

INTRODUCTION

1. – C'est en conclusion du Colloque « Les machines dans la linguistique », qui réunit à Prague en Avril 1966 les représentants de presque tous les Centres Européens s'occupant alors du domaine de la lexicologie et de la lexicographie assistées par ordinateurs, que fut formulée la proposition, accueillie à l'unanimité, d'organiser le prochain Congrès International en Italie sous le patronage du Centre National Universitaire de Calcul Electronique (CNUCE) de l'Université de Pise et de l'Académie de la Crusca à Florence. Cette proposition émanait de l'essor qu'à cette époque la collaboration étroite des deux Instituts avait apporté aux recherches dans ce domaine pour lequel l'Italie vantait déjà l'oeuvre de pionnier du P. R. Busa S. J. qui en 1951 avait publié les premiers index et concordances obtenus électroniquement.

En 1968, vu le nombre croissant des projets de traitement automatique des données linguistiques et littéraires, la Direction du CNUCE décidait d'instituer, dans le cadre du CNUCE, une section de linguistique ayant pour tâche de promouvoir et d'effectuer des recherches dans le domaine de la linguistique mathématique et computationnelle, d'assurer les relations scientifiques et l'échange d'informations entre les chercheurs italiens et étrangers, d'organiser et de coordonner l'activité didactique pour diffuser les expériences et les méthodologies et enfin de prêter assistance scientifique et technique aux Instituts intéressés. A la différence de l'assistance fournie à d'autres catégories d'utilisateurs, qui se limite en général à assurer les instruments et le temps de calcul, l'assistance réservée aux linguistes et aux humanistes comprenait également la consultation et la collaboration pour la planification scientifique et technique des projets, l'analyse et la mise au point des programmes nécessaires et l'exécution des élaborations. Cette décision a contribué de façon déterminante au développement qualitatif et quantitatif des recherches et des applications en Italie, illustré au moins en partie, dans ce volume, par quelques articles. Les savants italiens ont maintes fois

exprimé leur reconnaissance à la Direction du CNUCE, en particulier au Directeur A. Faedo et au Secrétaire Général G. Torrigiani.

Pour remplir les tâches assignées à la Section dans le domaine de la didactique et des relations internationales, il sembla opportun d'organiser une Ecole Internationale d'Eté et de la compléter en lui adjoignant le Colloque International proposé à Prague en 1966.

Pour établir ce programme les suggestions de nombreuses personnes m'ont été de grande utilité, mais il m'est difficile de les citer toutes ici sans commettre d'omission. La première suggestion m'a été faite par quelques visiteurs américains parmi lesquels J. Raben et R. Dyer qui avaient contribué à l'organisation de l'*American Philological Association Summer Institute on Computer Applications to Classical Studies*. Le programme fut discuté avec de nombreux participants de la 1969 *International Conference on Computational Linguistics* à Stockholm dont l'approbation constitua la première preuve de validité. Les facteurs déterminants qui nous décidèrent à le mettre à exécution furent l'assurance de la collaboration de l'Académie de la Crusca, avec le conseil éminent de G. Devoto, C. A. Mastrelli et G. Nencioni, l'approbation et le précieux guide des linguistes de l'Institut de glottologie de l'Université de Pise et en particulier de son directeur T. Bolelli, qui avait justement patroné en 1970 le commencement d'un cours officiel de *Linguistique Mathématique et Computationnelle*, le premier en Italie.

2. – Le Colloque International « l'Elaboration Electronique en Lexicologie et Lexicographie » a eu lieu à Pise les 24 et 25 Août 1970. Plus de 200 savants italiens et étrangers y ont pris part. Les travaux ont été ouverts par A. Faedo, Recteur de l'Université de Pise et Directeur du CNUCE, dont le discours est transcrit à la page VII. Après le salut du Maire de Pise, G. Devoto se chargea de la présidence des travaux. Le 25 Août la présidence fut assumée par C. A. Mastrelli et P. E. Arias. Le soir du 24 Août l'Université de Pise offrit une réception en honneur des participants.

3. – L'Ecole Internationale d'Eté, « l'Elaboration Electronique des données linguistiques et littéraires », a été ouverte le 17 Août 1970 par G. Torrigiani, Secrétaire Général du CNUCE, qui a souhaité la bienvenue aux participants en son nom et en celui de A. Faedo, en présentant les activités du CNUCE depuis son origine et en mettant particulièrement en relief l'engagement contracté par la Direction pour promouvoir et favoriser les études de linguistique automatique. Le nombre et la re-

présentativité scientifique des participants témoigne de l'accueil favorable réservé au CNUCE par le monde de la linguistique automatique de tous les pays, a conclu G. Torrigiani, qui a renouvelé ses remerciements et ses vœux de fécond travail aux participants de l'Ecole.

L'Ecole d'Eté fut conclue par G. Devoto et A. Faedo le 5 septembre au cours d'une réception offerte par le Maire de Pise aux participants et au Comité scientifique de l'Ecole. Les travaux se sont tenus tous les jours de 8h. 30 à 12h. 30 et de 15h. à 19h. 30, excepté le dimanche 23 Août dédié à la visite de la ville de Pise et le dimanche 3 Septembre consacré à la visite des antiquités étrusques et romaines de Volterra dont le Maire a reçu les participants. Les 24 et 25 Août les travaux de l'Ecole ont été fusionnés avec ceux du Colloque International auquel ont participé tous les inscrits à l'Ecole, à titre de conférenciers pour quelques-uns.

Il y eut 150 demandes d'inscription dont la répartition selon la nationalité est la suivante: Italie 95, France 15, Allemagne 7, Belgique 7, Angleterre 7, Hollande 4, Suède 2, USA 2, Roumanie 2, Danemark 1, Espagne 2, Yougoslavie 2, Suisse 2, Pologne 1, Finlande 1. La répartition selon la qualification professionnelle comprenait: 21 professeurs titulaires ou chargés de cours, 29 assistants d'Université, 31 chercheurs, 25 étudiants au niveau du 3ème cycle, 29 étudiants, 15 employés de l'industrie et de l'administration. Une enquête sur les recherches et les activités a donné les résultats suivants: Glottologie et Linguistique Générale 16, Information Retrieval 13, Littérature Italienne 13, Psycholinguistique, Psychologie, Pédagogie, Psychiatrie 13, Recherches de software et Physique 12, Langue et Littérature Anglaise 12, Linguistique Mathématique et Computationnelle 11, Langue Classique et Archéologie 11, autres Philologies et Langues 8 (allemande 4, slave 2, orientale 2), Enseignants de lycée 8, Démologie et Dialectologie 6, Philosophie 4, autres 8. 101 inscrits ont été admis. Les heures de cours ont été de 87, celles des séminaires et des conférence de 40.

4. – L'Ecole fut annoncée au mois de Juin 1970. L'importance numérique des demandes d'inscription parvenues malgré le bref temps à disposition, au lieu des 30 ou 40 prévues à l'origine, semble indiquer que l'initiative a répondu à une exigence largement diffusée en Europe parmi les humanistes.

Les buts de l'Ecole avaient été ainsi énoncés dans le programme: « *la computational linguistics* est aujourd'hui considérée non seulement comme une sous-spécialisation professionnelle, mais plutôt comme un

ensemble de méthodologies et de techniques utilisables par tous les linguistes et par tous ceux qui pour leurs études doivent faire des recherches et des opérations sur des textes en langue naturelle. Les élaborations des données linguistiques intéressent donc aujourd'hui beaucoup de secteurs variés comme la lexicographie, la stylistique, la critique de texte, la statistique linguistique, la formalisation en linguistique, l'analyse documentaire des textes, la pédagogie des langues, l'instruction automatique, la gestion des archives de langue et de dialectes, l'analyse psychologique et psychiatrique du langage, etc...

Les Centres et les Instituts spécialisés, surtout ceux qui ont été récemment constitués dans de nombreux pays pour la lexicographie, sont en train de préparer des « bibliothèque électroniques », toujours plus complètes, de textes enregistrés et analysés sur des supports utilisables par l'ordinateur. Il est sans cesse plus facile qu'un chercheur trouve déjà enregistrés électroniquement les textes qu'il veut soumettre à une étude linguistique, littéraire, historique etc... au moins en ce qui concerne les principales langues de culture. De là dérive une double exigence: d'un côté les chercheurs doivent être informés aussi bien des projets en cours que des méthodes et des techniques d'utilisation des *data base*, de l'autre il est nécessaire de faire un effort commun pour unifier les méthodologies sur le plan international (comme la Section de Linguistique du CNUCE est déjà en train de le faire sur le plan national) afin que les textes, qui peu à peu s'ajoutent à de telles bibliothèques soient à la disposition des divers chercheurs et soient élaborables avec des programmes normalisés. De nombreuses organisations qui s'occupent de l'élaboration des informations à une large échelle, dans les milieux administratifs, industriels, scientifiques etc... se tournent vers l'automation et tendent à construire des systèmes pour analyser et élaborer le contenu des documents en langue naturelle. Des applications comme *l'information retrieval*, des systèmes *fact-retrieval* ou *question-answering*, la *computer assisted instruction*, la *traduction mécanique* ont intérêt à l'élaboration automatique du langage à cause de sa fonction de véhicule primaire des informations. Les recherches et les applications ont en commun l'utilisation de quelques techniques computationnelles de base pour l'élaboration automatique des données non numériques et l'exigence d'introduire un certain degré de formalisation dans les données linguistiques et dans leur élaboration. L'Ecole, qui s'adresse principalement à des savants, des chercheurs et des étudiants de formation littéraire, se propose de porter les participants à la connaissance et à l'application immédiate des méthodologies et des techniques en usage,

et d'en illustrer les possibilités à travers l'expérience et les résultats obtenus par les savants qui travaillent depuis longtemps dans ces divers domaines. »

L'association de l'Ecole et du Congrès spécialisé, dans le cadre duquel les participants de l'Ecole ont pu assister et souvent prendre part à la discussion et à la confrontation des méthodes et des idées entre les chercheurs provenant de divers pays et représentant des courants et des écoles différentes, s'est révélée de grande utilité. Grâce à cette association, cette première Ecole a rempli non seulement une fonction propédeutique et formatrice, mais elle a répondu également à une autre exigence fondamentale de ce domaine en rapide développement et en continuelle expansion: l'échange des informations, la standardisation des méthodes, des programmes et des matériaux, la mise à jour continue et la collaboration interdisciplinaire. Ces activités, comme on l'a déjà dit, font partie des fins institutionnelles de la Section Linguistique du CNUCE.

5. – Ce volume contient les textes des communications présentées au Colloque et les textes des conférences et des leçons tenues à l'Ecole. Il a été décidé de n'établir aucune distinction entre elles, soit à cause de la réelle efficacité didactique et informatrice conférée par l'étroit lien qui unit les deux manifestations, soit afin de regrouper les textes selon les sujets traités. Le volume comprend les textes de toutes les conférences, à l'exception de celle de C. Tagliavini dont le texte a été inclus dans le volume U. Bortolini, C. Tagliavini, A. Zampolli: *Lessico di Frequenza della Lingua Italiana Contemporanea* (1971). En ce qui concerne les textes des cours, ceux de D.G. Hays, M. Gross, A. Zampolli n'ont pas été inclus, car ils font l'objet de publications de ces divers professeurs.¹ M. Gross a cependant envoyé un article qui développe l'un des sujets qu'il a traité.

6. – Sur le conseil du Comité Scientifique et d'après les résultats scientifiques et didactiques obtenus, l'on a décidé de rendre stable l'Ecole Internationale d'Eté. Le Comité Scientifique trouve opportun qu'elle se déroule tous les deux ans en alternant les programmes essen-

¹ D. G. HAYS, *Introduction to Computational Linguistics*, New York, 1967. M. GROSS, *Mathematical Models in Linguistics*, Englewood Cliffs, N. J. V., 1972. A. ZAMPOLLI, *Fondamenti di Linguistica Computazionale*, Pisa, 1970.

tiellement introductifs, de caractère général, et les programmes traitant des sujets spécifiques. L'Ecole d'Eté de 1970 a voulu simplement offrir aux chercheurs et aux étudiants, en particulier italiens, un premier panorama d'information sur les aspects principaux et sur les possibilités de l'élaboration électronique de données linguistiques et littéraires, et elle leur a fourni les instruments de recherche et la méthodologie.

Une vingtaine de projets nouveaux ont été proposés en Italie par les participants de l'Ecole, et d'autres nous ont été signalés en France en Angleterre, en Allemagne, en Espagne et en Belgique.

Je pense que l'Ecole d'Eté de 1972 pourra s'articuler selon les deux arguments que les participants ont indiqué à la fin de l'Ecole de 1970: *grammaires formelles des langues naturelles et traitement automatique*: étude formelle de la syntaxe, représentation des propriétés syntaxiques d'un lexique, relations entre lexique, syntaxe et sémantique, méthodes d'analyse syntaxique automatique et de vérification automatique des règles; *applications de l'informatique aux analyses lexicales et aux élaborations lexicographiques*: problèmes et méthodes de la lexicologie et de la lexicographie contemporaines assistées par ordinateur, archives lexicales et banque des mots, rédaction lexicographique assistée par ordinateur, technologie avancée pour l'usage interactif de l'ordinateur, statistique lexicale.

S'il m'est permis de jeter un regard sur la 3ème édition de l'Ecole qui aura lieu en 1974, je formule l'hypothèse que l'on s'occupera davantage de sémantique. L'intérêt renouvelé pour les recherches sémantiques dans de nombreuses écoles linguistiques, joint aux résultats de la logique et de la sémantique formelle, trouvera certainement un terrain développement et d'application dans le secteur du traitement automatique en sémantique, qui est aujourd'hui le moins développé et le plus limitatif de la « computational linguistics ».

7. – En ma qualité de Directeur de l'Ecole et de Coordonnateur du Colloque, je désire remercier la Direction du CNUCE qui en est l'initiatrice et les a patronné. Je remercie tout particulièrement A. Faedo et G. Torrigiani qui ont démontré une exacte compréhension des exigences et des perspectives de nos disciplines; G. Devoto qui de ses suggestions et ses conseils nous a guidé pour la préparation du programme scientifique et sa réalisation et qui, en sa qualité de Président de l'Académie Toscane des Sciences et des Lettres « La Colombaria », a promu la publication de ces Actes; l'Institut de Glottologie de l'Université de Pise et en particulier son Directeur T. Bolelli pour l'assistance émi-

nente qu'il a bien voulu nous accorder; tous les membres du Comité Scientifique pour leur précieuse consultation et leur patronage scientifique; l'Académie de la Crusca pour sa continuelle et indispensable collaboration; la Société de Linguistique Italienne et son Président T. De Mauro pour l'aide apportée à l'organisation de la participation des chercheurs italiens à l'Ecole; l'Académie de la Colombaria et son secrétaire F. Adorno qui ont patronné la publication de ces Actes et l'Editeur Leo S. Olschki qui a particulièrement soigné leur publication.

A. ZAMPOLLI

Pise, Janvier 1971.