

**L'Institut d'Etudes et de Recherches  
pour l'Arabisation espère avoir  
terminé avant la fin de l'année 84  
le stockage de près  
de 500.000 données**

L'Institut d'Etudes et de Recherches pour l'Arabisation (IERA), dirigé à Rabat par le professeur Lakhdar Ghazal, a mis au point un codage de l'arabe vocalisé et un matériel de saisie de données bi-alphabétiques (arabe/latin) qui permettront la création de la première banque de données arabe.

Le procédé du Pr Lakhdar a fait l'objet d'un exposé et de démonstrations à l'exposition internationale de linguistique (Expo-Langues) qui s'est tenue à Paris du 28 janvier au 1er février 1983.

L'élaboration du codage, indispensable au traitement par ordinateur des données linguistiques, a nécessité une réforme de la graphie, notamment en raison de l'inadaptation de l'alphabet arabe aux techniques standard de diffusion de l'informatique imprimée.

Cette réforme a donné lieu dans un premier temps à la réalisation d'une nouvelle machine à écrire, d'une casse typographique et d'une police standard de 108 signes.

Grâce à cette standardisation du caractère arabe et aux progrès réalisés dans le domaine de l'électronique et de l'informatique, l'IERA a réussi à développer un logiciel et des terminaux de saisie bi-alphabétiques qui ont été présentés à Paris.

L'IERA a utilisé un logiciel mis au point par l'Agence Spatiale Européenne et qui a été développé et adapté aux exigences linguistiques arabes. Ce travail a permis la mise au point du «Quest Lexar», premier logiciel documentaire capable de

gérer des données bi-alphabétiques dont l'arabe est entièrement vocalisé.

Grâce à ce logiciel, aux terminaux «EURAB» (abrégé de Euro-Arabe) qui lui sont adaptés et à son système de codage unifié, l'IERA a pu commencer l'indexation et la saisie de données se rapportant aux domaines des sciences fondamentales (mathématiques, physique, chimie...), des sciences naturelles (médecine, botanique, géologie...) et des sciences humaines (droit, économie, finance...). Plusieurs fichiers lexico-graphiques ont pu ainsi être constitués.

L'IERA espère avoir terminé avant la fin de l'année prochaine le stockage de près de 500.000 données qui constitueront la première banque de données représentative de l'état actuel des terminologies arabes.

L'Institut s'est défini comme objectifs à moyen terme de dégager les termes «unifiés» sur la base de la plus grande fréquence d'emploi, remplir les lacunes lexicales (absence d'équivalents arabes satisfaisants des termes techniques européens), dégager des formants terminologiques (préfixes, infixes, suffixes) et l'étude de la traduction assistée par ordinateur.

A plus long terme, l'IERA projette la création d'un «trésor» (lexique) de la langue arabe, des banques de données spécifiques et la traduction de banques de données déjà existantes.