alle wetenschappen van de Oudheid en voor de klassieke filologie in het bijzonder. Met variabelen heeft immers niet alleen de historicus, de epigrafist, de papyroloog te maken, maar ook alwie een literaire tekst wil uitgeven of de vorm of de inhoud ervan wil onderzoeken. Daarom is het niet verwonderlijk dat ook de verschillende disciplines van de Oudheid getracht hebben hun variabelen te tellen en te ordenen met behulp van de computer van zodra die vlugge rekenaar en rangschikker daartoe kon worden ingeschakeld. Sta me toe u eerst een overzicht te geven van de disciplines van de Oudheid waar de computer sedert een vijftiental jaren een plaats verworven heeft in het behandelen van hun meetbare variabelen. Dit overzicht wil niet chronologisch zijn of exhaustief, maar systematisch en exemplarisch.

Van onze materiële overblijfselen uit de Oudheid worden de zuiver archeologische vondsten met hun specifieke variabelen hier best buiten beschouwing gelaten, al is ook in de archeologie de computer lang geen onbekende¹. De tot ons gekomen geschreven bronnen, hetzij documentair, hetzij literair van aard, kunnen echter evenals de ongeschreven bronnen van archeologisch standpunt worden benaderd. Epigrafica en papyrologica zijn evengoed mate-

* Uitgewerkte versie van een voordracht gehouden te Brussel op de Iden van maart 1975. Om de kosten te drukken worden de Griekse woorden in de tekst getranslittereerd, waarvoor ons excuus.

1. Zo werd een beschrijvende code voor computeronderzoek van alle mogelijke variabelen in de Kretenzische munten van de VIde-IVde eeuw v.C. verwezenlijkt door de Fransman J. C. Gardin en zijn ploeg van het C.A.D.A. (*Centre d'Analyse Documentaire pour l'Archéologie*). Ook de schijnbaar eenvoudige, maar in feite complexe variabelen in de vazenkunst worden door hen bestudeerd. Zie Ouy, pp. 23, 28, 29. Op de tentoonstelling *Onder de zon van Amarna. Echnaton en Nefertiti*, Koninklijke Musea voor Kunst en Geschiedenis, Brussel 17 jan.-16 maart 1975, vernam ik ook dat men uit de *talatats* van de negende pyloon van de tempel te Karnak het oorspronkelijke reliëf tracht te reconstrueren met behulp van de computer (het *Akhenaton Temple Project*).

riële produkten van menselijke bedrijvigheid als sieraden of gebruiksvoorwerpen, en uiterlijk variëren ze evenzeer naar gelang van tijd en plaats.

Fysische of archeologische kenmerken van een inscriptie zijn bijvoorbeeld haar positie op het monument (links, midden, rechts, voorzijde, zijkanten of achterzijde), de versieringen (met palmblaren, klimopblaren, kransen), de grafie van de letters (verschillende vormen naar gelang van plaats, sociale klasse en tijd, schrijfrichting, leesbaarheid, tussenkomst van een andere hand, afkortingen, ligaturen), de aanwezigheid of afwezigheid van verselementen, enz.². Die variabelen kan de epigrafist van een passende code voorzien en na codering van een aantal inscripties kan de ordinator naar wens lijsten aanleggen met verwijzingen naar alle plaatsen waar een of andere variabele in het behandelde corpus voorkomt. Naast de fysische variabelen van de inscripties kunnen natuurlijk ook die van de papyrologische vondsten machinaal worden geordend³.

Een epigrafisch of papyrologisch document is echter meer dan een louter archeologische vondst: het document is ook drager van een mededeling. Die kan worden bestudeerd naar de inhoud en naar de vorm. Bij wijze van spreken ziet men ook hier niets dan variabelen. In epigrafica is er bijvoorbeeld sprake van verschillende goden, van verschillende personen met elk hun eigen praenomen, cognomen en gentilicium, van verschillende magistraturen en functies, tribus, keizers en geografica⁴; in papyrologische teksten gaat het bijvoorbeeld om verschillende soorten zaken, om verschillende belanghebbenden met elk hun naam en hoedanigheid, om verschillende bedragen, verschillende plaatsen en tijdstippen⁵. Ook die variabelen kunnen machinaal gerepertorieerd worden. Met de ordinator kunnen zelfs relaties tussen verschillende variabelen worden bestudeerd, bijvoorbeeld tussen personen, tussen personen en functies, tussen functie en plaats van uitoefening⁶. Behalve de inhoud van onze documentaire teksten wordt ten slotte ook hun taal machinaal onderzocht⁷, met bijzondere aandacht voor het herstellen van verminkte woorden, het aanvullen van lacunes en het ontdekken van stereotiepe formules⁸.

2. Cfr. Jory & Moore, pp. 8-9; Janon, pp. 5, 9-10 en 18-19.

3. Cfr. Evrard, *La papyrologie*, p. 93.

4. Cfr. Janon, p. 5.

5. Cfr. Evrard, *La papyrologie*, pp. 87-88.

6. Cfr. Janon, p. 10.

7. Zie bijvoorbeeld Tomsin, met enkele resultaten, pp. 65-66.

8. Cfr. Evrard, *La papyrologie*, pp. 93 en 95-96; Zarri, p. 68. Voor het gebruik van de computer bij de studie van het Romeinse recht, zie Müller. Ook in de mycenologie en bij het onderzoek van de maya-glieden doet men een beroep op de

Tot hiertoe is gesproken over het inschakelen van de computer bij de studie van allerhande variabelen in onze documentaire teksten. Hoe belangwekkend ook, zijn epigrafie en papyrologie voor de filoloog slechts hulpwetenschappen; in hoofdzaak zal hij teksten bestuderen die bewaard gebleven zijn in manuscripten. De mededeling met haar vorm en haar inhoud is echter niet het enige waarmee hij rekening moet houden. Als drager van een mededeling is een manuscript evengoed als een inscriptie of een papyrologisch document een archeologische getuige van menselijke bedrijvigheid. De drie genoemde soorten geschreven bronnen zijn in hun uiterlijke vormgeving niet noodzakelijk het werk van één man, maar bij de afwerking van een manuscript komt doorgaans het grootste aantal ambachtshandelingen kijken: de perkament- of papiermaker, de linieerder, de kopiïst, de corrector, de verluchter of schilder, de binder en soms de edelsmid; van al die vakmannen is de kopiïst voor de filoloog wel de interessantste. Daar al die mensen mogelijk verschillend werken naar gelang van plaats en tijd ontstaat een indrukwekkend aantal variabelen die voor computerbehandeling gecodeerd kunnen worden. Aldus kunnen de handschriften dank zij allerlei externe criteria voor de filoloog gedateerd worden. Om hem alle manuscripten te laten vinden die voor zijn studie van belang zijn, kan de codicoloog ook van alle collecties computercatalogi bezorgen met alle namen van auteurs en werken, met incipits, explicits, colofons, enz.⁹

Ter attentie van de filoloog, die zich toch vooral met in manuscripten bewaarde literaire teksten inlaat, kan hier wat uitvoeriger ingegaan worden op enkele studies van antieke auteurs, ondernomen met behulp van de computer.

Het gebruik van de machine is in het filologisch onderzoek binnengedrongen vanaf het eerste stadium ervan, de recensio. In tegenstelling met epigrafische of papyrologische documentaire teksten bestaan er van antieke werken in handschriften doorgaans veel kopieën. Naast louter orthografische varianten, die geïnspireerd kunnen zijn door verschillende orthografische conventies en verschillen in uitspraak naar gelang van de nationaliteit van de kopiïst of van de periode waarin hij leefde, constateert de tekstuitgever tussen verschillende kopieën van éénzelfde werk ook andere afwijkingen: de fouten en de *variae lectiones*. Daar van verscheidene variëteit lezingen de ene door de andere verklaard kunnen worden, zal de tekstuitgever om een kritische editie te kunnen bezorgen de manuscripten rangschikken op basis van de

machine. Voor de mycenologie zie Mater en Geiss, voor de maya-teksten zie Mater, p. 14, en Morgan.

9. Cfr. Ouy, pp. 28-35.

gemeenschappelijke fouten. Dat is de methode Lachmann. Dom Henri Quentin ontwierp echter een methode waarbij het begrip fout wordt vervangen door het begrip variante (de theorie van de vergelijking per drie). Zijn methode is grondig gekritiseerd. Toch heeft dom Jacques Froger van Solesmes op een inzicht van Quentin gesteund om zijn theorie over de automatische seriëring van manuscripten op te bouwen¹⁰. In navolging van Quentin meent Froger in een filiatie van manuscripten een onderscheid te moeten maken tussen de onderlinge banden tussen de manuscripten (*l'enchaînement des copies*) en de oriëntatie van hun stemma. In zijn theorie verloopt de stemmatisering als volgt. Na collationering van de kopieën kiest men er willekeurig één als origineel en bepaalt men welke groepen woorden er in de andere manuscripten variant zijn ten opzichte van dat „origineel”. Dan rangschikt men de variante groepen in dalende lijn van het aantal handschriften welke die variante groepen vertonen. Volgens hem staat die werkwijze toe de onderlinge verwantschap tussen de verschillende kopieën te ontdekken. Om nu de oriëntatie van het stemma te kennen, wil hij opnieuw het begrip fout invoeren. Dan komt het er natuurlijk op aan in zijn variante groepen de varianten te onderscheiden die als fouten mogen worden beschouwd. Juist daarin zit hem dikwijls de moeilijkheid¹¹. Een computerprogramma om manuscripten automatisch te classificeren en de stemmata te oriënteren heeft Froger echter niet uitgewerkt¹². Daartoe hebben te Göttingen Dietmar Najock, Jürgen Mau en Heinz-Jörg Ahnert het AFFILI-programma opgesteld¹³. De laatst genoemde heeft als testmateriaal de collaties van Dionusios de Areopagiet door Andresen en Ritter mogen gebruiken¹⁴.

Van zodra de tekst van een auteur vastgesteld en uitgegeven is, leent hij zich tot taalkundig, inhoudelijk en stilistisch onderzoek. Bij elk van die benaderingswijzen kan men een beroep doen op de computer. Die machine „met bronzen ingewanden” bezit een sterker geheugen en een grotere betrouwbaarheid dan de grammaticus Didumos¹⁵. Met grote snelheid legt de ordinator ook hier de gewenste lijsten aan. De computer kan men zelfs automatisch Latijnse en Griekse teksten laten coderen, morfologisch, syntactisch

10. Als testmateriaal gebruikte Froger een tekst van Molière, waarvan hijzelf acht „kopieën” maakte met vooraf bepaald stemma, waarbij elke „kopiist” de fouten van zijn model bewaarde en er enkele nieuwe invoerde, cfr. Evrard, *La critique des textes*, p. 74.

11. Id., pp. 72-75.

12. Daarover heb ik althans geen gegevens.

13. Cfr. Najock, *Orientation en Automatic classification*, Mau en Ahnert.

14. Cfr. Ahnert, p. 54, n. 5.

15. De associatie is van de Amerikaan Nathan A. Greenberg, cfr. Greenberg, p. 1.

en metrisch. Een programma om de woorden uit een Latijnse tekst automatisch te lemmatiseren en te analyseren werd in 1965 op punt gesteld in het Luikse *Laboratoire d'Analyse statistique des Langues anciennes* (LASLA). Het programma was ontworpen door Albert Bodson en Suzanne Govaerts¹⁶ en was er in gebruik tot in 1970, toen het werd verbeterd door Joseph Denooz¹⁷. Een gelijkaardig programma voor het Grieks werd in 1970 uitgewerkt door Etienne Evrard¹⁸. Of het nieuwe systeem van Denooz intussen ook op het Grieks wordt toegepast, heb ik nergens gelezen. De morfologische analyse gebeurt door vergelijking van elk woord uit een tekst (*woordvorm, mot-forme, item of token* genoemd) met een lexicon van lemmata (*lexemen, lemmes, lemmata, vocabula of types*) en een lexicon van uitgangen. Voor het onderzoek van de syntaxis en de stilistiek wordt elk item gereferentieerd naar zijn plaats in lijn of vers, in zin, paragraaf en hoofdstuk, en wordt er ook onderscheiden of een werkwoord in een hoofdzin dan wel in een bijzin voorkomt. Het oudste programma gaf alle mogelijke morfologische determinaties op, het jongste sluit amfibologieën uit. Voor de moeilijkste Latijnse teksten is het doeltreffend in 91 % van de gevallen¹⁹.

Bij het metrisch onderzoek werd de machine oorspronkelijk ingezet om volledig manueel gescandeerde verzen en woorden te triëren. Aldus werd de ordinator in 1964 door de Amerikaan T. McDonough gebruikt voor zijn studie van de metriek bij Homeros en in 1967 door Robert R. Dyer Jr., eveneens Amerikaan en ook voor een studie van Homeros. Wilhelm Ott van Tübingen en de Amerikaan Nathan A. Greenberg hebben echter vanaf 1967 programma's gepubliceerd voor een zuiver automatische scansie van de Latijnse hexameter. Een tegenhanger voor het Grieks werd in 1972 uitgewerkt door Evrard. Het gescandeerde pensum bestond uit vijf *Idyllen* van Theokritos. Zijn programma laat de ordinator elk vers vergelijken met 32 mogelijke versschema's. In 98 % van de gevallen levert de machine het juiste resultaat; zij levert het zonder menselijke hulp in 92 % van de geanalyseerde verzen²⁰.

Naast volledige woordindices zoals het LASLA er biedt²¹, kan de machine allerlei andere integrale of selectieve, enkelvoudige of meervoudige lijsten produceren van delen van woorden (lettergrepen of klanken), afzonderlijke woorden of woordseries. De

16. Cfr. Bodson & Govaerts.

17. Cfr. Denooz, p. 5.

18. Cfr. Evrard, *L'automatisation*.

19. Cfr. Denooz, pp. 55-56.

20. Cfr. Evrard, *Scansion automatique*, pp. 2, 3, 7 en 18.

21. Voor enkele realisaties van het LASLA zie Delatte, *Computers*, pp. 101-103 en Govaerts, *L'étude*, p. 101.

studie van delen van woorden kan interessant zijn voor de linguïstiek, de metriek, de stilistiek. In een artikel over computeronderzoek op Homeros spreekt Dyer onder andere over de mogelijkheid om met de computer de bouw van het woord te bestuderen²² en Annette Paquot van de U.C.L. heeft het over een Plautus-project waarin onder meer problemen van de fonetica aan bod komen: de graad van redundantie van een taal op het niveau van de klanken, het functionele rendement en de diachronie van consonantgroepen of *clusters*, aspecten van de syntactische fonetica (bijvoorbeeld de probabiliteit van epanastrofe, de stijlfiguur waarbij de eindklank van een woord terugkeert aan het begin van het volgende), de graad van muzikaliteit van de taal bij een of andere taalgebruiker, het voorkomen van alliteratie, homoioteleuta en rijm²³. De machine kan ook een vocabularium geheel rangschikken in dalende lijn van frequentie, of gedeeltelijk (bijvoorbeeld de 100 belangrijkste betekenisdragende woorden); van een vocabularium kan ze ook elk woord in een meervoudige lijst opdelen volgens zijn frequentie per rol (als het een toneelstuk betreft), per gedicht (als het om een bundel gaat, bijvoorbeeld per elegie, ode of ecloge), enz. Dank zij de ordinator kan de filoloog ook complete lijsten van de metrische variabelen verkrijgen, bijvoorbeeld van alle dactylen en spondeeën, van alle cesuren en hiaten en hun plaats in het vers, van de woorden met verschijnselen als de bovengenoemde alliteratie, homoioteleuton en rijm, enz.²⁴. Lijsten van woordseries kunnen dienen voor het opsporen van formules, bijvoorbeeld epische formules²⁵, en voor de studie van de zinsbouw²⁶.

Aan de hand van de zopas besproken hulpmiddelen (woordindices en allerhande lijsten) kan de filoloog terdege de mededeling van een tekst aanpakken. Daar de woorden de inhoud en de vorm van de mededeling dragen, is het normaal dat de keuze en de opeenvolging van de woorden afhangen van de inhoud, van conventies naar gelang van het genre, alsook van het temperament en de leeftijd van de auteur. Dat is een empirische waarheid, maar een en ander is ook statistisch nagegaan²⁷. Eerst wordt hier de woordkeuze behandeld in haar relatie met het onderwerp, met het genre en met de persoonlijkheid van de auteur; vervolgens wordt er iets gezegd over de opeenvolging van de woorden.

22. Cfr. Dyer, pp. 4-14.

23. Paquot, pp. 44-47.

24. Cfr. Ott, pp. 17-18; Evrard, *Scansion automatique*, p. 4. Sindsdien zijn er van Ott reeds verscheidene metrische analyses van Latijnse dichters verschenen.

25. Cfr. Dyer, pp. 17-32.

26. Cfr. Delatte, *L'ordinateur*, vooral pp. 8-13.

27. Cfr. Coppens-Ide, p. 188.

Dat het onderwerp van een geschrift invloed uitoefent op de keuze van de woorden is geen nieuwe ontdekking. Regelrecht frequentieonderzoek van alle woorden staat echter toe dat verschijnsel beter te evalueren. Daartoe kan een traditioneel kaartenstelsel worden aangewend ofwel de machine. Enkele voorbeelden. Na het computeronderzoek van de werken van Seneca philosophus in het LASLA werd het Louis Delatte nog duidelijker welke belangrijke vocabulariumverschillen er bestaan tussen bijvoorbeeld de *Quaestiones naturales* en de *Brieven aan Lucilius*²⁸. Ook onderzocht Evrard dat een bepaald hoofdstuk uit de *Consolatio ad Marciam* van dezelfde auteur zich van de rest van het werk onderscheidt door een buitengewone concentratie van woorden met betrekking op de leeftijden²⁹, en na het LASLA-onderzoek van de *Oden* van Horatius constateerde hij dat *mare* en aanverwante woorden uitzonderlijk frequent zijn in het Eerste Boek van dat werk³⁰. E. D. Kollmann van de universiteit van Tel Aviv merkte dank zij het LASLA-onderzoek van Vergilius' *Eclogen* belangrijke verbale verschillen op tussen die verschillende gedichten. Ze verschillen soms onderling in belangrijke mate naar het aantal gebruikte pronomina, vocatieven, imperatieve, tussenwerpsels, abstracta, Griekse woorden en Romeinse en volgens het Grieks geïnflecteerde eigennamen, verschillen die volgens hem aan de inhoud te wijten zijn³¹. Er zijn dus niet alleen belangrijke divergenties in het voorkomen van de betekenisdragende woorden, maar ook van sommige grammaticale woorden en grammaticale categorieën. Voor het Grieks werd niet veel vergelijkbaar materiaal gevonden. Toch dit. Bij de manuele voorbereiding van een Grieks frequentievocabularium viel het Evrard op dat het woord *éphēbos* in zijn uitgelezen teksten wel vrij frequent was, maar het kwam alleen voor bij Xenofoon en dan nog in één enkele passus van de *Kuroēpaideia*³². Bij dezelfde gelegenheid ondervond hij ook dat *phēmi* veel frequenter is bij Xenofoon en Platoon dan bij de andere onderzochte auteurs. Het gebruik van het werkwoord *hēgēomai* *leiden* was haast uitsluitend vertegenwoordigd bij Xenofoon, terwijl Lusias, Platoon en Demosthenes *hēgēomai* haast uitsluitend in de zin van *menen, beschouwen als* gebruiken. De woorden *psēphos* en *psēphizō* bleken vooral lysiaans te zijn, *aretē* platonisch³³. Uit een

28. Cfr. Delatte, *Key-words*, p. 67, n. 3.

29. Cfr. Evrard, *Deux programmes*, p. 86.

30. Cfr. id., p. 83.

31. Cfr. Kollmann, *A study*, pp. 15-16 (Table 10).

32. Cfr. Evrard, *Deux programmes*, p. 83.

33. Cfr. Evrard, *Le Laboratoire*, pp. 111-113. Het corpus teksten dat aan de basis lag van zijn geplande frequentievocabularium (bij mijn weten is het echter niet

artikel van Leonard Brandwood uit Manchester over computeronderzoek bij Platoon ten slotte heb ik genoteerd dat de verspreiding van *ean* in de verschillende Boeken van de *Wetten* zeer heterogeen is naar gelang van het gebruik van de formule „Als die en die overtreding begaan wordt, is die en die straf van toepassing”³⁴.

Behalve met de inhoud staat het vocabularium ook met het genre in verband. Steunend op de manueel vervaardigde frequentielijst van de Amerikaan Paul B. Diederich heeft Kollmann nageteld dat er Latijnse woorden zijn die alleen in het proza gebruikt worden, terwijl andere uitsluitend in de poëzie voorkomen³⁵. Verder karakteriseert hij woorden aan de hand van hun frequentie als vooral prozaïsch, vooral poëtisch of vooral middellatijns³⁶. Govaerts en Denooz hebben in hun machinaal vervaardigde Latijnse basiswoordenschat het gebruik van het vocabularium bij verscheidene auteurs ook grafisch voorgesteld. Daardoor wordt het verschil tussen proza- en poëzietaal nog duidelijker. De helling van de curve toont aan dat de dichter een rijkere, meer gediversifieerde taal gebruikt met minder *mots-thèmes* (gewone woorden zonder de welke men geen gedachte zou kunnen uitdrukken) en meer *mots de caractérisation*³⁷. Nochtans zijn er ook graden van poëtische en prozaïciteit. Zo merkt Kollmann op dat de woordkeuze in Vergilius' *Eclogen* en in Horatius' *Oden* meer aansluit bij de doorsnee-literatuurtaal, Tibullus' en Propertius' woordkeuze minder³⁸. Aan de andere kant hebben Govaerts en Denooz onderzocht dat Tacitus' woordgebruik iets heeft van het poëtische genre³⁹. Naast verschillen tussen proza en poëzie zijn er dus ook divergenties binnen elk van beide. Op het niveau van de lemmata zijn er wat de frequentie van de betekenisdragende woorden betreft merkwaardige verschillen tussen de *Eclogen* van Vergilius, de *Oden* van Horatius en de *Elegieën* van Tibullus en Propertius, verschillen die Kollmann aan het verschil in literair genre wijt⁴⁰. Tussen de *Oden* en de *Eclogen* vindt hij ook belangrijke afwijkingen op het niveau van de woordcategorieën, de naamvallen, de

gepubliceerd) was geëxcerpeerd uit „*la prose attique et atticisante*”, zonder verdere details, cfr. id., p. 107.

34. Cfr. Brandwood, p. 13.

35. Cfr. Kollmann, *Word frequencies*, pp. 6-8 (Table 3 en 4).

36. Id., pp. 9-11 (Table 5).

37. Cfr. Govaerts & Denooz, *Présentation*, p. 24. De grafieken vindt men in bijlage bij dat artikel en ook, gedeeltelijk, in de eerste versie van hun Latijnse basiswoordenschat, cfr. id., *Lexique*.

38. Cfr. Kollmann, *Word frequencies*, p. 12, sub 1.

39. Cfr. Govaerts & Denooz, *Présentation*, p. 24.

40. Cfr. Kollmann, *A study*, pp. 8-9 (Table 5).

wijzen en de tijden, en de pronomina⁴¹. Over het verband tussen vocabularium en genre in het Grieks dat zou vastgesteld zijn met behulp van de computer heb ik minder informatie. Bij zijn reeds boven gemeld manueel frequentieonderzoek heeft Evrard wel een kenmerk van oratorische stijl herkend in het gebruik van *autos* bij Demosthenes. In het corpus van zijn excerpten waren de occurrenties van dat pronomen in dalende lijn van frequentie als volgt verdeeld over de drie gebruikswijzen: 1. in een casus obliquus (*eius*, enz.), 2. met de betekenis *ipse*, en 3. met het lidwoord (*idem*). Bij Demosthenes, de redenaar, bekleedt de betekenis *ipse* niet langer de tweede, maar de eerste plaats⁴². Met de computer heeft de Canadees Gilles Maloney de frequentie en de plaats van de werkwoordelijke vormen bij Thoekudides bestudeerd. De werkwoorden zijn bij die auteur significant frequenter in de redevoeringen dan in het verhaal. Er is dus duidelijk een verschil qua genre. Opmerkelijk is ook het overvloediger gebruik van de infinitief in Thoekudides' redes en van de losse genitief in de verhalende gedeelten van diens werk⁴³. Zoals het geval was voor het Latijn, vertonen ook in het Grieks echter niet alleen de betekenisdragende, maar ook de grammaticale woorden verschillen in frequentie naar gelang van het genre. Zo constateerde Brandwood dat in de dialogen van Plato's *Apologie* per zin minder dan half zoveel *kai*'s voorkomen als in het doorlopend proza, en in de dialooggedeelten van de *Timaios* en van de *Kritias* zijn er eveneens significant minder *dé*'s dan in de gedeelten doorlopend proza⁴⁴.

De persoonlijkheid van een auteur, zijn belangstellingssfeer, zijn temperament is een derde factor die zijn woordkeuze en woordgebruik bepaalt. Uit een lijst van alle woorden van een werk, gerangschikt in dalende lijn van frequentie, kan men de zogenaamde *sleutelwoorden* halen. De zeer talrijk voorkomende *mots-outils* (de grammaticale woorden) en de *mots-thèmes* en *mots de base* worden buiten beschouwing gelaten. In aanmerking komen de frequente *mots de caractérisation* (woorden met een zeer precieze en beperkte betekenis), die het talrijkst zijn in aantal lemmata, maar het minst talrijk in occurrentie. De min of meer frequente woorden uit hetzelfde semantisch veld kunnen dan worden gegroepeerd en achteraf kan men aan de hand van de kontekst van elk van die woorden nagaan of er in het bestudeerde werk soms niet bewust of onbewust een of meer thema's verscholen zijn. Men kan

41. Cfr. id., p. 4 (Table 1) over de woordcategorieën, pp. 5-6 (Table 2) over de naamvallen, pp. 6-7 (Table 3) over de wijzen en de tijden, en p. 7 (Table 4) over de pronomina.

42. Cfr. Evrard, *Le Laboratoire*, p. 113.

43. Cfr. Maloney, p. 91.

44. Cfr. Brandwood, pp. 6 en 8.

integendeel bijvoorbeeld ook met de tekst in de hand bij de frequentste twintig karakteristieke woorden aanstippen tot welk veld ze behoren om daarna te onderzoeken hoe de voorgedeede woorden van eenzelfde semantisch veld over het werk verspreid zijn. Beide werkwijzen zijn voorbeelden van *content-analysis*. Door de eerste methode wordt het temperament van de behandelde auteur wel het best belicht. Een voorbeeld daarvan is Delattes opstel over sleutelwoorden bij Propertius en Tibullus⁴⁵. Een voorbeeld van de tweede soort content-analysis is een artikel van Kollmann over de bouw van de *Ecloges*⁴⁶. Voor het Grieks heb ik alleen een studie gezien van de manier waarop een auteur, met name Iohannes Ludos, als bureaucraat zijn omgeving zag⁴⁷. Bedoeld artikel geeft hoofdzakelijk een voorstelling van bruikbare manueel in te vullen rasters voor de studie van verhoudingen. Die verhoudingen kunnen dan echter wél met de computer in grafieken en diagrammen uitgebeeld worden⁴⁸.

De karakteristieke woorden in het werk van een auteur bestuderen is één middel om iets meer te weten te komen over diens persoonlijkheid, een ander middel houdt een studie in van *alle* woorden die in een werk voorkomen. Men kan namelijk ook nagaan hoe dikwijls een auteur elk van zijn vocabula bezigt, m.a.w. hoe rijk of hoe arm zijn taalgebruik is. De verhouding tussen het aantal woordvormen en het aantal vocabula wordt *moyenne d'emploi* of *gemiddeld gebruik* genoemd. Dat gemiddeld gebruik verschilt soms erg van auteur tot auteur⁴⁹. Interessanter nog is het te

45. Cfr. Delatte, *Key-words*. Uit de key-words en hun context leidt hij af dat Tibullus teergevoelig is, aandacht heeft voor de details van de menselijke fysiek en onstuimig is, wat op zijn zinnelijkheid en zijn jeugd wijst. Uit Tibullus' gebruik van wijzen en tijden wordt geconcludeerd dat hij aangegrepen is door het statisch karakter van de dingen en dat hij in het heden leeft (pp. 36-44). Alhoewel men kan zeggen dat de dichter blijkbaar veel waarde gehecht heeft aan magie, was hij toch te jong en te rechtlijnig om veel thema's te verwerken (pp. 64-66). Propertius is Tibullus' tegenbeeld: hij bezit geen teder gevoel noch aandacht voor fysieke details, is bezadigd, beschrijft altijd beweging, leeft van het verleden (pp. 37-44) en speelt bewust op morbiede wijze met het thema van de dood, zelfs van de zelfmoord (pp. 45-51). Bovendien is hij volgens de woorden van Delatte geobsedeerd door het water, een relict uit zijn jeugd (pp. 52-68). Voor Propertius' water-thema zie ook reeds Delatte, *L'ordinateur*, pp. 18-19.

46. Cfr. Kollmann, *A study*.

47. Cfr. Carney.

48. Een voorbeeld van structuuranalyse met behulp van de computer vindt men in een opstel van twee Amerikanen over de bouw van *Les liaisons dangereuses*. Cfr. Allais (Christiane en Claude).

49. Voor het Latijn zijn er gegevens voorhanden uit deze onderzochte pensa: 1. een selectie uit Phaedrus, 2. excerpta uit de *Metamorfozen* van Ovidius, 3-5. Vergilius' *Bucolica*, *Georgica* en *Aeneis* I-VI, 6. Catullus, 7. de *Oden* van Horatius, 8. Caesar *de bello Gallico*, 9-10. Sallustius' *Catilina* en *Jugurtha*, 11. exc. philosophica uit Cicero, 12. Seneca's *Brieven aan Lucilius*, 13. exc. uit Tacitus' *Historiae* en *Annales*, 14. exc. uit Livius. Cfr. Govaerts & Denooz, *Présentation*, pp. [31]-[57].

weten of de verhouding tussen het aantal woordvormen of *items* en het aantal lemmata bij een auteur constant is dan wel varieert van woordzone tot woordzone (*mots-outils*, *mots-thèmes*, *mots de caractérisation*). Om werken van eenzelfde of van verschillende auteurs op dat punt te vergelijken, kiest men er het best van ongeveer gelijke lengte, d.w.z. van een ongeveer gelijk aantal items. Zo neemt Govaerts ter vergelijking Seneca's *Consolatio ad Marciam* en Tibullus' *Elegieën*⁵⁰. Wat de rijkdom van de woordenschat betreft, constateert Govaerts dat Tibullus' taal wat minder rijk is dan Seneca's, want bij de eerste ligt het gemiddeld gebruik wat hoger⁵¹. De spreiding van het vocabularium stelt ze voor als een semi-logaritmische curve. Als de curve naar boven of naar onder knikt, verraadt zich een overgang van de ene zone naar de andere. Zo blijkt Tibullus' gebruik van de *mots-outils* rijker dan dat van Seneca, terwijl diens taal armer is aan *mots de caractérisation*. Tibullus' veelvuldig herhalen van karakteristieke woorden verklaart Govaerts als een trek van diens persoonlijkheid en niet als een beperktheid van zijn lexicon: in verhouding tot Seneca gebruikt hij immers meer woorden slechts éénmaal. Als hij meer dezelfde woorden herhaalt, wijst dat op een zekere rechtlijnigheid in zijn temperament (in navolging van Delatte⁵² gebruikt ze het woord *obsession*), omdat hij namelijk dezelfde idee steeds door hetzelfde woord wil weergeven. Men moet er echter volgens haar ook rekening mee houden dat Tibullus niet een werk van één adem geschreven heeft zoals Seneca, maar afzonderlijke gedichten⁵³. Voor Griekse auteurs kan men dergelijke vocabulariumonderzoeken evengoed op touw zetten, maar een voorbeeld heb ik niet gevonden.

De „computerfilologie” waagt zich op de bekende paden van de traditionele filologie: de kritiek en de hermeneutiek. Boven is al gesproken over de computer in de tekstrestitutie en in de studie van de mededeling van een tekst. Een eerste benadering van de mededeling was het onderzoek van de woordkeuze in functie van het onderwerp, het genre en de persoonlijkheid van de auteur. De notie gemiddeld gebruik alsook grafieken en diagrammen als middel om de frequentie van taalverschijnselen te visualiseren zijn echter in de traditionele filologie minder gewoon. Bij andere benaderingswijzen van de mededeling doet de computerfiloloog een beroep op nog andere verworvenheden van de statistiek, namelijk allerlei tests. Zo komen bij de vergelijking van de zinsbouw

Een eenvoudige deling van het aantal items door het aantal lemmata levert telkens de te vergelijken waarden voor het gemiddeld gebruik op.

50. Cfr. Govaerts, *Quelques observations*, pp. 20-21.

51. Cfr. id., p. 21.

52. Cfr. supra, n. 45.

53. Dit is een résumé van het gehele artikel.

bij verschillende auteurs allerhande berekeningen kijken, die men eveneens door de machine kan laten uitvoeren.

Dat er tussen verschillende auteurs ook verschil is in de manier waarop ze hun woorden elkaar laten opvolgen, is al lang geweten. Nieuw is echter de manier waarop de kenmerken van de stijl gemeten worden. De conditio sine qua non van de nieuwe stylo-metrie is de referentiëring van alle voorkomende woorden naar hun plaats in lijn of vers, in zin, paragraaf en hoofdstuk. Zo kan men de computer gemakkelijk bijvoorbeeld laten nagaan hoe dikwijls vormen van het werkwoord vooraan, middenin of achteraan in de zin voorkomen en welke wijzen bij voorkeur een of andere positie in de zin innemen. Men kan zich voorstellen dat zich in die voorkeur verschillen laten gevoelen tussen verschillende auteurs en zelfs tussen verschillende delen van het werk van eenzelfde auteur. Althans van de laatste mogelijkheid is er een voorbeeld voorhanden, en nog wel voor het Grieks. Ik bedoel het artikel van Gilles Maloney over de frequentie en de plaats van de werkwoordsvormen bij Thoeukudides. Maar verder meer daarover. Een ander kenmerk van het computeronderzoek in de filologie is het vergelijken van de frequentie van een bepaald taalverschijnsel in een deel van een werk of in een bestanddeel van een corpus met zijn frequentie in het geheel van dat werk of van dat corpus, en het berekenen of een verschil tussen de verwachte en de vastgestelde occurrence zo miniem is, dat het slechts aan het toeval te wijten kan zijn, dan wel belangrijk genoeg om het aan een andere factor te moeten toeschrijven; bij één auteur kan die factor een verschil in onderwerp of genre zijn, in een corpus bestaande uit werken van verschillende auteurs zal die factor met de stijl van elke auteur verband houden. Of een verschil significant is of niet, wordt besloten op grond van een test uit de probabiliteitsleer: de χ^2 -test. Boven zijn twee computeronderzoeken van Latijnse werken vermeld (Seneca's *Consolatio ad Marciam* en Horatius' *Oden*) waaruit werd afgeleid dat er bij vergelijking van een bepaald deel met het geheel, in dat deel een belangrijke concentratie van woorden uit een bepaald semantisch veld optreedt (resp. dat van de leeftijden en dat van de zee). In verband met dit soort van vergelijkingen moet hier, zoals aangekondigd, even uitgebreid worden over een artikel van Maloney. In dat opstel wordt naast de frequentie van Thoeukudides' werkwoordsvormen in het verhaal, in de rede in de oratio obliqua en in de rede in de oratio recta ook de plaatsing van de werkwoordsvormen in elk van die „delen" nagegaan, namelijk hoeveel er telkens vooraan, middenin of achteraan in de zin voorkomen. Vergelijking van de aantallen in de delen met de aantallen in het geheel leidt tot conclusies waarover dadelijk meer. Bij

Maloney gaat het dus, zoals in de pas vermelde Latijnse onderzoeken, om één auteur en om vergelijking van delen met een geheel. Een voorbeeld van vergelijking van telkens één bestanddeel van een corpus met het gehele corpus vindt men in het opstel van Jacques De Bie over syntaxis en woordorde bij Isokrates, Thoekudides en Demosthenes. Daarin wordt de frequentie van een aantal taalverschijnselen in een pensum uit elk van die auteurs vergeleken met de frequentie die ze bereiken in de som van de pensa uit alle drie de auteurs. Maar eerst de verschuldigde resultaten van Maloney's onderzoek.

Bij Thoekudides verkiezen de werkwoorden van de hoofdzin of van de onafhankelijke zin het begin van de zin, zoals ook de losse genitief dat doet, terwijl de participia zich in het midden concentreren en de infinitieven aan het einde. Wat de relatiefzinnen en de ondergeschikte zinnen betreft, is er geen regelmaat in de plaats van de werkwoordelijke vormen te bespeuren. Als er echter van één genre naar een ander overgegaan wordt, treden er wel enkele veranderingen in de plaatsing op. In de indirecte stijl zijn de ondergeschikte zinnen wat minder frequent in het begin van de zin dan elders. Het verhaal onderscheidt zich van de andere twee groepen hierin, dat de betreffende bijzinnen er zelden vooraan in de volzin voorkomen en dat de participia even frequent zijn in het begin als in het midden van de zin. Nog meer dan de plaatsing van de werkwoordsvormen wisselt bij overgang van één genre naar een ander hun frequentie. Thoekudides gebruikt namelijk minder ondergeschikte zinnen in het verhaal dan elders, maar hij gebruikt er evenveel in beide soorten redevoeringen; het omgekeerde is het geval voor de hoofdzinnen en de onafhankelijke zinnen. Dat resultaat bevestigt het vermoeden van elke Thoekudideslezer: de hypotaxis is frequenter in de redevoering, de parataxis in het verhaal. Maloney constateert verder dat er meer infinitieven en minder losse genitieven voorkomen van zodra Thoekudides van het verhaal naar de redevoering overgaat, terwijl de overgang van het verhaal en van de rechtstreekse rede naar de indirecte stijl gepaard gaat met een vermindering van het aantal participia⁵⁴. De Bie's corpus bestaat uit twee redevoeringen van Isokrates (XIII en VII), de *epitafios* bij Thoekudides (II 35-46) en twee redevoeringen van Demosthenes (V en IV). In vergelijking met het gehele corpus blijkt Demosthenes lange woorden en lange zinnen

54. Cfr. Maloney, p. 106 (over de plaats van de werkwoordsvormen in het algemeen), p. 102 (over de plaats van de werkwoordsvormen bij overgang van één genre naar een ander), p. 104 (over de frequentie van de soorten zinnen in de drie genres), p. 90 en 106-107 (resp. over de frequentie van de infinitieven en losse genitieven, en van de participia).

te vermijden, in tegenstelling met Isokrates. Men verwacht dus complexere zinnen bij Isokrates, dat wil in de eerste plaats zeggen met meer werkwoordelijke vormen. Nochtans bevatten de behandelde teksten van Isokrates minder werkwoorden dan verwacht (die van Demosthenes veel meer). Dat gebrek aan werkwoorden bij Isokrates wordt echter niet gecompenseerd door meer adjectieven of substantieven (wat wèl het geval is bij Thoeukudides). Bij Isokrates komen ook veel minder vervoegde werkwoorden en infinitieven voor dan verwacht, en meer bij Demosthenes. Bij Isokrates zijn de participia integendeel veel sterker vertegenwoordigd dan verwacht en bij Demosthenes minder sterk. Bij Isokrates telt De Bie significant minder ondergeschikte en nevenschikte werkwoorden dan verwacht, en bij Demosthenes meer. Als Demosthenes' zinnen korter zijn, dan moet dat komen doordat ze minder elementen bevatten die met andere overeenkomen. Bij Isokrates is natuurlijk het omgekeerde waar : bij Isokrates zijn er veel meer met elkaar verbonden elementen⁵⁵. Isokrates' moeilijkheid zit dus in de „syntaxis” in de ware zin van het woord.

Een andersoortig computeronderzoek van de stijl is er een dat de opeenvolging van metrische elementen bestudeert. Daar ik wel over een voorbeeld beschik voor het Latijn, maar niet voor het Grieks, zal ik hier bondiger zijn. Het voorbeeld betreft een artikel van Pol Tordeur over Prudentius en enkele andere laat-Latijnse dichters van dezelfde periode, waaronder sommigen voor verzorgde stilisten doorgaan en anderen voor slordige. Voor elk van beide groepen heeft Tordeur machinaal een reeks metrische kenmerken verzameld ; nadien werd Prudentius' metriek onderzocht en werd nagegaan bij welke groep zich elk van de behandelde werken lijkt aan te sluiten. Het resultaat is dat Prudentius de groep van de „verzorgden” of van de „slordigen” vervoegt naar gelang van het genre waarin hij zijn onderwerpen giet. Tordeur geeft ook nog een conclusie over de plaats die zijn auteur dan wel inneemt onder de dichters van zijn tijd. De tegenstelling verzorgd-onverzorgd zal wel in nog andere literaire periodes terug te vinden zijn en men kan ook de stijlkenmerken diachronisch onderzoeken, namelijk binnen elk genre. Dat zal Tordeur echter in een andere publikatie uit de doeken doen⁵⁶.

Tot slot van dit deel over het gebruik van de computer in het onderzoek van zins- en versbouw kan hier het volgende geconcludeerd worden : de machine kan men inzetten om stijlverschillen te

55. Cfr. De Bie, pp. 21-25. Op 18 juni 1975 is Jacques De Bie te Leuven gepromoveerd op het proefschrift *Problemen, methoden en resultaten van statistisch en computationeel taalonderzoek, toegepast op de woordorde in klassiek Attisch proza*.

56. Cfr. Tordeur, pp. 20-21 en 32-33.

meten tussen delen van het werk van één enkele auteur of tussen verschillende auteurs die eerst tot één corpus verzameld werden; met hetzelfde hulpmiddel kan men de stijl tevens synchronisch bestuderen (bij auteurs van eenzelfde periode) of diachronisch (bij auteurs van eenzelfde genre).

Uit allerlei manueel en ook machinaal uitgevoerde stijlonderzoeken is voldoende gebleken dat de stijl verschilt van auteur tot auteur. Heeft iemand niet gezegd: „Le style est l'homme même.”? Als de stijl het auteurschap verraad, is het niet verwonderlijk dat de filoloog de stijl zal hanteren als criterium van authenticiteit. Wellicht vertrouwt de computerfilologie nergens méér op de mogelijkheden van de machine dan in de echtheidskritiek. Als vergelijkingsmateriaal voor elk verdacht deel van een werk of voor elk betwist integraal werk kunnen immers hele reeksen als authentiek beschouwde werken woord voor woord in het geheugen van de computer opgeslagen worden, waarna die gegevens naar wens vliegensvlug weer te voorschijn geroepen kunnen worden. Het komt er in authenticiteitsonderzoeken wél op aan de juiste discriminanten te vinden. Bij dergelijke onderzoeken op moderne auteurs is er in dat verband wel eens aan de lengte van de zinnen gedacht. Op grond daarvan de onechtheid bewijzen van een *antiek* werk (of een *passus* eruit) wordt echter voor onmogelijk gehouden. Het is immers zo, dat er in onze teksten geen interpunctie overgeleverd is. De interpunctie van onze uitgaven hangt dikwijls genoeg van de willekeur van de uitgevers af⁵⁷. Het vocabularium en de grammaticale categorieën voldoen als criterium evenmin. Het gebruik van het vocabularium en het voorkomen van de grammaticale categorieën verraden natuurlijk een en ander over de stijl, maar die factoren verschillen ook van het ene als authentiek beschouwde werk van eenzelfde auteur tot het andere en zelfs vertonen ze verschillen in frequentie binnen eenzelfde authentiek werk⁵⁸. Er moeten dus andere discriminanten gevonden worden, die men liefst

57. Cfr. Delatte, *A propos de la mesure*, pp. 51 en 62. Daarom en ook omdat de lengte van de zinnen erg kan wisselen, werkt Morton met als „lengtemaat” de tussenruimte tussen twee *kai's*, cfr. Michaelson & Morton, *The spaces in between*, pp. 23 en 26-27. Voor reserves tegenover de zinslengte als criterium zie ook Coppens-Ide, p. 188 en Kasher, pp. 13-14. Voor willekeur van de uitgevers inzake elisie zie Michaelson & Morton, *Elision*, pp. 42-43, waar ze beweren te kunnen aantonen wanneer een uitgever voor elisie verantwoordelijk is.

58. Cfr. supra, met noten 28-34, 43, 44 en 54. Voor homogeniteitsverschillen binnen de als authentiek beschouwde stukken van Tibullus zie Govaerts, *Le Corpus Tibullianum*, p. 293 (in verband met het vocabularium) en p. 321 (in verband met de grammaticale categorieën). Er zijn ook homogeniteitsverschillen geconstateerd voor het vocabularium in verschillende delen en rollen van Euripides' *Elektra*, cfr. Rigo, *Ajustement*, pp. 82-83 en voor vocabularium en enkele grammaticale categorieën in verscheidene rollen van Seneca's *Thyestes*, cfr. Germain, pp. 42-45.

gecombineerd op de betwiste tekst(en) zal toepassen. Of bepaalde taalverschijnselen als echtheidscriteria voldoen, moet blijken na ijking. Die ijking voert men uit aan de hand van de frequentie van die taalverschijnselen in werken die als authentiek worden aangezien. Om geldig te zijn, moeten die criteria aan volgende voorwaarden beantwoorden: verhoudingsgewijs mag er geen significant frequentieverschil zijn in hun occurrence over het geheel en over elk deel van eenzelfde werk, er mogen geen significante frequentieverschillen zijn in hun voorkomen in verschillende werken van eenzelfde auteur, maar er moeten er wél zijn tussen hun occurrence in werken van verschillende auteurs⁵⁹. Om de kans dat zijn discriminanten geldig zijn aanzienlijk te vergroten, zal de onderzoeker er ook voor zorgen dat hij werken uitkiest met verschillend onderwerp, met verschillende compositiedatum (als het werken van één auteur betreft) van verschillende literaire vorm (genre), met verschillende stijl (verzorgde en onverzorgde), enz.⁶⁰. Met die streng geselecteerde discriminanten gaat men ten slotte het probleem van de authenticiteit aanpakken. Verschilt voor een reeks geijkte criteria het betwiste werk (resp. de verdachte passus) op significante wijze van de authentieke werken van de betrokken auteur (resp. van de authentieke delen van het betrokken werk), dan neemt men een verschil van auteur aan. Het is dus mogelijk op grond van tests uit de probabiliteitsleer een werk aan een auteur *niet* toe te kennen. Het omgekeerde gaat echter niet aan. Twee werken die volgens een bepaald criterium niet of niet significant van elkaar verschillen, moeten daarom nog niet van dezelfde hand zijn. Gelijke lichaamslengte impliceert toch ook geen identiteit van de persoon⁶¹?

De aard van de discriminanten die in het computeronderzoek van de authenticiteit gebruikt kunnen worden alsook hun toepassing op echtheidsproblemen bij Griekse teksten kunnen hier wat nader omschreven worden. De lengte van de zinnen, het vocabularium en de grammaticale categorieën worden zoals gezegd als authenticiteitscriterium geweerd. Welke woorden er dan wél aanvaard? De Brit Rev. A. Q. Morton, die op Homeros, de geschriften van de apostel Paulus en Platoons *Zevende Brief* gewerkt heeft⁶², houdt geen rekening met betekenisdragende woorden omdat ze te zeer aan het onderwerp gebonden zijn. Hij verkiest woorden met hoge frequentie zoals sommige voorzetsels, sommige voornaamwoorden, het werk-

59. Cfr. Michaelson & Morton, *The spaces in between*, p. 27.

60. Cfr. id. *ibid.*, p. 27.

61. Cfr. id., *ibid.*, p. 28.

62. Cfr. Coppens-Ide, p. 193, n. 6 (bibliografie tot 1971), Brandwood, p. 22, n. 1, en Michaelson & Morton, *The spaces in between* en *Elision*.

woord *ênai* en sommige partikels⁶³. De frequentie van die criteria wordt beschouwd niet binnen de ruimte van één zin, maar binnen de ruimte tussen twee opeenvolgende *kai*'s. Al zijn discriminanten hanteert Morton in combinatie met elkaar. Zijn conclusies blijven nochtans niet onaangevochten. Zo concludeert Brandwood op basis van Mortons gegevens, maar in tegenstelling tot Morton, dat Platoons *Zevende Brief* wèl authentiek is. Morton zou in zijn berekeningen enkele redeneringsfouten hebben gemaakt⁶⁴. Mortons methode zelf valt Brandwood echter niet aan.

Na computeronderzoek van teksten in verband met de authenticiteit zijn volgende resultaten mogelijk: soms kunnen compleet verschillende auteurs verondersteld worden voor twee werken of voor delen van één werk; ingrijpende veranderingen in een werk kunnen soms aan een latere revisie te wijten zijn en in relatie met de context is het soms ook mogelijk onderliggende oudere bronnen te ontdekken⁶⁵. In verband met het bronnenonderzoek vraag ik me af of men de computermethodes ook op varienschrijvers zou kunnen toepassen. Zou het vaak grote aantal van de bronnen die ze geëxcerpeerd hebben, bekende en onbekende, niet verwarrend kunnen werken? Ook het computeronderzoek van fragmenten van verloren werken in hun verhouding tot overgeleverde is problematisch. Als een auteur gedachtengoed van een ander wil gebruiken, hoeft hij hem immers niet noodzakelijk letterlijk te citeren. Hij kan ook de inhoud van diens mededeling met min of meer succes in zijn eigen stijl weergeven. Een grammaticus citeert misschien wèl letterlijk, maar dan bij voorbeeld om één of andere stijleigenaardigheid, die bij de geëxcerpeerde auteur niet eens zo frequent

63. Cfr. Coppens-Ide, p. 190, Michaelson & Morton, *The spaces in between*, p. 23, en *Elision*, p. 33. Oorspronkelijk hielden echter ook Morton en zijn medewerkers rekening met de lengte van de zinnen, cfr. Brandwood, p. 1.

64. Cfr. Brandwood, pp. 20-21 (conclusie) en 1, 4-6 en 8 (redeneringsfouten). Andere kritiek: McArthur bij Radday & Shor, p. 40, n. 2.

65. Zo komt Morton tot de conclusie dat van de *Brieven van Paulus de Pastorale Brieven* (d.i. *Timotheus I en II* en *Titus*), *Thessalonicenzen I en II*, *Kolossenzen*, *Efeziërs* en *Filippenzen* niet van de hand van de apostel zelf zijn, cfr. Radday & Shor, p. 25. Zie ook Michaelson & Morton, *The spaces in between*, pp. 46-47 (*Efeziërs*, *Filippenzen* en *Kolossenzen*), p. 47 (*Thessalonicenzen I en II*), p. 48 (*Pastorale Brieven*), pp. 49-50 (*Hebreeërs*), en verder *Elision*, p. 45 (*Romeinen en Korinthiërs II*). *Korinthiërs II* bestaat uit Paulinisch materiaal en stukken met andere kenmerken, wellicht wegens een latere revisie, cfr. Michaelson & Morton, *The spaces in between*, pp. 45-46. Over stijlverschillen tussen Flavius Iosephus' *Joodse Oorlog*, wat hij eerst in het klad schreef, waarschijnlijk in het Aramees, waarna assistenten hem het werk in het Grieks hielpen omzetten aan de ene kant en diens *Joodse Oudheden*, wat hijzelf schreef, zie Michaelson & Morton, *Elision*, pp. 41-42. Over de revisie van Diodorus Siculus' *Historiai* cfr. id., *ibid.*, p. 41. Over twee oudere bronnen in *Korinthiërs I* zie Michaelson & Morton, *The spaces in between*, p. 42. Over de citaten en excerpten in Iosephus' *Vita* en *Contra Apionem* cfr. id., *Elision*, p. 41.

hoeft te zijn. Maar dergelijke problemen kent de traditionele filologie evengoed⁶⁶.

* * *

De titel en de inhoud van het eerste deel van dit opstel wekken bij de lezer allicht het dubbel vermoeden dat schrijver dezes zelf een frequentieonderzoek op zijn geweten heeft en dat hij voor zijn onderzoek een beroep gedaan heeft op de computer. Het eerste moet hier bevestigd worden, het tweede echter ontkend. Ondanks alle mogelijkheden van de computerfilologie heb ik me niet aan de computer zelf gewaagd. Daar waren een viertal redenen voor. Vooreerst was mijn eigen onderzoek in oorsprong te bescheiden van opzet: het pensum tekst ging niet te zwaar wegen en aan het materiaal zouden niet te veel vragen gesteld worden (het was een onderzoek voor onderwijsdoeleinden). Ten tweede wist ik in het begin niet dat er ook voor het Grieks een automatische analyse mogelijk was (ze was toen ook nog maar weinig gebruikt). Gaandeweg bleken er ook voor een bepaald onderzoeksgebied (de stamsystemen van het werkwoord) geen codes voorzien te zijn, en ten slotte werd het me ook duidelijk dat het voorbereiden van het materiaal voor de computer geen tijdsbesparing zou betekend hebben⁶⁷. Uit publikaties van computerfilologen viel er echter heel wat te leren. Cijfermateriaal uit computeronderzoeken kon ik best gebruiken om het met gegevens uit mijn eigen auteur te vergelijken; ook methodes van onderzoek kon ik eruit overnemen. Zo ben ik al lezende onder de indruk gekomen van de mogelijkheden van de computer in de filologie, wat meteen verklaart waarom ik me in deze bijdrage niet heb kunnen beperken tot enkel die methodes en die resultaten van het computeronderzoek, die rechtstreeks met mijn eigen frequentieonderzoek in verband gebracht konden worden.

Het frequentieonderzoek waarvan sprake, slaat op een keuze excerpten uit de Boeken I en IV van de *Anabasis* van Xenofon die in het Middelbaar Onderwijs het meest gelezen worden⁶⁸. Het corpus komt overeen met c. 62 bladzijden van een Budé-uitgave; nauwkeuriger uitgedrukt, bevat het pensum 11.148 items, die van

66. Enkele bedenkingen over de waarde van een fragment als staal van de stijl van een auteur vindt men bij Pennacini.

67. Over het tijdrovend aspect van de voorbereidingen voor computeronderzoek cfr. Greenberg, p. 2; Evrard, *La papyrologie*, p. 89; De Poerck, p. 104; Engels, pp. 250-251.

68. De uitgelezen fragmenten zijn: *An.* I 1, 2, 3, 4 (4-9 en 11-17), 5 (1-3 en 7-9), 6, 7 (10-13), 8, 9 en 10 (1-3 en 16-19); IV 1 (19-28), 2 (1-23), 3, 4 (11-13 en 19-22), 5 (1-33), 6 (10-16), 7 (8-14 en 19-27) en 8 (19-21 en 25-28). Voor een verantwoording van die keuze zie Goerlandt, *Een frequentieonderzoek*, pp. 157-158.

1.653 lemmata afhangen (zie tabel I). Dit frequentieonderzoek werd ondernomen op verzoek van de Rijksinspecteurs dr. F. De Raeve en wijlen dr. P. Noyen, en onder de hogere leiding van Prof. dr. H. Van Looy. Het was de bedoeling na te gaan wat er bij de lectuur van de eerste Griekse schoolauteur zo frequent is aan vocabularium, morfologie en syntaxis, dat het reeds in de initiatiecursus aangeleerd moet worden. Het probleem was bijzonder acuut sinds het invoeren van het Vernieuwd Middelbaar Onderwijs, waarvan het leerplan voor het eerste jaar Grieks (in het tweede oriënteringsjaar) de *frequente* taalverschijnselen voorschreef zonder exhaustief te kunnen opsommen waarin die frequente taalverschijnselen dan wel bestonden.

Voor het excerpteren van een frequentievocabularium zijn er tot hiertoe verscheidene methodes aangewend⁶⁹. Hoewel de LASLA-methode ook niet helemaal volmaakt is⁷⁰, vind ik ze toch de wetenschappelijk meest verantwoorde. Ze werd geconcipeerd door Etienne Evrard, en uitgewerkt door Arthur Bodson en verder door Suzanne Govaerts en Joseph Denooz⁷¹. In haar meest gedegen vorm ziet ze eruit als volgt. De woorden die in een tekst onder verbogen of vervoegde vorm voorkomen (de *items*), worden gegroepeerd onder hun respectief semantisch trefwoord of *lemma*. Een fictief voorbeeld voor het Grieks: vijfmaal *basileus*, driemaal *basiléōs* en tweemaal *basileî* betekent één lemma *basileus* met tien items. Natuurlijk zijn er ook onverbuigbare woorden, zoals bij voorbeeld *pollakis*. In dat geval zijn lemma en item identiek. Voor het excerpteren van een frequentievocabularium volgens LASLA worden alle lemmata gerangschikt in dalende lijn van frequentie. De frequentie van het lemma met het grootste aantal items (bv. 400) vertegenwoordigt van het totale aantal items van een bepaalde tekst (bv. 10.000) een zeker percentage (in dit fictief voorbeeld: 4%), de som van de occurrenties van het lemma met het grootste aantal items en van dat met het op één na grootste aantal items vertegenwoordigt een wat groter percentage (bv. 400 plus 350 op 10.000, of 7,5%), enz. Het numeriek belangrijkste lemma van een bepaalde tekst vertegenwoordigt van het totale aantal lemmata (bv. 1.500) slechts een gering deel (in ons fictief voorbeeld: 1 op 1.500, of 0,066...%), het belangrijkste en het op één na belangrijkste samen een iets groter deel (hier 2 op 1.500, of 0,133...%). Het voorkomen van één of meer lemmata met een

69. Cfr. Goerlandt, *Een frequentieonderzoek*, pp. 151-157.

70. Cfr. Goerlandt, *Frequentieonderzoek en ...*, p. 16 (het probleem van de composita; zie hierover nu ook Hessinger, p. 23 van een onder n. 100 geciteerd artikel).

71. Cfr. Delatte, *Le Laboratoire*, p. 103; Evrard, *Le Laboratoire*, pp. 111-112; Bodson; Govaerts & Denooz, *Lexique*.

gelijk aantal items brengt voor de berekening geen complicaties teweeg. Als er in onze lijst van lemmata bijvoorbeeld op de derde plaats van belangrijkheid twee lemmata staan met een frequentie 325, dan stelt de som van de tot hiertoe besproken lemmata 4 eenheden op 1.500 voor, of 0,266... %, terwijl de som van alle overeenkomstige items nu 1.400 eenheden op 10.000 bedraagt, of 14 %. Naarmate verder wordt afgedaald, stijgen de procenten in de reeks van de items veel sneller dan die in de reeks van de lemmata : in het begin van de lijst komen met relatief veel items relatief weinig lemmata overeen. Op de duur gaat de stijging van de procenten in de reeks van de lemmata echter met de stijging van de procenten in de reeks van de items genivelleerd worden. Die nivellering treedt des te sneller in, naarmate er steeds meer lemmata geteld worden die slechts weinig items dekken⁷². De Luikse onderzoekers gaan er nu van uit dat de leerling op de lectuur van een tekst wordt voorbereid door vooraf alle belangrijke lemmata in te studeren. Ze stellen echter vast dat er een limiet is voorbij dewelke de leerling veel nieuwe lemmata zou moeten memoriseren om van de voorbereide tekst slechts een gering aantal nieuwe items te kunnen begrijpen. Die rendementslimiet ligt eigenaardig genoeg in alle onderzochte teksten steeds omstreeks 85 % van het totale aantal items⁷³. Daarmee komt naar gelang van de stijl van de auteur meer of minder dan grosso modo 30 % van het totale aantal lemmata overeen⁷⁴. De laagste frequentie tot dewelke men in de reeks van de lemmata is afgedaald om in de reeks van de items de rendementslimiet te bereiken, wordt de *limiet-frequentie* genoemd. Die zal ook afhankelijk zijn van de lengte van de geëxcerpeerde tekst. Wat nu de praktische uitwerking van hun (Latijnse) frequentievocabularia betreft, gaan de Luikse onderzoekers als volgt te werk. Eerst lemmatiseren ze alle items uit een aantal werken voor middelbare schoollectuur (hetzij integrale werken, hetzij excerpten) en bepalen ze op de zожuist geïllustreerde manier voor elk onderzocht werk afzonderlijk de limietfrequentie. Vervolgens worden alle vocabula die in *elk* onderzocht werk frequent zijn, in één *gemeenschappelijk* of *basisvocabulary* gegroepeerd. Vocabula die in één bepaald werk frequent genoeg zijn (en eventueel ook in andere), maar niet in *alle* geëxcer-

72. Reële voorbeelden van deze methode vindt men in Govaerts & Denooz, *Lexique* en Goerlandt, *Lexique*.

73. Cfr. Govaerts & Denooz, *Lexique*, p. 17: „au-delà de 85 %, le gain de compréhension est très faible comparativement à l'effort de mémorisation qu'il exigerait”.

74. In de publikatie vermeld onder n. 49 kan men goed nagaan welke percentages van de items het dichtst de rendementslimiet benaderen en met welk percentage van de lemmata de „renderende” items telkens overeenkomen.

peerde werken, worden ondergebracht in een *bijzonder vocabularium* of *microlexicon* voor dat werk.

Voor mijn eigen woordfrequentieonderzoek op Xenofoon heb ik de methode van het LASLA gevolgd. Met de empirisch vastgestelde rendementslimiet van 85 % voor ogen kon de limietfrequentie voor het corpus van de excerpten op 4 bepaald worden. Van de 1.653 lemmata bereiken er 460 de frequentie 4 of meer, dat is slechts 27,83 %. Van die vocabula hangen niet minder dan 9.428 van de 11.148 items af, of 84,57 %⁷⁵. Die 460 woorden vormen natuurlijk geen echte *basiswoordenschat*, omdat er geen andere Griekse schoolauteurs bij kwamen kijken. Het is best mogelijk dat echt belangrijke woorden in mijn uittreksels toevallig de limietfrequentie niet halen, terwijl woorden die normaal in een microlexicon voor de *Anabasis* zouden thuishoren (of eigen zijn aan de historiografie in het algemeen), in de excerpta de limietfrequentie toevallig wél bereiken⁷⁶.

Wie een tijdje in de literatuur omtrent frequentieonderzoek graa duint, moet onvermijdelijk op het probleem van de homogeniteit stoten. Zo kwam ik ertoe bij elk lemma uit mijn corpus na te gaan hoe frequent het is in de uittreksels uit Boek I en in die uit Boek IV. In volgende voorbeelden vindt men tussen haakjes de aldus gedifferentieerde frequenties: *adelphos* (6-0), *anthrōpos* (6-12), *kōmē* (1-14), *lophos* (0-9), *nux* (1-11), *hodos* (9-20), *polis* (29-1), *pûr* (0-10), *hupozugion* (1-14), *khiōn* (0-14), *khrēma* (12-0), *kbronos* (5-1). Was de *Anabasis* volkomen homogeen, dan zou volgens de verhouding van de sommen van de items (7.175-4.823) de verhouding telkens 6-4 moeten zijn. In de frequenties van bovenstaande reeks betekenisdragende woorden herkent men duidelijk de sporen van het verschil in onderwerp tussen beide ge-excerpeerde Boeken. Die voorbeelden illustreren tevens dat het eigenlijk fout zou zijn bij de berekening van de limietfrequentie geen onderscheid te maken tussen de „Boeken” I en IV. Verscheidene vocabula die in beide „Boeken” samen merkkelijk meer dan de limietfrequentie halen, komen alleen of haast uitsluitend in de excerpten uit Boek IV voor. Om te vermijden dat een leerling die alleen uittreksels uit het Eerste Boek zal lezen, „verloren” zou leren, werd een limietfrequentie berekend voor de excerpten uit Boek I en voor die uit Boek IV afzonderlijk. De 220 vocabula die zowel in de ene als in de andere groep de limietfrequentie bereiken,

75. Deze en verdere cijfers steunen op een verbeterde versie van mijn gegevens. Lichte verschillen met de getallen in mijn *Lexique* zijn mogelijk.

76. Daarom is het interessant mijn resultaten te vergelijken met het basisvocabularium van Meyer & Steinthal. Voor een bespreking van hun basislexicon zie Goerlandt, *Een frequentieonderzoek*, pp. 151 en 156-157.

vertegenwoordigen telkens een klein percentage van het aantal lemmata, maar nog een aanzienlijk aantal items⁷⁷.

Het dozijn betekenisdragende woorden dat zopas als voorbeeld werd gegeven van het gebrek aan homogeniteit binnen de *Anabasis* mag niet de indruk wekken als zou het met de grammaticale woorden anders gesteld zijn. Enkele voorbeelden om dergelijke indruk weg te werken: *eis* (76-20), *en* (46-41), *epi* (45-57); *epei-toen* (26-23), *epeidē-toen* (5-11); *anō* (1-8), *autoū* (3-6), *euthus* (4-11), *prōton* (8-0); *an* (38-13), *ge* (13-0), *dē* (18-20).

Tussen de excerpten uit de Boeken I en IV is er niet alleen op het niveau van de woorden een gebrek aan homogeniteit vast te stellen, maar ook op het niveau van de grammaticale categorieën. De verdeling van de woordenschat per grammaticale categorie voor de gezamenlijke uittreksels is voorgesteld in tabel I. Na differentiëring van de frequenties werden tussen de „Boeken” I en IV significante verschillen vastgesteld voor lidwoord, voornaamwoorden en telwoorden. In het „Eerste” Boek komt het lidwoord minder voor dan verwacht, voornaamwoorden en telwoorden integendeel meer dan verwacht. Voor het „Vierde” Boek geldt natuurlijk het omgekeerde. Het overschot aan voornaamwoorden in het „Eerste” Boek is vooral te wijten aan het grote aantal persoonlijke voornaamwoorden, terwijl bij de telwoorden vooral de hoofdtelwoorden voor significante verschillen zorgen⁷⁸.

Als er op het gebied van het vocabularium belangrijke divergenties kunnen bestaan tussen delen van eenzelfde werk van één auteur, dan moeten er *a fortiori* ook divergenties voorkomen tussen werken van verschillende auteurs. Het voorkomen van afwijkingen tussen verschillende auteurs op het niveau van de grammaticale categorieën is reeds gesignaleerd bij de bespreking van frequentieonderzoeken in het eerste deel van dit opstel. Wat betreft de verschillen op het niveau van de woorden wou ik nog even terugkomen op wat hierboven vermeld is in verband met de grammaticale woorden. Brandwood heeft voor enkele werken van Platoon opgemerkt dat er belangrijke verschillen zijn in de occurrence van *kai* en *de*. Mag men dan ook niet veronderstellen dat er tussen *verschillende* auteurs divergenties zullen bestaan in het voorkomen van hun grammaticale woorden? Het zou inderdaad een vergissing zijn te menen dat auteurs die over verschillende onderwerpen schrijven alleen hun betekenisdragende woorden op verschillende wijze zouden gebruiken. De enige gegevens om dat enigszins te illustreren heb ik uit twee artikels van de computeradepten

77. Voor meer details zie Goerlandt, *Een frequentieonderzoek*, pp. 159-161.

78. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 47, n. 28.

Michaelson en Morton moeten puren (zie tabellen II en III)⁷⁹. Door de verhouding tussen de occurrenties van *alla* en *ou* bij enkele auteurs (tabel II) en door de gemiddelde frequenties van *alla* in enkele andere tekstgeheelen (tabel III) blijkt dat althans die grammaticale woorden over de Griekse literatuur niet homogeen verdeeld zijn.

Zoals gezegd, beperkt dit frequentieonderzoek op de *Anabasis* van Xenofoon zich niet tot het vocabularium. Ook de morfologie en de syntaxis werden, althans gedeeltelijk, onderzocht. Voor zover mogelijk, werden ook de resultaten gerelativeerd door vergelijking met resultaten uit andere frequentieonderzoeken.

Van de morfologie werd reeds dat deel onderzocht, dat de leerling wel de meeste moeilijkheden oplevert, namelijk de morfologie van het werkwoord. De spreiding van de werkwoorden-items over enkele grammaticale categorieën (wijs, tijd, vox, persoon, stamsysteem) zijn voorgesteld in een aantal tabellen (tabellen IV-VIII). Telkens werden absolute waarden en procenten medegedeeld. Het belang van dergelijk cijfermateriaal overstijgt het louter wetenswaardige: het is duidelijk dat een Griekse initiatiecursus die op de frequente taalverschijnselen afgestemd is, met de occurrenties van de grammaticale categorieën rekening moet houden. Van wetenschappelijk standpunt is het interessant te constateren dat van de stamsystemen het praesenssysteem in de onderzochte passussen met het leeuwenaandeel wegloopt (haast 2/3 van het aantal werkwoorden-items). Dat systeem *moet* wel het belangrijkste zijn, omdat het als enige in de rij alle wijzen, tijden en voces kan aannemen⁸⁰. Bij Xenofoon is het praesenssysteem echter nog belangrijker dan verwacht, en wel hierdoor. Bij hem komen er veel historische praesentia en beschrijvende imperfecta voor, die de plaats innemen van narratieve aoristen⁸¹. Het spreekt vanzelf dat het veelvuldig gebruik van het historisch praesens en het beschrijvend imperfectum de frequentie van het praesenssysteem opdrijft en die van de systemen waaronder de aorist kan vallen

79. Cfr. Michaelson & Morton, *Elision*, pp. 45-55 en *The spaces in between*, pp. 55-59.

80. In H. W. Smyth, *Greek Grammar*, p. 109 vindt men deze indeling van de werkwoordelijke morfologie in negen stamsystemen: I. praesens-syst. (praes. en impf.), II. futurum-syst. (fut. A en M), III. eerste aorist-syst. (1ste ao. A en M), IV. tweede aorist-syst. (2de ao. A en M), V. eerste perfectum-syst. (1ste perf., 1ste plqpf. en fut. ex. A), VI. tweede perfectum-syst. (2de perf. en 2de plqpf. A), VII. perfectum mediaal-syst. (perf., plqpf. en fut. ex. M en P), VIII. eerste passief-syst. (1ste fut. P en 1ste ao. P), IX. tweede passief-syst. (2de fut. P en 2de ao. P).

81. Cfr. Herrmann, p. 213; van alle Griekse auteurs uit de klassieke tijd gebruikt Xenofoon in zijn *Anabasis* trouwens het grootste aantal praesentia historica, cfr. id., p. 208.

(vooral III en IV), doet dalen. In verband met de stamsystemen moet nog opgemerkt worden dat de grammatica van Smyth met de term „2de aorist-systeem” (systeem IV) de aoristus secundus en de aoristus tertius bedoelt. Na splitsing van dat systeem in zijn twee componenten werd vastgesteld dat de echte tweede aorist niet minder frequent is dan de eerste aorist, wat interessante implicaties kan hebben voor de methodiek: wegens de frequentie van de tweede aorist, de gelijkenis van zijn uitgangen met die van het imperfectum en het verschil in aspect van beide tijden is het verantwoord de tweede aorist in te voeren tegelijk met het imperfectum⁸².

Zojuist werd er wel op gewezen dat een Griekse initiatie cursus in de geest van de nieuwe leerplannen rekening moet houden met de frequentie van de taalverschijnselen. Er is echter nog niet geantwoord op deze voor de hand liggende vraag: welke dan wel de „essentiële” wijzen, tijden, voces en personen zijn. Het ligt al even voor de hand de oplossing te gaan zoeken in de richting van de empirisch vastgestelde rendementslimiet van 85 % zoals hierboven beschreven voor het vocabularium. Volgens die methode zouden de frequentste wijzen (resp. tijden, voces en personen) die samen 85 % van het aantal werkwoorden-items omvatten, „essentieel” zijn. De tabellen tonen het volgende aan: in dalende lijn van frequentie zijn aldus van de *wijzen* de indicatief, het participium en de infinitief het belangrijkste, van de *tijden* het praesens, de aorist en het imperfectum, van de *voices* het actief en het mediaal, en van de *personen* de derde enkelvoud en de derde meervoud. Als limietfrequentiebepaling is dergelijke werkwijze nochtans waardeloos, omdat die berekening op vier verschillende categorieën wordt toegepast. Een echte limietbepaling is volgens de LASLA-methode alleen dan mogelijk, wanneer met slechts één soort gegevens gewerkt wordt. Is het niet zo, dat bij de boven beschreven methode voor het vocabulariumonderzoek met slechts één klasse gegevens wordt geopereerd, namelijk de lemmata? Zoals het lemma de semantische kern was bij het onderzoek van de woordenschat, zo moest er nu een morfologische kern worden gevonden. Als kern voor de werkwoordelijke morfologie werd eerst gedacht aan de combinatie *wijs-tijd-vox-(persoon)*. Men kan inderdaad alle werkwoordelijke items herleiden tot elementen als: indicatief-imperfectum-mediaal-derde persoon enkelvoud, infinitief-aorist-actief, enz. Binnen de vakjes welke door die drie of vier categorieën worden omschreven, treffen er nochtans vormen samen die morfologisch verschillend gebouwd zijn. Een fictief voorbeeld: zo kan

82. Dat doet Holtermann in zijn *Ianua Linguae Graecae* vanaf les 9.

epaideueto samentreffen met *etheâto*, *ekhrêto*, *edeïto*, *histato*, enz.; *paideûsai* kan het gezelschap krijgen van *timêsai*, *kalêsai*, *baleîn*, *doûnai*, enz. Er moet dus ook rekening gehouden worden met de stamsystemen en hun morfologische wetten. Uiteindelijk werd als kernelement voor de werkwoordelijke morfologie de formule *stamsysteem-morfologische wet-wijs-tijd-vox-(persoon)* aangenomen⁸³. Zoals er bij het vocabulariumonderzoek eerst werd nagegaan hoeveel items er van elk lemma afhangen, zo moet er nu telkens nageteld worden hoeveel items er bij de verschillende morfologische kernelementen horen. Vervolgens moeten de morfologische kernelementen in dalende lijn van frequentie gerangschikt worden, om zoals boven beschreven, weer een limietfrequentie te kunnen berekenen.

Bij het bepalen van een limietfrequentie voor de morfologie kan men zich natuurlijk niet beperken tot de werkwoordelijke morfologie. Daarnaast zijn er ook de naamwoorden, waarvoor er een morfologische kern moet worden gevonden, alsook de onverbuigbare items. Een echte limietfrequentiebepaling zou dus rekening moeten houden met de werkwoordelijke en de naamwoordelijke morfologie, alsook met de onverbuigbare woorden. Tot hiertoe werd echter enkel de werkwoordelijke morfologie onderzocht. Een gedetailleerde uiteenzetting daarover hoop ik elders te publiceren⁸⁴. Wel kan daarvan hier een soort algemene conclusie meegedeeld worden. Belangrijk zijn: de vervoeging van het regelmatig paradigma *paideuô* (behalve het futurum passief en het perfectum), de vervoeging van de contracta (behalve *douloô*), de tweede aorist op *-on/-omên*, *eimi-zijn*, *histêmi*, de werkwoorden op gutturaal- of dentaalstam (vooral in de aorist), alsook *eîmi-gaan* (als praesens).

Homogeniteitsverschillen werden er ook voor de morfologie vastgesteld, meer bepaald binnen enkele grammaticale categorieën van het werkwoord. In de spreiding van de *voces* over de excerpten uit het Eerste en uit het Vierde Boek vallen er geen noemenswaardige afwijkingen van de theoretische (6-4)-verhouding waar te nemen. Significante verschillen met de theoretische waarden werden wél vastgesteld bij één wijs, één tijd en drie personen. Bij de *wijzen* wordt de bijzondere positie ingenomen door het participium. Die wijs komt namelijk in de uittreksels uit het Eerste Boek in belangrijke mate minder frequent voor dan in die uit het Vierde. Bij de

83. Cfr. Goerlandt, *Frequentieonderzoek en ...*, pp. 18-22.

84. De werkwoordelijke morfologie van de onderzochte *Anabasis*-excerpten werd intussen uitvoerig behandeld in een publikatie uitgegeven door het Ministerie van Nat. Opv. en Ned. Cult., cfr. Goerlandt, *Resultaten*, pp. 9-18 en 29-54. Op pp. 19-25 en 55-72 van hetzelfde werk werd nogmaals aandacht besteed aan het frequentievocabularium.

tijden geldt dezelfde vaststelling voor het imperfectum. In verband met de *personen* doet zich een merkwaardig balanseffect voor. Aan de ene kant zijn de eerste en de derde persoon enkelvoud op significante wijze frequenter in het „Eerste” Boek dan in het „Vierde”; aan de andere kant komt de derde persoon meervoud in belangrijke mate méér voor in het „Vierde” dan in het „Eerste” Boek⁸⁵. Die verschillen in de spreiding van de persoonsvormen over de twee geëxcerpeerde Boeken moeten aan de inhoud te wijten zijn. In het „Eerste” Boek is het vooral Kuros die handelt en spreekt (resp. in de derde en de eerste persoon enkelvoud), terwijl het „Vierde” Boek eerder (in de derde persoon meervoud) een relaas geeft van de lotgevallen van Grieken en barbaren.

Als de *Anabasis* ook op morfologisch gebied geen homogeniteit vertoont, mag men ook weer van auteur tot auteur een verschillende spreiding van de morfologie verwachten. Dat gebrek aan homogeniteit kan hier aangetoond worden door vergelijking van gegevens bij Xenofoon met enkele resultaten uit vier computeronderzoeken. Daar van de morfologie bij Xenofoon enkel de werkwoordelijke grondig werd onderzocht, kan een vergelijking alleen daarop betrekking hebben.

In De Bie's corpus uit Isokrates, Thoeukudides en Demosthenes, waarvan hierboven al sprake was, dekken de werkwoorden-items 21,2 % van het totale aantal items. In mijn gezamenlijke excerpten uit Xenofoon is de verhouding 24,9 %. Dat morfologisch verschil tussen beide corpora moet significant genoemd worden. Verwonderlijk is dat verschil niet, als men weet dat er zelfs binnen de tekstenverzameling van De Bie significante verschillen optreden⁸⁶.

Als men de interimresultaten van een bepaald onderzoek van Alfred Tomsin ter vergelijking mag nemen, dan zou het gebruik van de wijzen in de Griekse documentaire papyri haast uitsluitend significante verschillen vertonen met hun gebruik in mijn corpus uit Xenofoon. Nochtans lijkt hier voorzichtigheid geboden, aangezien Tomsin's resultaten op slechts een zeer klein aantal werkwoorden-items betrekking hebben⁸⁷.

Enkele gegevens uit het eerder al besproken artikel van Maloney bieden meer houvast. Na vergelijking van het aantal werkwoordsvormen bij Xenofoon en Thoeukudides blijkt de eerste veel rijker aan vervoegde werkwoorden dan de tweede. De participia (met

85. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 50, n. 50.

86. Voor het percentage van de werkwoorden-items in het corpus De Bie cfr. De Bie, p. 22, sub (5) en voor de homogeniteitsverschillen zie supra.

87. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 24 en p. 51, n. 56. Mijn Xenofoononderzoek betreft 2.747 werkwoorden-items (of 2.777, de verbaaladjectieven inbegrepen), het onderzoek van Tomsin slechts 139, cfr. Tomsin, p. 65 en pp. 59-61 (verantwoording van de keuze).

inbegrip van de losse genitieven) en de infinitieven daarentegen waardelijke, betrekkelijke en temporele, mededelende zinnen met *hoti* of *hōs*, en gevolg-, doel- en redengevende zinnen). Er werd zijn veel talrijker bij Thoekudides dan bij Xenofoon. Van al zijn werkwoordsvormen zijn trouwens bij Thoekudides ruim de helft hetzij participia hetzij infinitieven; bij Xenofoon ligt dat percentage beduidend lager⁸⁸.

Over het voorkomen van de tijden van de indicatief en het gebruik van de personen van de indicatief in de *Elektra* van Euripides zijn er waardevolle gegevens verzameld door G. Rigo⁸⁹. Van alle tijden van de indicatief blijken in de *Elektra* en in het corpus uit de *Anabasis* alleen de frequenties van de aorist ongeveer overeen te stemmen. Euripides gebruikt in vergelijking met Xenofoon veel vaker het praesens, het futurum (simplex en exactum) en het perfectum (en plusquamperfectum); het imperfectum daarentegen is veel belangrijker bij Xenofoon. Van de personen is de frequentste bij Xenofoon de derde enkelvoud, in aantal gevolgd door de derde meervoud. In Euripides' *Elektra* is de voornaamste persoon eveneens de derde enkelvoud, maar de tweede en de derde plaats worden ingenomen door de eerste en de tweede persoon enkelvoud, terwijl de derde persoon meervoud pas op de vierde plaats komt⁹⁰. Al die belangrijke verschillen in het voorkomen van de werkwoordelijke morfologie moeten verband houden met het verschil in onderwerp en genre: Xenofoon schrijft (vooral in de derde persoon) over verleden gebeurtenissen, terwijl een tragedie (in de eerste persoon) personages laat optreden en stelling laat kiezen tegenover gebeurtenissen in verleden, heden en toekomst, en tegenover elkaar⁹¹.

Hierboven werd uiteengezet hoe de LASLA-methode voor het samenstellen van een basisvocabularium op de morfologie kan worden toegepast. In een voordracht voor de discussiegroep „V.S.O. en Grieks” op het *Colloquium didacticum classicum quintum* (Brussel, 1973) heb ik ook een mogelijke limietbepaling voor de syntaxis voorgesteld. Daar er echter met de voorgestelde „syntactische kern” nog niet geopereerd werd en de beschrijving ervan niet in een bondige formule gegoten kan worden (wat wél het geval is voor de morfologische kern), verwijs ik daarvoor liever naar de publikatie van die toespraak⁹². Van de syntaxis werd nochtans wel

88. Cfr. Goerlandt, *Frequentieonderzoek en...*, pp. 24-25 en p. 52, n. 59 en 60.

89. Cfr. Rigo, *Projet*.

90. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 25 en p. 52, n. 61 en 62.

91. Voor een gedetailleerde verklaring van het gebruik van tijden en personen in de indicatief in Euripides' *Elektra* zie Rigo, *Projet*, pp. 51-52.

92. Cfr. Goerlandt, *Frequentieonderzoek en...*, pp. 26-27 (sub 2.3.1.).

een en ander onderzocht, namelijk de frequentie van participium- en infinitiefconstructies, alsook van een aantal bijzinnen (voornagegaan of er voor die syntactische verschijnselen geen significante frequentieverschillen bestaan tussen de uittreksels uit het Eerste Boek van de *Anabasis* en die uit het Vierde aan de ene kant en tussen de gezamenlijke uittreksels uit Xenofoon en een soort doorsnee-Grieks aan de andere kant. Met de naam *doorsnee-Grieks* mag men wel het corpus teksten betitelen waarvan Friedrich Maier de syntaxis heeft onderzocht⁹³. Het omvat de 97 Griekse versio's die in de jaren 1901-1964 aan Beierse abiturienten werden voorgelegd en die gekozen zijn uit 22 prozaschrijvers van de Attische redenaars tot Prokopios⁹⁴. Tot slot werden nog enkele gegevens uit De Bie en Maloney met gegevens uit Xenofoon vergeleken.

Voor de onderzochte syntactische verschijnselen werden er niet veel homogeniteitsverschillen geconstateerd tussen de „Boeken” I en IV. De verschillen vertonen om welke reden ook steeds hetzelfde patroon: telkens zijn de afwijkende syntactische fenomenen in belangrijke mate minder frequent dan verwacht in het „Eerste” Boek en meer in het „Vierde”. De divergenties worden aangetroffen in de participium- en de infinitiefconstructies, niet in de onderzochte bijzinnen. Bij de participia gaat het om de temporaal gebruikte participia coniuncta, om de attributief gebruikte participia in het algemeen en de gesubstantiveerde attributieve participia in het bijzonder. Bij de infinitiefconstructies betreft het alleen de accusativus cum infinitivo bij onpersoonlijke uitdrukkingen⁹⁵.

Talrijker zijn de verschillen tussen de syntaxis van de *Anabasis* en die van het doorsnee-Grieks van Maier. Een uitvoerige vergelijking van beide corpora vindt men in mijn colloquiumvoordracht⁹⁶.

De Bie en Maloney hebben in tegenstelling tot Maier met de computer gewerkt. Daar hun onderzoek niet zoals dat van Maier voor onderwijsdoeleinden bestemd was, vallen er bij wijze van spreken, althans wat de syntaxis betreft, slechts weinig kruimels van hun tafel. Alleen dit. De Bie heeft genoteerd dat de onderschikkende voegwoorden in Demosthenes' redevoering *De Vrede* en in diens *Eerste Philippische* (samen) 3,12 % van het totale aantal items vertegenwoordigen; in Isokrates' *Areopagitikos* is dat

93. De term *doorsnee-Grieks* wordt niet door Maier zelf gebruikt. De resultaten van zijn syntactisch onderzoek vindt men in Maier, pp. 89-93.

94. Cfr. id., *ibid.*, p. 87, n. 2 en p. 88.

95. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 33 en p. 55, n. 85 en 86.

96. Cfr. id., *ibid.*, pp. 29-31.

slechts 1,94 %⁹⁷. Voor Xenofoon vond ik nog een andere verhouding, namelijk 2,91 %. Volgens Maloney vertegenwoordigen de losse genitieven in zijn corpus uit Thoe kudides 2,6 % van het aantal werkwoorden-items⁹⁸. Bij Xenofoon is de losse genitief met zijn 1,16 % van de werkwoorden-items minder dan half zo vaak gebruikt. Niet alleen in vergelijking met Thoe kudides schijnt Xenofoon met de losse genitief zuinig om te springen, maar ook in vergelijking met de 22 auteurs uit Maiers doorsnee-Grieks: in het corpus excerpten van Maier komt die constructie verhoudingsgewijs in significante mate meer voor dan in mijn excerpten uit de *Anabasis*⁹⁹.

Als besluit bij dit punt over de syntaxis mag dus wel genoteerd worden dat er inzake syntaxis evenmin als inzake vocabularium en morfologie in de Griekse literatuur enige homogene verdeling verwacht mag worden¹⁰⁰.

* * *

97. Cfr. De Bie, p. 27, n. 8 en 11.

98. Cfr. Maloney, p. 90 (tableau I B).

99. Cfr. Goerlandt, *Frequentieonderzoek en...*, p. 53, n. 74, sub (b).

100. Sinds het afsluiten van deze bijdrage kwamen me nog volgende publikaties onder ogen die met mijn onderwerp in verband staan:

- Delatte L., Govaerts S. & Denooz J., *Quelques mots des activités du L.A.S.L.A.*, in *Revue*, 1975/4, pp. 1-5;
- Duchesne-Degey M., *Un programme d'ordinateur pour la coupe syllabique en latin*, in *Revue*, 1975/1, pp. 37-57 (om Latijnse én Griekse woorden te splitsen in lettergrepen als aanloop tot stilistisch onderzoek);
- Duroisin P., *Un usage nouveau des index mécanographiques*, in *Revue*, 1975/1, pp. 59-69 (gebruik van de LASLA-indices op Seneca philosophus om na te gaan wat van diens gedachtengoed Henry de Montherlant gebruikt heeft en hoe);
- Hessinger J. J., *A statistical supplement to an index verborum*, in *Revue*, 1975/4, pp. 11-56 (een methode om meer linguïstische informatie te puren uit een LASLA-index op Caesar de bello Gallico);
- Maloney G. & Villeneuve Winnie F., *Mots longs et mots courts en grec*, in *Revue*, 1975/3, pp. 1-34 (definitie van lange, middellange en korte woorden en hun frequentie in alle traktaten van het *Corpus Hippocraticum* behalve XXIII en XXX);
- Mattea Michela, *Statistical researches in the „verbum“ lexical field on the Frontonian rhetorical works, De orationibus and De eloquentia*, in *Revue*, 1975/3, pp. 35-[55] (een lijst van Fronto's key-words bewijst welk belang hij hechtte aan het woord en aan de welsprekendheid; hij moet Marcus Aurelius in diens retorische studies danig gestimuleerd hebben);
- Nicolova-Burova Anna, *On the chronology of Seneca's philosophical dialogues*, in *Revue*, 1975/2, pp. 1-30 (op basis van bepaalde klassieke stijlcriteria alsook van de frequentie en de spreiding van de woordsoorten wordt een stijlevolutie geschetst van *De ira* I en II over *De ira* III naar *De clementia*);
- Packard D. W., *A Greek computer at Chapel Hill*, in *Revue*, 1975/4, pp. 7-10 (doeleinden: a) computer aided teaching of Greek and Latin, b) searching large collections of Greek and Latin texts, c) text editing and typesetting);
- Russo Frederica, *Utilisation de la mesure des longueurs de phrases pour la détermination de la structure de la prose de Fronton dans le „De eloquentia“ et le „De orationibus“*, in *Revue*, 1975/2, pp. 31-48 (als criterium voor de lengte van de zin werd het aantal lettergrepen tussen twee clausules genomen; na

Tot slot van deze uiteenzetting wou ik graag nog even stilstaan bij de vraag naar het nut van het frequentieonderzoek.

Bij allerlei problemen van de filologie is het verantwoord, ja soms noodzakelijk, aan frequentieonderzoek te doen. In het eerste deel van dit opstel werd erop gewezen hoe men in de hulpwetenschappen van de klassieke filologie alsook in alle mogelijke stadia van het filologisch onderzoek zelf, bij het verwerken van zijn variabelen zelfs een beroep gaat doen op de computer. Machinaal maar ook manueel uitgevoerde statistische onderzoeken dienen zeker de wetenschap. Een laatste mogelijke dienst aan de wetenschap vanwege het frequentieonderzoek werd hierboven nog niet met zoveel woorden behandeld, maar steunt wel op het zoeken naar de basisfenomenen van de taal zoals beschreven in het tweede deel van deze bijdrage. Ik bedoel het opstellen van een soort „basisvocabularium”, „basismorfologie” en „basissyntaxis” voor alle auteurs van de Latiniteit en van de Graeciteit. Alvorens dergelijk plan verwezenlijkt kan worden, zal er eerst nog veel water naar de zee moeten vloeien. Maar dergelijke realisatie zou zeker van wetenschappelijk belang zijn. Zo zouden taal en stijl van elke auteur bestudeerd kunnen worden in hun afwijkingen tegenover een „norm”. Zo zou de filoloog woordenschat, morfologie en syntaxis synchronisch en diachronisch kunnen bestuderen. Zo zou hij ook beter de eigenheden van elke tijd en van elk genre kunnen bepalen en ook de evolutie van het taalgebruik zelf gemakkelijker kunnen volgen, bij voorbeeld welke oude woorden er mettertijd uit de taal verdwijnen, welke archaische wendingen er weer opgerakeld en welke neologismen er ingevoerd worden.

Het nut van frequentieonderzoeken voor het onderwijs zal wellicht niet iedereen zonder meer willen erkennen. Misschien zullen sommigen zelfs vrezen dat een beperking van vocabularium en grammatica noodzakelijkerwijze een verschrompeling, ja een aftakeling van de goede oude klassieke studies moet betekenen. Ze moeten echter bedenken dat de suprematie van de Oude Talen in ons onderwijs definitief tot het verleden behoort en men zijn eisen moet matigen. Ik wil het frequentieonderzoek hier niet voorstellen als het middel om de klassieke studies te redden, maar zich niet

toepassing van dat criterium op enkele stalen uit Fronto's genoemde werken en uit Cicero's *Orator* laat de auteur zien hoe Fronto Cicero's breedheid van de periode verlaat ten gunste van de expressiviteit van het woord);

- Weil G. E., Salem A. & Serfaty M., *Le livre d'Isaïe et l'analyse des sources textuelles. Recherches sur des méthodes d'analyses statistiques descriptives (l'analyse des correspondances et la classification automatique)*, in *Revue*, 1976/2, pp. 1-85 (à suivre) (tekstanalyse van een omstreden werk op basis van de spreiding van de 100 frequentste woorden, de eigennamen uitgezonderd, over elk van de capita van dat werk; evaluatie van de gegevens van Kasher).

willen aanpassen aan de tijdsgeest, die toch naar rendement vraagt, zal zeker de vroegere „gouden tijd” niet herstellen. Als de klassieke studies in het Middelbaar Onderwijs behouden moeten blijven, dan zal dat zeker niet kunnen door de oude, deductieve methode, met haar overdreven aandacht voor niet-frequente woorden en zeldzame morfologie, te willen handhaven. Het kan paradoxaal klinken, maar eventuele tegenstanders van het benutten van frequentieonderzoeken in het onderwijs moeten er niet voor beducht zijn dat de voorschriften van de nieuwe leerplannen inzake frequentie pijnlijk nauwkeurig zullen worden toegepast. Dat zou trouwens ook moeilijk kunnen. De ontwerper van een handboek met leeslessen volgens de inductieve methode zal zich voor de moeilijkheid geplaatst zien dat hij met de meest frequente taalverschijnselen alléén geen zinnige Griekse teksten kan samenstellen. De woorden die het talrijkst voorkomen zijn immers de grammaticale (dat zijn lidwoord, voornaamwoorden, voorzetsels, voegwoorden, bijwoorden en partikels). In mijn excerpten uit de *Anabasis* vertegenwoordigen de grammaticale woorden ongeveer 53 % van het totale aantal items (zie tabel I). De zwakke positie van de betekenisdragende woorden (substantieven, adjectieven en werkwoorden) blijkt echter niet alleen uit hun kleiner aantal items (c. 47 %), maar vooral uit de massa vocabula die er nodig zijn om hun kleiner aantal items te dekken: het totale aantal lemmata bestaat uit c. 80 % betekenisdragende woorden en c. 20 % grammaticale. De verhouding tussen grammaticale en betekenisdragende woorden wordt misschien wel het best geïllustreerd door vergelijking van hun gemiddeld gebruik (het coëfficiënt van de verhouding tussen aantal items en aantal lemmata). Zo blijkt het gemiddeld gebruik van de grammaticale woorden meer dan viermaal zo hoog te liggen als dat van de betekenisdragende (18,69 tegen 3,92). Om zinnige, en wat meer is, interessante teksten te kunnen schrijven zal de ontwerper van een handboek ook minder frequente woorden moeten inlassen. In deze context kan nog even het probleem aangeraakt worden waarop de initiaticursus de leerling eigenlijk dient voor te breiden. Op alle schoolauteurs, op het gehele oeuvre van Xenofon, de gehele *Anabasis* of op excerpten daaruit die men het volgende jaar werkelijk wil lezen? In het ene uiterste geval kan men de leerling misschien te veel hooi op zijn vork geven, in het andere moet men hem op Xenofon voorbereiden met Xenofon. De samensteller van een begincursus naar de nieuwe voorschriften zal dus een grote soepelheid aan de dag moeten leggen. Soepelheid is trouwens ook vereist als men de mentaliteit van de jonge leerling wil respecteren. Het aanbrenge van de leerstof moet ook didactisch verantwoord zijn, volgens het principe van de geleidelijkheid

met anticipatie en herhaling. Men kan niet zonder meer eerst aanleren wat het meest frequent is en zijn leergang verder opbouwen in functie van de frequentie. Moeilijkheden zijn er dus genoeg. Maar toch moet het mogelijk zijn een initiatiecurcus te vrijwaren van de overtolligheden die vroegere leerboeken ontsierden.

Tot slot mag ik wel concluderen dat het frequentieonderzoek zijn nut heeft én voor de wetenschap én voor het onderwijs. Het is echter wèl zo, dat de toepassing van de zuivere wetenschap in het onderwijs door allerlei factoren kan worden bemoeilijkt.

TABEL I.

Verdeling van de woordenschat per grammaticale categorie bij Xenofoon, *Anabasis* I en IV (excerpta).

Categorie	Absolute waarden		%		Orde van belangrijkheid		Gemiddeld gebruik
	Voc- bula	Items	Voc- bula	Items	Voc- bula	Items	
Lidwoord	1	1200	0,06	10,76	10	4	1200,—
Substantieven	446	1596	26,98	14,32	2	3	3,58
Adjectieven	234	870	14,16	7,80	3	6	3,72
Voornaamwoorden	51	1113	3,09	9,98	5	5	21,82
Telwoorden	38	183	2,30	1,64	6	10	4,82
Werkwoorden	657	2777	39,75	24,91	1	1	4,23
Voorzetsels	35	745	2,12	6,68	8	7	21,29
Voegwoorden	37	390	2,24	3,50	7	9	10,54
Bijwoorden	114	577	6,90	5,18	4	8	5,06
Partikels	40	1697	2,42	15,22	9	2	42,43
Totalen	1653	11148	100,—	100,—			6,74
Betekenisdragende	1337	5243	80,88	47,03	1	2	3,92
Grammaticale	316	5905	19,12	52,97	2	1	18,69
Totalen	1653	11148	100,—	100,—			6,74

TABEL II.

Frequentie van *alla* en *ou*
(naar S. Michaelson & A. Q. Morton, *Elision...*).

Auteur, werk		Occurrenties		Procentuele verhouding	
		<i>alla</i>	<i>ou</i>	<i>alla</i>	<i>ou</i>
1.	Isokrates, rede I-XXI (alle)	503	901	36	64
2.	Lusias, rede I-XXXV (alle)	452	798	36	64
3.	Demosthenes, rede I-X	111	250	31	69
4.	Xenofon, 6 minora	111	118	48	52
5.	Aristoteles, <i>De Anima</i>	220	314	41	59
6.	Aristoteles, <i>De Caelo I-II</i>	40	37	52	48
7.	Aristoteles, <i>De Interpretatione</i>	66	188	26	74
8.	Homerus, <i>Ilias</i>	740	793	48	52
9.	Homerus, <i>Odusseia</i>	582	691	46	54
10.	Thoekudides, (selecta)	109	359	23	77
11.	Herodotos, (selecta)	83	254	25	75
12.	Ploetarchos, (selecta)	123	157	44	56

Legende

1. zie : S. Michaelson & A. Q. Morton, *Elision...* : p. 47 ; 2. p. 48 ; 3. p. 49 ; 4. p. 50 (*Hieroon, Agesilaos, Spartaanse instellingen, Poroi, Cavaleriecommandant, Rijkunst*) ; 5.-7. p. 54 ; 8.-9. p. 55 ; 10. p. 45 (eerste 200 zinnen van elk Boek) ; 11. p. 46 (eerste 200 zinnen van elk Boek) ; 12. p. 53 (*Isis en Osiris 351d-356f, E te Delfoi 384d-390d, Orakels te Delfoi 394e-401e, Oude mannen 783b-794h*).

TABEL III.

Gemiddelde frequentie van *alla* in stalen uit Griekse auteurs
(naar S. Michaelson & A. Q. Morton, *The spaces in between*, pp. 55-59).

	Auteur	Aantal <i>alla's</i>	Aantal tussenruimten	Gemiddelde fre- quentie op 100
1.	Herodotos	81	900	9,-
2.	Thoekudides	18	800	2,3
3.	Isokrates	376	1600	23,5
4.	Demosthenes	133	800	16,6
5.	Lusias	140	700	20,-
6.	Ploetarchos	59	500	11,8
7.	Xenofoon	40	600	6,7
8.	Iosepos	32	400	8,-
9.	Aristoteles	77	700	11,-
10.	Diodoros	12	400	3,-
11.	Filoon	31	400	7,8

Legende

1. in de som van de eerste 100 tussenruimten tussen 2 *kai's* van elk Boek ; 2. id. ; 3. id., ... van de redevoeringen II-XII, XIV en XVI-XIX ; 4. id., ... van de red. II-IV en VI-X ; 5. id., ... van de red. I-III, V, XII, XIV en XIX ; 6. in de som van de eerste tussenruimten tussen 2 *kai's* van *Ballingschap, Oude mannen en Goddelijke wraak*, en van de eerste 200 van *Fortuin van de Romeinen* ; 7. in eerste 100 van *Agésilao's, Spartaanse instellingen, Poroi, Cavaleriecommandant, Rijkunst, Jager* ; 8. id., ... van *Joodse oorlog I en V, Leven I en Tegen Apioon I* ; 9. in eerste 100 van *De Caelo I-IV en Delen van de dieren* en eerste 200 van *Voortbeweging van de dieren* ; 10. eerste 100 van Boeken I-II en XXI-XXII ; 11. id., ... van *Schepping, Allegorie, Cherubim en Kain en Abel*.

TABEL IV.

Verdeling van de werkwoordsvormen per wijs
bij Xenofoon, *Anabasis I & IV* (excerpta).

Wijs	Absolute waarden	%
Indicatief	1323	47,64
Conjunctief	61	2,20
Optatief	159	5,73
Imperatief	10	0,36
Infinitief	459	16,53
Participium	735	26,47
Verbaaladjectief	30	1,08
Totalen	2777	100,—

TABEL V.

Verdeling van de werkwoordsvormen per tijd
bij Xenofoon, *Anabasis* I & IV (excerpta).

Tijd	Absolute waarden	%
Praesens	1106	40,26
Imperfectum ind.	631	22,97
Futurum	85	3,09
Aorist	810	29,49
Perfectum	95	3,46
Plusquamperfectum ind.	20	0,73
Totalen	2747	100,—

TABEL VI.

Verdeling van de werkwoordsvormen per vox
bij Xenofoon, *Anabasis* I & IV (excerpta).

Vox	Absolute waarden	%
Actief	1895	68,98
Mediaal	619	22,53
Passief	233	8,48
Totalen	2747	100,—

TABEL VII.

Verdeling van de werkwoordsvormen per persoon
bij Xenofoon, *Anabasis* I & IV (excerpta).

Persoon	Absolute waarden	%
1 enk.	65	4,19
2 enk.	15	0,97
3 enk.	889	57,24
1 meerv.	28	1,80
2 meerv.	19	1,22
3 meerv.	537	34,58
Totalen	1553	100,—

TABEL VIII.

Verdeling van de werkwoordsvormen per stamsysteem
bij Xenofoon, *Anabasis* I & IV (excerpta).

Stamsysteem	Absolute waarden	%
I PRAESENS - s.	1743	63,31
II FUTURUM - s.	85	3,09
III EERSTE AORIST - s.	347	12,60
IV TWEDE AORIST - s.	388	14,09
V EERSTE PERFECTUM - s.	21	0,76
VI TWEDE PERFECTUM - s.	40	1,45
VII PERFECTUM MEDIAAL - s.	54	1,96
VIII EERSTE PASSIEF - s.	61	2,22
IX TWEDE PASSIEF - s.	14	0,51
Totalen	2753	100,—

BIBLIOGRAFIE

- Ahnert H.-J., *Affiliation programs (Part II). Using the program „AFFILI“*, in *Revue (de l'Organisation internationale pour l'étude des langues anciennes par ordinateur)*, 1972/4, pp. 35-54.
- Allais Christiane & Claude, *A method of structural analysis with an application to „Les liaisons dangereuses“*, in *Revue*, 1968/2, pp. 13-33.
- Bodson A., *Vocabulaire latin*, Liège, C.I.D.E.L.A., L.A.S.L.A., 1968.
- Bodson A. & Govaerts Suzanne, *Codification d'un texte latin sur cartes mécanographiques IBM-80 colonnes*, in *Revue*, 1966/1, pp. 1-50. (+ *Corrigenda*, in *Revue*, 1966/2, pp. 47-48).
- Brandwood L., *Plato's Seventh Letter*, in *Revue*, 1969/4, pp. 1-25.
- Carney T. F., *Looking for a writer's „Picture of reality“*, in *Revue*, 1968/2, pp. 35-81.
- Coppens-Ide Hilda, *Authorship problems and the computer*, in *Revue*, 1971/3, pp. 187-194.
- De Bie J., *Recherche sur la syntaxe et l'ordre des mots en grec*, in *Revue*, 1970/1, pp. 17-27.
- Delatte L., *A propos de la mesure des longueurs de phrases dans les textes classiques*, in *Revue*, 1966/1, pp. 51-65.
- Id., *Computers and the Classics*, in *Revue*, 1967/4, pp. 97-103.
- Id., *Key-words and poetic themes in Propertius and Tibullus*, in *Revue*, 1967/3, pp. 31-79.
- Id., *Le Laboratoire d'Analyse statistique des Langues anciennes à l'Université de Liège et l'établissement d'un vocabulaire de base et d'une grammaire de fréquence du latin*, in *Colloquium voor de studie van de problemen van het onderwijs in het Grieks en het Latijn in de landen van de Europese Gemeenschap (Gent 17-21 dec. 1963). Verslag en Mededelingen*, Gent, 1964, pp. 96-106.
- Id., *L'ordinateur et la création littéraire. Communication faite au Colloque International de l'Esthétique et des Sciences connexes, Liège, octobre 1966*, in *Revue*, 1966/3, pp. 5-20.
- Denooz J., *Recherches sur le traitement automatique de la langue latine*, in *Revue*, 1973/1, pp. 1-89.
- De Poerck G. & Zwaenepoel R., *Traitements automatiques de textes littéraires en ancien français. Une expérience: le Charroi de Nîmes*, in *Revue*, 1971/2, pp. 97-111 (pp. 97-106 en 109-110: De Poerck, pp. 107-108: Zwaenepoel).
- Dyer R. R., *Towards computational procedures in homeric scholarship*, in *Revue*, 1967/4, pp. 1-54.
- Engels L., *The computer and the humanities. Summary of the Colloquium*, in *Revue*, 1971/3, pp. 249-251.
- Evrard E., *Deux programmes d'ordinateur pour l'étude quantitative du vocabulaire*, in *Revue*, 1967/3, pp. 81-95.
- Id., *La critique des textes et son automatisation: à propos d'un livre récent*, in *Revue*, 1968/1, pp. 69-81.
- Id., *La papyrologie au L.A.S.L.A.: réalisations et projets*, in *Revue*, 1967/4, pp. 87-96.

- Id., *L'automatisation de l'étude des textes grecs au LASLA*, in *Revue*, 1971/2, pp. 89-95.
- Id., *Le Laboratoire d'Analyse statistique des Langues anciennes à l'Université de Liège. L'établissement d'un vocabulaire de base du grec*, in *Colloquium...*, pp. 107-114. (Cfr. Delatte, *Le Laboratoire*).
- Id., *Scansion automatique de l'hexamètre grec*, in *Revue*, 1972/4, pp. 1-33.
- Geiss H., *Einige Probleme der Mykenologie und ihre Lösungsmöglichkeiten mit Hilfe von datenverarbeitenden Maschinen*, in *Revue*, 1968/3, pp. 19-33.
- Germain Agnès, *Un programme de comptage et de distribution automatique des fréquences des mots par rôle dans les tragédies de Sénèque*, in *Revue*, 1974/3, pp. 33-48.
- Goerlandt E., *Een frequentieonderzoek op Xenofoons Anabasis*, in *Kleio*, 3 (1973-1974), pp. 145-164.
- Id., *Frequentieonderzoek en de studie van het Grieks*, in *DiCG*, 14 (1974), pp. 5-64.
- Id., *Lexique de base grec*. Traduit par S. Govaerts & J. Denooz, Liège, C.R.I.D.E.L.A., L.A.S.L.A., 1973.
- Id., *Resultaten van een frequentieonderzoek op de Anabasis van Xenofoon*, Brussel, Ministerie van Nationale Opvoeding en Nederlandse Cultuur, Bestuur Hoger Onderwijs en Wetenschappelijk Onderzoek, Afdeling bijscholing en pedagogisch onderzoek, [1976] (= *Documentatie 101, Grieks*).
- Govaerts Suzanne, *Le Corpus Tibullianum. Index verborum et Relevés statistiques. Essai de méthodologie statistique*, La Haye, Mouton, 1966, XIII-335 p. (= *Travaux publiés par le Laboratoire d'Analyse statistique des Langues anciennes*, Université de Liège, Faculté de Philosophie et Lettres, Fasc. 5).
- Id., *L'étude des textes latins au LASLA*, in *Revue*, 1971/2, pp. 97-102.
- Id., *Quelques observations à propos de la courbe cumulative de la distribution du vocabulaire chez un poète latin: Tibulle*, in *Revue*, 1973/4, pp. 19-27.
- Govaerts Suzanne & Denooz J., *Lexique de base latin*, Liège, C.R.I.D.E.L.A., L.A.S.L.A., 1972.
- Id., *Présentation d'un vocabulaire de base latin publié par le Laboratoire d'Analyse statistique des Langues anciennes de l'Université de Liège*, in *Revue*, 1974/1, pp. 17-29 (+ bijlagen).
- Greenberg N. A., *Vergil and the computer. Forth foot texture in AENEID I*, in *Revue*, 1967/1, pp. 1-16.
- Herrmann E., *Das Praesens historicum in Xenophons Anabasis*, in *Nachrichten von der Akademie der Wissenschaften in Göttingen, Philologisch-historische Klasse*, Jahrgang 1944, n° 8, pp. 183-214.
- Holtermann H., *Ianua Linguae Graecae. I. Lees- en Oefenboek, II. Woordverklaring*, in de bewerking van dr. Em. Vandevelde, Leuven, s.d.
- Janon M., *Epigraphie et informatique*, in *Revue*, 1974/2, pp. 1-24.
- Jory E. J. & Moore D. W., *An index to CIL VI. Two aspects*, in *Revue*, 1966/2, pp. 7-16.

- Kasher A., *The Book of Isaiah: characterization of authors by morphological data processing*, in *Revue*, 1972/3, pp. 1-62.
- Kollmann E. D., *A study of the vocabulary of Vergil's Eclogues*, in *Revue*, 1973/3, pp. 1-24.
- Id., *Word frequencies in latin literature*, in *Revue*, 1973/4, pp. 1-17.
- Maier F., *Stilübungen und Interpretation im Griechischen. Zum Problem des Übersetzens an Universität und Gymnasium*, München, Max Hueber Verlag, 1967. (= *Sprachen der Welt*, Nr. 6025).
- Maloney G., *La fréquence et l'ordre des formes verbales dans l'œuvre de Thucydide*, in *Revue*, 1970/3, pp. 87-109.
- Mater E., *Die Verwendung von datenverarbeitenden Maschinen bei der Untersuchung von Linear-B-Inschriften*, in *Revue*, 1968/3, pp. 1-17.
- Mau J., *Affiliation programs*, in *Revue*, 1972/3, pp. 63-76.
- Meyer Th. & Steinthal H., *Grund- und Aufbauwortschatz Griechisch*, Stuttgart, 1973. (= *Klettbuch* 6632).
- Michaelson S. & Morton A. Q., *Elision as an indicator of authorship in Greek writers*, in *Revue*, 1973/3, pp. 33-56.
- Id., *The spaces in between. A multiple test of authorship for Greek writers*, in *Revue*, 1972/1, pp. 23-77.
- Morgan R., *The computer and the maya glyphs*, in *Revue*, 1969/1, pp. 29-38.
- Müller L., *L'ordinateur et les textes de droit romain*, in *Revue*, 1970/4, pp. 65-82.
- Najock D., *Automatic classification of texts by methods of multivariate statistics*, in *Revue*, 1973/2, pp. 31-54.
- Id., *Orientation of text-stems*, in *Revue*, 1972/2, pp. 39-56.
- Ott W., *Metrical analysis of latin hexameter by computer*, in *Revue*, 1966/4, pp. 7-24.
- Ouy G., *Quelques problèmes posés par le projet de code descriptif pour les manuscrits médiévaux („Questionnaire codicologique“)*, in *Revue*, 1970/3, pp. 19-38.
- Paquot Annette, *Détails sur une nouvelle entreprise. Index Verborum inverse de Plaute avec relevés statistiques*, in *Revue*, 1973/2, pp. 35-47.
- Pennacini A., *Le fragment comme échantillon*, in *Revue*, 1968/1, pp. 1-9.
- Radday Y. T. & Shor H., *And in Isaiah*, in *Revue*, 1974/2, pp. 25-41.
- Rigo G., *Ajustement d'une courbe de distribution du vocabulaire*, in *Revue*, 1972/3, pp. 77-91.
- Id., *Projet d'étude statistique de l'Electre d'Euripide*, in *Revue*, 1971/4, pp. 45-55.
- Smyth H. W., *Greek Grammar*, Harvard University Press, Cambridge, Massachusetts, 1968⁵.
- Tomsin A., *Etablissement d'un programme d'analyse automatique des textes papyrologiques documentaires grecs*, in *Revue*, 1970/1, pp. 57-69.
- Tordeur P., *Essai d'analyse statistique de la métrique de Prudence*, in *Revue*, 1972/2, pp. 19-37.
- Zarri G. P., *Une expérience pour l'automation des recherches de papyrologie*, in *Revue*, 1967/4, pp. 55-86.