

Montréal – Canada



XVII^{èmes}
Journées
d'études ^{sur} la
PAROLE

Actes



Université de Montréal
Faculté des arts et des sciences
Département de linguistique et philologie

28-31 mai 1990

MARS : Un Système de Reconnaissance de l'Arabe Moderne

M. Djoudi

D. Fohr

J. P. Haton

CRIN — INRIA Lorraine
 Campus Scientifique
 B.P. 239
 54506 Vandœuvre-lès-Nancy CEDEX
 France

Résumé

Nous proposons dans cet article une architecture d'un système de reconnaissance automatique de l'Arabe moderne. Le système a pour but la compréhension de phrases dans un contexte multilocuteur. Sa mise en œuvre tient compte bien sûr des particularités de la langue. Nous allons décrire dans ce qui suit le module de décodage phonétique qui n'est autre que le système SAPHA développé pour le décodage acoustico-phonétique de l'Arabe [7]. Ensuite, nous aborderons le décodeur linguistique en évoquant les aspects de la morphologie de la langue, de la transcription phonétique-orthographique et de la structure syntaxico-sémantique des phrases. Enfin, nous parlerons de l'importance de la prosodie dans un système de compréhension de l'Arabe parlé et les points qu'il faut résoudre.

1 Introduction

L'arabe moderne, dite aussi littéraire ou standard a fait l'objet de plusieurs travaux anciens ou récents, ayant trait soit à l'aspect phonétique [1] [2] [10] [5] soit à la composante linguistique de la langue [15] [12] [14] [9]. Toutefois, le problème de reconnaissance automatique n'a été que très peu abordé jusqu'à présent [6]. Nous allons décrire dans cet article le système MARS (Modern Arabic Recognition System) qui se veut un système de reconnaissance de phrases en parole continue et dans un contexte multilocuteur. Le système reçoit en entrée une phrase et rend en sortie une (ou éventuellement plusieurs) interprétation sémantique de la phrase prononcée en utilisant des sources de connaissances diverses: phonétiques, phonologiques, morphologiques, syntaxiques, sémantiques et prosodiques.

2 Architecture du système

La structure générale du système fait apparaître deux grands sous systèmes. D'une part, un système de décodage acoustico-phonétique, en l'occurrence SAPHA et d'autre part, le décodeur linguistique SALAM. Chaque sous système utilise diverses sources de connaissances. Le décodeur phonétique fournit au système linguistique, en mode proposition, un treillis phonétique. Il peut être réactivé, en mode vérification, par le décodeur linguistique pour savoir si un phonème donné est l'étiquette d'un morceau de signal donné.

3 Le décodeur phonétique : SAPHA

Le système SAPHA [7] est composé d'un ensemble de modules et englobe les trois étapes : acoustique, phonétique et phonologique que comporte le niveau inférieur d'un système de reconnaissance automatique de la parole [11]. Ces modules sont :

3.1 Le module acquisition

C'est la première étape dans tout processus de reconnaissance de la parole. A ce niveau, l'acquisition de la parole est faite à partir d'un microphone ou d'une cassette. De même, on peut écouter un morceau de parole et jouer sur la valeur de la fréquence d'échantillonnage. Les fonctions de lecture, de sauvegarde, de coupure et d'affichage du signal sur une console graphique sont aussi prévues à ce niveau [13].

3.2 Le module acoustique

Ce module se charge d'extraire les paramètres acoustiques à partir du signal temporel, il s'agit en particulier de :

- L'énergie du signal.
- La densité de passages par zéro par seconde.
- Les pics à partir des coefficients LPC.
- La fréquence fondamentale.
- Les caractéristiques fréquentielles en utilisant un algorithme de transformée de Fourier rapide (FFT) directement du signal ou bien à partir des coefficients LPC.

3.3 Le module de segmentation

Il consiste à segmenter le signal de parole en grandes classes phonétiques [8] [3] en utilisant des algorithmes procéduraux non contextuels et reposant sur des critères simples, le but essentiel de la segmentation est de :

- réduire l'explosion combinatoire lors de la reconnaissance,
- offrir un cadrage pour le module d'étiquetage .

Les classes concernées sont les voyelles, les plosives et les fricatives. Les autres phonèmes seront considérés comme faisant partie de la classe des inconnus.

3.4 Le calcul d'indices

L'extraction des indices phonétiques pertinents est une étape très importante dans le processus de décodage phonétique. Les valeurs des indices seront utilisées lors de la phase d'étiquetage. Ces indices sont: la durée du segment, le degré de voisement, la position de la barre d'explosion [4], les caractéristiques de cette barre, le suivi de formant, les transitions formantiques, le centre de gravité et la limite inférieure du bruit de friction.

3.5 Le module d'étiquetage

C'est un système expert à base de règles de production et c'est à ce niveau que se fait le décodage proprement dit. En partant des segments fournis par le module de segmentation, le module tente de trouver les bons phonèmes prononcés en utilisant les indices extraits lors de l'étape précédente et les connaissances se trouvant dans la base de règles.

3.6 Le module phonologique

La sortie du module d'étiquetage est un ensemble de phonèmes par segment. Le module phonologique essaye d'éliminer certaines solutions au regard du contexte et des règles phonologiques qui régissent la langue arabe. Dans le système, le module phonologique est intégré dans le module d'étiquetage,

et les règles phonologiques sont contenues dans la base de connaissances du système expert.

4 Le décodeur linguistique : SALAM

SALAM comme Système Approprié pour le décodage Linguistique de l'Arabe Moderne est conçu pour recevoir en entrée une suite de phonèmes sous forme de treillis phonétique pour donner en résultat l'interprétation sémantique de la phrase prononcée. Le décodeur linguistique SALAM comporte plusieurs modules :

4.1 Le module morphologique

L'Arabe possède un système morphologique régulier qui n'a pas beaucoup changé depuis plusieurs siècles. En effet, il existe trois types de racines : tri-radical, quadriradical et 5-radical. Les verbes, les noms et autres adjectives sont dérivés à partir de ces racines selon un modèle régulier, généralement, parmi les modèles suivants :

- Insertion de voyelles spécifiques.
- Préfixage.
- Suffixage.
- Ajout d'une consonne médiane.

Toutefois, les verbes sont dérivés uniquement des racines tri et quadriradicales. Cette régularité devrait être exploitée par le module morphologique. Les quelques irrégularités peuvent être prises comme des exceptions que le module doit stocker dans une table à part.

4.2 Le transcripteur phonético-orthographique

L'algorithme de passage de la transcription phonétique à la forme orthographique du mot doit prendre en compte certaines caractéristiques de la langue en particulier :

- La quantité de la voyelle (brève ou longue).
- La gémination et le tanwin.
- L'assimilation ou non du /l/ avec les consonnes lunaires et solaires.
- Les lettres non prononcées.
- La prononciation du /t/ "fermé" à la fin des noms et des adjectives.
- Les mots irréguliers.

L'absence d'un clavier bilingue, nous contraint de

négliger cette phase et de travailler sur la structure phonétique interne du mot.

4.3 Le module syntaxico-sémantique

La structure grammaticale de l'Arabe standard est rigide, et obéit à des règles qu'on peut formaliser. Cette rigidité n'est pas évidente lorsqu'il s'agit du langage parlé. Comme restriction, nous prenons, dans un premier temps, une grammaire artificielle qui engendre la plus grande partie des phrases. Une phrase correcte de l'Arabe peut être soit nominale (commence par un nom) ou verbale (commence par un verbe). Dans les deux cas, on peut bien introduire une "adat" pour exprimer par exemple une négation ou une interrogation. Le verbe s'accorde en genre et en nombre avec son sujet seulement lorsque le sujet le précède. Dans le cas contraire, il n'y a accord que dans le genre.

Le module syntaxico-sémantique se charge de traduire le treillis de mot fournis par le module phonologique en des phrases ayant un sens en utilisant les règles syntaxiques et sémantiques. Il travaille sur les racines des mots d'une part et d'autre part, sur le résultat du système phonétique sur les voyelles. Les consonnes ont plus de rôles à jouer sur ce plan. Les voyelles servent à régler certains problèmes d'accord.

La vocalisation d'un message, qui consiste à mettre des voyelles sur les consonnes n'est nécessaire en Arabe que pour enlever certaines ambiguïtés dans le sens de la phrase.

4.4 Le système prosodique

Le rôle principal de la prosodie est de déterminer la nature de la phrase. Les indices prosodiques à extraire sont :

4.4.1 La durée

La durée relative d'un phonème dépend de son environnement, de la vitesse d'élocution et d'autres facteurs et elle est très significative dans la langue arabe. Le calcul de la durée vocalique moyenne permet d'avoir une idée sur la vitesse d'élocution.

4.4.2 L'accent

Les syllabes ne sont pas produites avec la même intensité, elles peuvent être accentuées ou non. Nous distinguons en Arabe deux types d'accents :

- Un accent inhérent ou primaire qui se situe au niveau du mot et dépend de la structure syllabique de ce dernier.
- Un accent global ou secondaire qui apparaît au niveau de la phrase.

Les règles qui régissent les syllabes accentuées d'un mot de l'Arabe seront formalisées et utilisées par le module prosodique.

4.4.3 L'intonation

La fréquence fondamentale ou pitch est fonction du voisement des phonèmes. Le modèle du pitch dépend de la phrase. En Arabe, il existe cinq types de phrases :

- Les phrases déclaratives.
- Les commandes.
- Les questions.
- Les appels.
- Les exclamations.

L'extraction du modèle du pitch permet d'avoir une idée sur la nature de la phrase prononcée.

5 Conclusion

Nous avons présenté dans cet article une architecture d'un système de reconnaissance de l'Arabe moderne. Un système de reconnaissance de la parole continue dans un contexte multilocuteur. La partie décodage phonétique est en grande partie opérationnelle; pour trois locuteurs masculins, le système SAPHA arrive à faire la reconnaissance analytique des phonèmes avec un taux global de l'ordre de 80 pourcent. Il reste à introduire de nouvelles connaissances pour améliorer le pourcentage de reconnaissance phonétique et développer les algorithmes du décodeur linguistique.

Références

- [1] Salman H. AL. Ani. *ARABIC PHONOLOGY An Acoustical and Physiological Investigation*. Mouton & Co N.V., The Hague, 1970.
- [2] J. F. Bonnot. *Recherche expérimentale sur la nature des consonnes emphatiques de l'Arabe classique*. Technical Report 9, Travaux de l'institut de phonétique de Strasbourg, 1977.
- [3] A. Callec, S. Monne, M. Querre, O. Travarain, and G. Mercier. Automatic segmentation of phonetic units and training in the real speech recognition system. In *IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 2000-2003, Paris, 1982.
- [4] M. Djoudi. *Détection et localisation de la barre d'explosion en parole continue et dans un contexte multilocuteur*. Technical Report, Centre de Recherche en Informatique de Nancy, 1986.

- [5] M. Djoudi. *Etude phonétique de l'Arabe standard*. Technical Report 89-R-057. Centre de Recherche en Informatique de Nancy, 1989.
- [6] M. Djoudi, D. Fohr, and J. P. Haton. Phonetic Study for Automatic Recognition of Arabic. In *Proceedings of European Conference on Speech and Technology*, pages 268-271, Paris, September, 1989.
- [7] M. Djoudi, D. Fohr, and J. P. Haton. SAPHA : un système expert pour le décodage acoustico-phonétique de l'arabe standard. In *Première Conférence Maghrébine sur le Génie logiciel et l'Intelligence Artificielle*, Constantine, September, 1989.
- [8] D. Fohr. APHODEX : Un système expert en décodage acoustico-phonétique de la parole continue. *Thèse de Doct. Univ. de NANCY 1*, 1986.
- [9] S. Ghazali. Elements of Arabic Phonetics. In *Applied Arabic Linguistics and Signal & Information Processing*, pages 51-58, 1987.
- [10] A. Giannini and M. Pettorino. *The Emphatic Consonants in Arabic*. Giardini editori e stampatori, 1982.
- [11] Gillet and et al. SERAC : Un système expert en reconnaissance acoustico-phonétique. *Actes du 4^{ème} congrès Reconnaissance des Formes et Intelligence Artificielle*, 1984.
- [12] Ibn Jinni. *Sirr SinaaEat Al IEraab*. Mustapha Al Halabi, 1954.
- [13] Y. Laprie. *Notice d'utilisation de Snorri*. Technical Report, Centre de Recherche en Informatique de Nancy, 1988.
- [14] S. De Sacy. *Grammaire arabe*. De Sacy, 1810.
- [15] Sibawayh. *EL KITAB, traité de grammaire arabe*. H. Derembourg, 1889.

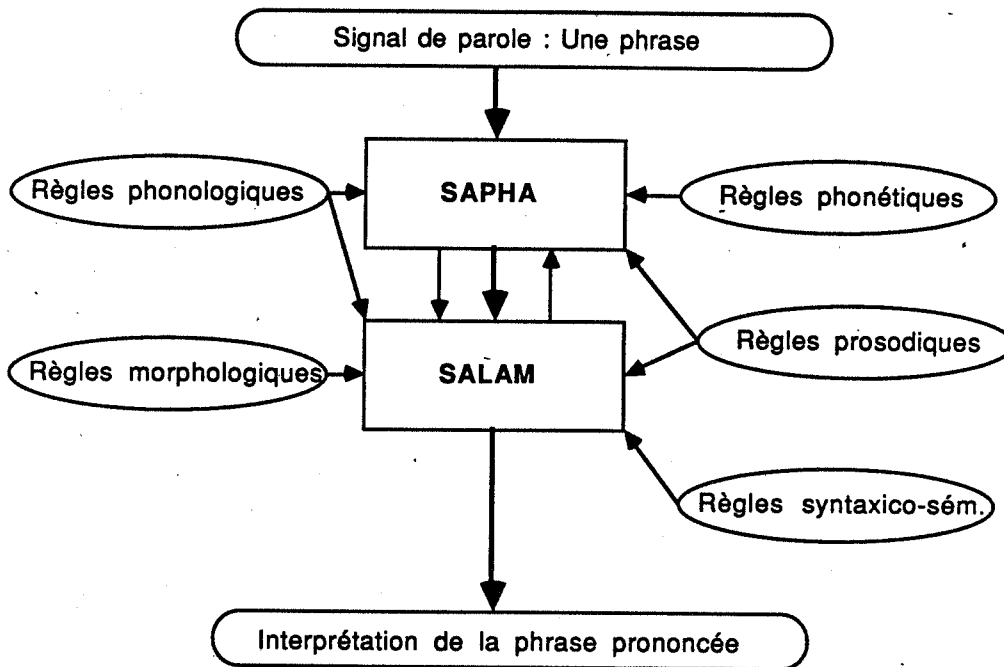


Figure 1 : Architecture de MARS

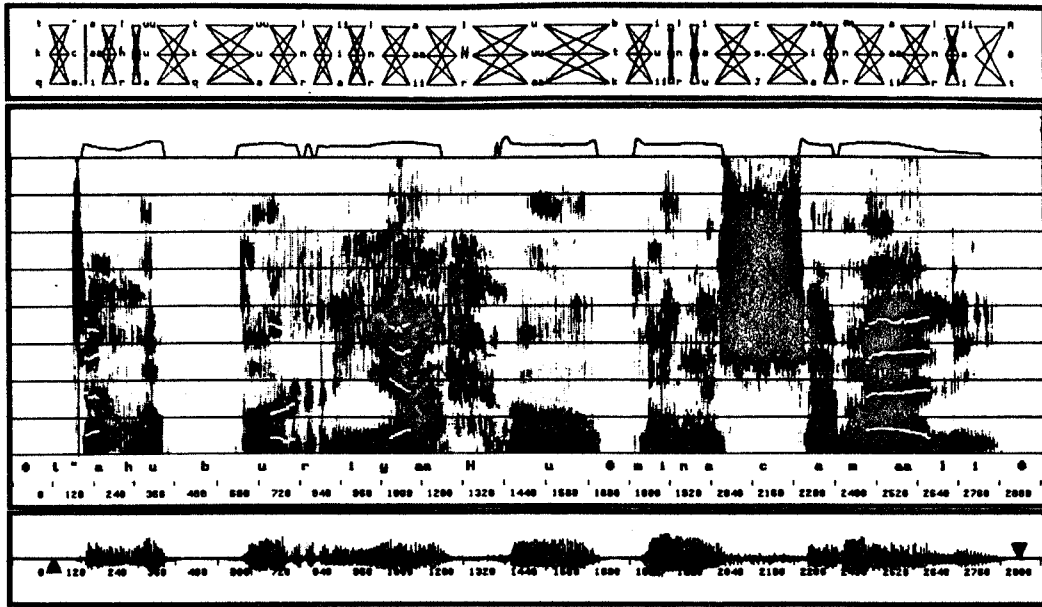


Figure 2 : Etiquetage phonétique d'une phrase

Traduction : "Les vents soufflent du nord"