

J.-P. HATON

Le traitement automatique de la parole est un domaine de recherche pour lequel un effort important a été consenti au cours des trois dernières décennies. Les problèmes à résoudre sont considérables et de nature fondamentale. Ils sont de plus par essence pluridisciplinaires : traitement du signal, reconnaissance des formes, intelligence artificielle, informatique, phonétique, linguistique, ergonomie, neurosciences interviennent à des degrés divers dans les solutions apportées.

Ces travaux de longue haleine donnent néanmoins naissance dès à présent à des produits intermédiaires qui trouvent leur place dans des applications pratiques dans le cadre de la communication homme-machine. C'est le cas aussi bien en synthèse de la parole (processus de production vocale), qu'en reconnaissance (perception vocale) ou en vérification de locuteur (identification de la personne qui parle).

Ce numéro spécial est consacré exclusivement à la reconnaissance automatique de la parole. Ce domaine donne lieu à des projets de grande envergure un peu partout dans le monde (Europe, Japon, USA) tandis que des produits de bonne qualité sont déjà proposés ou sur le point d'apparaître. Le moment est donc bien choisi pour faire le point sur la question et il faut remercier Bruno Georgel qui a eu l'idée de consacrer un numéro spécial de la revue TS à ce thème. Ce numéro comprend neuf articles qui abordent les aspects les plus importants du problème, tant sur le plan de la recherche que des applications et qui illustrent bien sa diversité. Tous ces articles sont issus de laboratoires participant au GRECO-PRC Communication Homme-Machine (pôle Parole), structure de coordination mise en place par le CNRS et le Ministère de la Recherche et de la Technologie.

L'article de J. Mariani constitue une bonne introduction, complète et documentée, au problème, à ses techniques et à ses applications.

La reconnaissance de mots, prononcés isolément ou de façon enchaînée, constitue la quasi-totalité des applications existantes, dans la mesure où il s'agit d'une simplification du problème général de la reconnaissance de la parole naturelle prononcée continûment. La plupart des systèmes actuels sont monolocuteurs et nécessitent un apprentissage préalable pour tout nouveau locuteur. Néanmoins, de bonnes performances sont maintenant possibles en

multilocuteurs (cf. C. Gagnoulet - D. Jovet et J. C. Junqua) et en adaptation automatique au locuteur (cf. J. P. Tubach et coll.). Les systèmes de reconnaissance de mots sont pour l'essentiel fondés sur des modèles formels qui ont été progressivement optimisés : programmation dynamique (cf. J. C. Junqua, qui utilise par ailleurs un ensemble de connaissances phonétiques), modèles stochastiques de Markov, les plus performants actuellement (cf. C. Gagnoulet - D. Jovet), réseaux neuronaux, plus récents mais prometteurs (cf. l'article de J. P. Haton et coll. où un modèle neuronal est utilisé pour un autre problème, le décodage acoustico-phonétique).

La reconnaissance de phrases en parole continue est un problème encore plus complexe que la reconnaissance de mots et fait nécessairement appel à des connaissances (phonétiques, linguistiques, etc.) relatives à la communication parlée. Une étape importante est constituée par le décodage acoustico-phonétique. L'article de J. P. Haton et coll. illustre différentes approches de ce problème et les questions qui se posent. La segmentation du semi-continuum vocal en unités de différents niveaux (mots, syllabes, phonèmes, etc.) est une opération délicate, indispensable dans tout système de reconnaissance de parole continue. M. Rossi présente ce problème et ses différentes implications.

Un système de reconnaissance automatique de la parole doit être intégré dans un système plus général d'interaction homme-machine, éventuellement multimédia, et comportant une véritable composante de dialogue. De nouveaux problèmes se posent pour réussir une telle intégration ; ils sont présentés dans l'article de J. M. Pierrel et coll. La plupart des travaux relatifs au dialogue sont à long terme. Néanmoins des applications existent déjà : elles peuvent être simples comme le système Mairievox de C. Gagnoulet - D. Jovet, ou plus complexes comme les systèmes DIRA décrit dans l'article de J. Caellen et coll. et DIAPASON décrit dans l'article de E. Gallais et coll.

Pour terminer cette introduction, je tiens à remercier J. Malbos pour son amabilité et son efficacité ainsi que tous les lecteurs qui ont aidé à améliorer la qualité des articles : P. Alinat, G. Bailly, C. Bourjot, J. Caelen, R. Carré, M. Eskenazi, J. Guizol, M. Guyomard, J. C. Junqua, J. Mariani, H. Maynard, H. Meloni, G. Mercier, M. A. Morel, F. Néel, G. Perennou, J. M. Pierrel, J. Siroux.