

UN MODULE DE TRAITEMENT DU TEXTE ECRIT EN FRANCAIS EN VUE  
DE LA SYNTHÈSE AUTOMATIQUE PAR DIPHONÈME

Daniel Teil et Bernard Prouts, LIMSI - CNRS 91406 Orsay, France

Ce module, destiné à remplacer celui implanté dans notre système de synthèse (Teil et al., 1972), détermine à partir d'un texte écrit en français orthographié les éléments nécessaires à sa synthèse par diphonèmes: suite des phonèmes, valeurs du pitch, valeurs temporelles.

Le programme orthoépique se résume à l'exploitation d'un lexique contenant indifféremment les règles de prononciation et les exceptions. Dans ce lexique sont inclus des marques de liaison et des marques prosodiques. Le traitement des homographes qui ne peut être fait que par analyse syntaxique et sémantique n'est pas traité actuellement.

L'interdépendance de la prosodie avec la syntaxe n'étant pas encore clairement établie nous avons choisi une méthode lexicale de découpage du texte en groupes prosodiques. Les ponctuations déterminent les coupes fortes; les coupes faibles séparent des groupements de mots établis en fonction d'une liste de mots "outils" (Choppy et al., 1975).

La courbe mélodique est calculée à partir d'un schéma en dent de scie appliqué au niveau du groupe mélodique et au niveau de la phrase en fonction de la ponctuation. Le rythme est actuellement traité de façon assez sommaire.

Conclusion

Les résultats obtenus sont assez encourageants et nous incitent à parfaire l'algorithme surtout au niveau de la prosodie.

Références

- C. Choppy, J.S. Liénard et D. Teil (1975): "Un algorithme de prosodie automatique sans analyse syntaxique", 6èmes Journées d'Etude sur la Parole, Toulouse.
- D. Teil, J. Sapaly et J.S. Liénard (1972): "Conception et réalisation d'un terminal vocal d'ordinateur à vocabulaire illimité et réponse immédiate", 3rd International Congress on Data Processing in Europe, Salzbourg.