

CONSTRUCTION DE REFERENCES-MOTS A L'AIDE DE DEMI-SYLLABES

APPLICATION A L'ITALIEN

Brigitte Flocon^{*}, Louis Sauter^{*}, Maura Colombo^{**}

^{*} Laboratoires de Marcoussis, Route de Nozay, F-91460 Marcoussis

^{**} CSELT, 274 Via Romoli, I-10148 Turin

ABSTRACT

Speaker-independent demi-syllable based recognizer has been developed for a French minimal pair vocabulary of 18 words. In the framework of an ESPRIT project, the algorithms have been extended to the Italian language. This paper reports the modifications that have been introduced, in order to perform this experiment and presents the first results. It also points out the kind of problems that may arise when one system has been developed for one language and is extended to another one.

INTRODUCTION

Un système de reconnaissance de mots isolés multi-locuteur utilisant des références globales obtenues par concaténation de demi-syllabes a été développé aux Laboratoires de Marcoussis. Ce système a été testé sur un vocabulaire de 18 mots formé de paires minimales. Cette étude a été réalisée dans le cadre d'un projet Européen du programme ESPRIT, le projet SPIN : Speech Interface at Office Workstation. En collaboration avec le CSELT à Turin, nous étudions l'extension de ce système de reconnaissance à la langue Italienne. Bien que toutes deux langues latines, le Français et l'Italien présente certaines disparités qui ont amenés des modifications dans les algorithmes de segmentation. Les problèmes de co-articulation ont également été examinés. Les premiers résultats des tests mettent en évidence certaines différences dues aux caractères particuliers de chacune des deux langues. Les études sont actuellement poursuivies afin de mieux tenir compte des éléments mis en valeur lors de cette première partie de l'étude.

LE SYSTEME DE RECONNAISSANCE

Le système de reconnaissance décrit ici a fait l'objet de plusieurs présentations par le passé [1], [2]. Nous reprendrons ici uniquement les points principaux, afin de situer précisément les travaux effectués sur l'Italien.

Le système comprend les modules suivants :

Analyse du signal de parole

Le signal est échantillonné à 8 KHz après filtrage passe-bas (fréquence de coupure : 3400Hz) et fenêtre de Hamming. Un jeu de 9 paramètres représentant les coefficients du cepstre (Mel Frequencies Cepstrum Coefficients) est déterminé toutes les 16 millisecondes.

Création des références

Un corpus de 35 mots [3] prononcés par 20 locuteurs est segmenté en demi-syllabes à l'aide d'un algorithme semi-automatique (supervisé par l'opérateur). L'algorithme de segmentation utilise les extrema de l'énergie du signal [4]. Il permet d'obtenir un ensemble de 87 demi-syllabes provenant de 20 locuteurs différents.

Dans une première phase, on recrée par concaténation et lissage chaque mot du vocabulaire de test à partir des demi-syllabes correspondantes pour chacun des 20 locuteurs.

Dans une deuxième phase, un algorithme de classification, utilisant la méthode des nuées dynamiques, détermine 5 classes pour chaque mot-test.

Dans une dernière phase, une référence est obtenue pour chaque classe, grâce à un algorithme de moyennage [3].

Algorithme de reconnaissance

L'ensemble de références est formé de 5 représentants par mot-test. Chaque mot inconnu est comparé à toutes les références à l'aide d'un algorithme de comparaison dynamique (Sakoe et Chiba type 1). Le mot reconnu est celui dont la distance au mot inconnu est la plus petite.

Performances

Le vocabulaire de test qui a été choisi comprend des paires minimales. Il est donné dans la table I. Un corpus de test a été enregistré avec 10 nouveaux locuteurs. Le taux de reconnaissance du système est de 92%.

ADAPTATION A L'ITALIEN

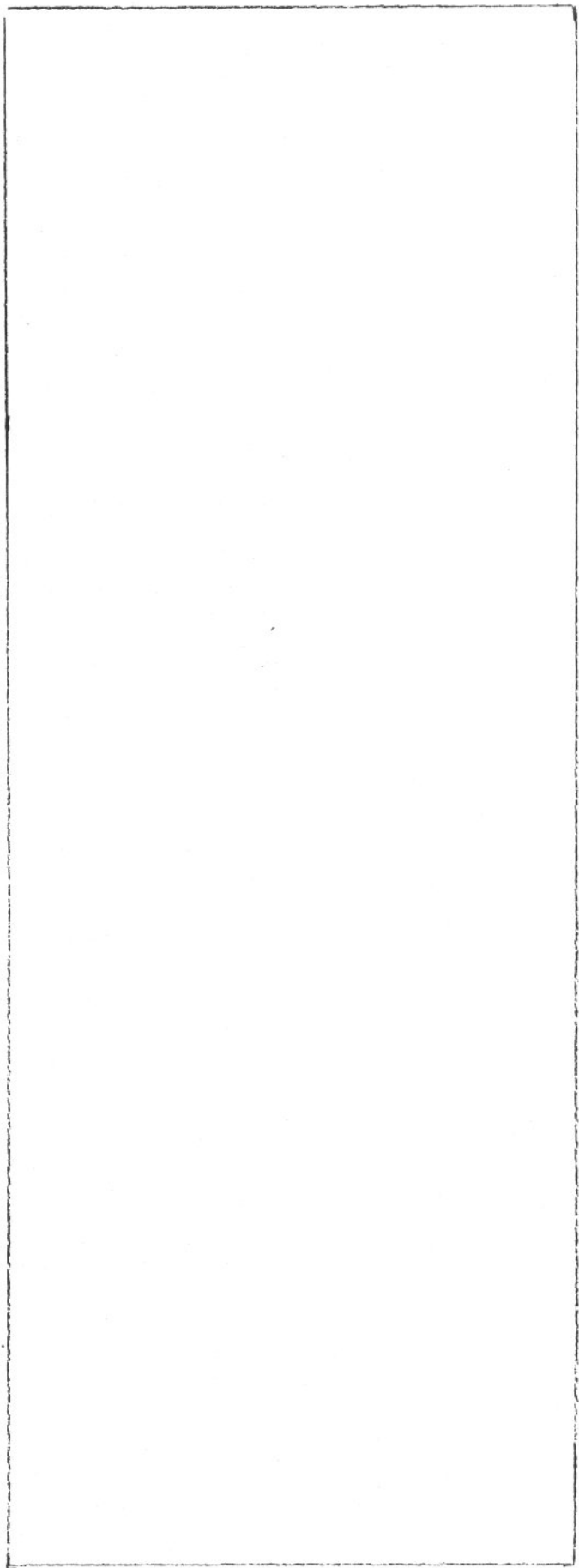
Analyse

La chaîne d'enregistrement de parole disponible dans le laboratoire Italien comporte un filtrage passe-bande (fréquences de coupure : 300 et 3400 Hz). Afin de tenir compte de cet élément, il a été décidé de focer à zéro les sorties des deux premiers filtres utilisés dans le calcul des MFCC.

Création des références

L'analyse du programme de segmentation automatique a fait apparaître certaines différences entre le Français et l'Italien. Certaines modifications ont du être apportées, afin de réaliser une segmentation correcte. Elles sont détaillées ci-après.

Programme de segmentation



TESTS COMPARATIFS

Les tables I et II donnent la liste des vocabulaires utilisés pour tester le système dans les deux langues.

Vocabulaire de test Français		
Cassette	Paté	Passoire
Termite	Traire	Traître
Divin	Vitre	Mité
Mitre	Titre	Cité
Douter	Sente	Santé
Quantité	Fete	Fez

Table I : vocabulaire de test Français

Vocabulaire de test Italien		
ACCETTA	CALCE	CORSO
DANZANTE	DATA	FETTA
FISSO	LANCI	MELA
PASSO	PIATTA	QUELLA
RICCA	RINCORRERE	RINUNCIARE
SOLA	TANTE	TORRE
TRATTA	VELA	VETRI

Table II : vocabulaire de test Italien

Le taux de reconnaissance sur l'Italien est d'environ 75%. Ce chiffre est nettement inférieur à celui obtenu pour le Français.

Une analyse des erreurs a été réalisée ; elle fait apparaître certaines erreurs sur des mots phonétiquement proches (confusion du type MELA/VELA, TRATTA/DATA, RICCA/FETTA...). D'autres erreurs sont dues à des problèmes de rapport de longueur entre le mot inconnu et les mots de références.

En effet, lors de la création de l'ensemble de références et afin d'éliminer tout effet de co-articulation d'une demi-syllabe sur l'autre, certaines demi-syllabes ont été fortement écourtées. Il en résulte, lors de la concaténation que plusieurs références sont anormalement courtes, comparées à des prononciations réelles des mêmes mots. Ce point est en particulier frappant pour les demi-syllabes de fin de mots, qui, pour l'Italien sont essentiellement des demi-syllabes de type CV.

D'autres études sont en cours de réalisation sur l'influence de la longueur des demi-syllabes en fonction de leur position dans le mot.

CONCLUSION

L'étude est arrivée à un point où les premiers résultats sont disponibles. Ces résultats sont intéressants, même s'ils ne permettent pas encore de donner les performances finales du système de reconnaissance pour l'Italien. Une étude plus approfondie de l'italien est en cours afin de pouvoir tenir compte des caractéristiques phonétiques particulières de cette langue par rapport au Français.

REFERENCES

- [1] L. SAUTER, "Isolated word recognition using a segmental approach", ICASSP, Tampa, March 1985.
- [2] L. SAUTER, "Reconnaissance segmentale multilocuteur de mots isolés", Congrès AFCET-Matériels et logiciels pour la 5^{ème} génération, Paris Mars 85.
- [3] B. FLOCON, N. BRIANT, "SYRIL : Système de reconnaissance de mots indépendant du locuteur", 4^{ème} Congrès AFCET-RFIA, Paris, Janvier 84.
- [4] D. FOHR, J.P. HATON, F. LONGCHAMP, L. SAUTER, "Methodes de segmentation syllabique en reconnaissance de la parole", 14^{èmes} d'études sur la parole (GALF), Paris, Juin 1985.