

Anis Zouaghi & Mounir Zrigui
Labo RIADI, Université de Monastir, Faculté des Sciences de Monastir, Tunisie
Mohamed Ben Ahmed
Labo RIADI, Université de la Mannouba
Ecole Nationale des Sciences de l'Informatique, Tunisie

Présentation d'un modèle numérique pour la compréhension de la parole arabe spontanée

Résumé

Notre travail s'intègre dans le cadre du projet intitulé "Oréodule": un système de reconnaissance, de traduction et de synthèse de la parole arabe. L'objectif de cet article est de présenter le modèle probabiliste que nous avons employé pour le décodage sémantique de la parole arabe spontanée, ainsi que les modifications que nous avons apportées à ce modèle pour améliorer ses performances. Ce modèle permet d'utiliser le contexte jugé pertinent, et d'intégrer différents types de données contextuelles pour déduire le sens d'un mot donné. Les résultats trouvés sont satisfaisants, nous avons atteint un taux d'erreur de l'ordre de 25.8%.

Mots-clefs

Compréhension automatique - Parole arabe spontanée - Analyse sémantique - Modèle numérique - Extraction automatique - Contexte pertinent - Information mutuelle moyenne.

1. Introduction

Nous distinguons généralement deux grands courants d'approches pour la compréhension de la parole: les approches symboliques linguistiques (ou par règles), et les approches numériques (ou probabilistes). L'utilisation des modèles numériques pour la reconnaissance [1] et la compréhension de la parole [2] ont l'avantage de réduire fortement le recours à l'expertise humaine. En plus ils sont plus portables vers d'autres domaines ou vers des applications multilingues [3]. La traduction ou dans notre cas l'association automatique à chaque mot de l'énoncé l'ensemble de traits sémantiques correspondant noté TSE [4] et [5], en se basant sur de tels modèles nécessite généralement la considération du contexte d'énonciation. Les informations contextuelles jouent un rôle primordial dans la sélection du TSE adéquat. Ces informations permettent de lever les ambiguïtés d'interprétation et améliorer les performances du système de compréhension [5] et [6].

Dans l'approche standard, le décodage du sens d'un mot est généralement déduit en analysant le contexte qui le précède ou/et qui le succède immédiatement. Or, dans le cas de la compréhension de l'arabe spontané, ceci n'est pas toujours optimum. En effet, nous avons atteint des taux d'erreur assez élevés égales à 57% et à 48,6% [5], en considérant respectivement un décodeur sémantique basé sur un modèle statistique considérant seulement le sens du mot ou des deux mots précédents le mot à interpréter. Afin de remédier à ce problème, nous avons décidé

de ne considérer que les sens des mots pertinents pour la sélection du TSE adéquat à la description de la signification du mot à interpréter. Nous tenons aussi compte du type de l'acte illocutoire accompli par l'énoncé auquel appartient le mot à interpréter (refus, demande, etc.) et de sa nature (par exemple demande de réservation, d'horaires, etc.) pour la prédiction du TSE à utiliser.

2. Les difficultés du décodage sémantique de la parole arabe

Le décodage sémantique des énoncés en langue arabe est une tâche très difficile vu sa richesse sémantique. Cette complexité est due aux spécificités de cette langue, telles que:

- Un mot arabe peut signifier toute une expression en anglais ou en français [7]. Par exemple le mot "أرأيت" (ara ayta), permet d'exprimer en français "Est ce que vous avez vu? ". Ainsi l'interprétation automatique de tels mots nécessite une segmentation préalable de ceux-ci qui n'est pas une tâche facile.

- Le double rôle de la lettre "و", qui peut être employée comme une première lettre d'un mot (par exemple ورق (warak) "feuilles"), ou comme une conjonction de coordination (voir énoncé E). Ce phénomène entraîne une ambiguïté au niveau de la distinction entre les deux rôles possibles de "و", ce qui rend difficile la tâche d'identification automatique des différentes propositions d'un énoncé en vue leur interprétation sémantique.

(E) أريد معرفة توقيت الفطار الذهاب إلى تونس وحجز مكان.

ouridou ma'rifata tawqyta elqitar athaheb ila tuwnis wa hajza makan

- L'ordre des mots dans une phrase arabe est relativement variable par rapport par exemple à la langue française, dont l'ordre d'agencement des mots dans une phrase respecte parfaitement l'ordonnement SVO (Sujet Verbe Objet) [8]. En arabe, nous avons toujours le libre choix entre le terme que nous voulons mettre en valeur au début de la phrase [9]. En plus de cela, l'oral spontané ne respecte pas généralement la grammaire à cause des hésitations, des autocorrections, ... du locuteur. Ceci complique la tâche de construction de la grammaire qui sera utilisée dans le processus d'interprétation sémantique. Il faut prévoir dans la grammaire toutes les règles de combinaisons possibles d'inversion de l'ordre des mots dans une phrase.

- Certaines lettres de la langue arabe comme par exemple : ف - ح - خ - ض - ذ - ظ : sont prononcées au moyen d'une forte expiration, ainsi la qualité du microphone peut influencer sur les résultats de reconnaissance de la parole [10].

- La nature de prononciation de certaines lettres arabes comme par exemple: ر - غ [10].

- La possibilité d'existence de plusieurs réalisations graphémiques pour un même phonème, ou de plusieurs réalisations phonétiques pour un même graphème. Certains graphèmes peuvent même ne peut pas être pris en compte lors de la prononciation [11]. Ce phénomène rend la tâche de reconnaissance plus ardue.

3. Le modèle proposé

3.1. Les modèles couramment utilisés

Dans la littérature plusieurs modèles ont été proposés pour le décodage sémantique de la parole spontanée. Certains utilisent les modèles de Markov cachés [6], d'autres les réseaux de neurones [12], les modèles de langage N-grammes [13], le λ -calcul ou encore les logiques, etc.

Contrairement au latin, la compréhension automatique de la parole arabe spontanée reste encore très peu abordée au niveau de la recherche scientifique. Durant les deux dernières décennies les efforts ont été plutôt concentrés sur la réalisation des analyseurs morphologiques et syntaxiques pour l'arabe tel que par exemple le système *AraParse* [14]. Malgré l'importance de la représentation et de l'analyse sémantique pour la réalisation de n'importe quel système de compréhension, il n'existe que quelques travaux qui s'intéressent à ce domaine en vue du traitement automatique de la langue arabe écrite et non pas parlée [15], [16], [17], [18].

3.2. Présentation du modèle

Pour réaliser un décodeur sémantique de la parole arabe spontanée, nous avons opté pour les choix suivants:

- Une représentation componentielle du sens des mots: chaque mot significatif pour l'application est représenté à partir d'un ensemble de traits sémantiques noté TSE = {domaine, classe sémantique, trait micro sémantique} et un ensemble de traits syntaxiques noté TSY = {genre, nombre, nature}. Les traits de l'ensemble TSE indiquent respectivement le domaine de l'application noté D_i , la classe sémantique à laquelle appartient le mot à interpréter notée C_i , et le dernier trait c'est un trait micro sémantique noté TM_i , qui permet de différencier le sens des mots appartenant à une même classe sémantique. Nous signalons que les mots synonymiques ou possédant un même rôle sémantique possèdent le même ensemble de traits TSE. En appliquant cette représentation, le sens du mot *الذاهب* "allant" par exemple est décrit comme suit:

الذاهب "ethaheb" \rightarrow TSE = {(transport) نقل "naql", (mouvement) حركة "haraka", (destination) وجهة "wijha"} + TSY = {(masculin) مذكر, (singulier) مفرد, (nom) اسم}.

- L'algorithme des K-means [19] pour extraire les classes sémantiques (qui correspondent au deuxième trait des TSE) d'une application finalisée donnée : cet algorithme permet de faciliter la tâche d'identification des classes sémantiques. En plus il a l'avantage d'être rapide face à des données de taille importante (converge à une vitesse linéaire), et simple à implémenter. Cependant, le problème principal de cette méthode est la dépendance du résultat du classement final des informations données en entrée. En effet, les k centres de gravité des k classes sémantiques à déterminer sont choisis d'une manière totalement arbitraire. Dans le cadre de notre travail, cette limite ne pose pas de problème. L'algorithme des k-means est utilisé rien que pour aider et donner une idée à l'utilisateur (surtout si cet utilisateur n'est

pas un expert du domaine) sur la classification possible des mots de l'application d'un point de vue sémantique. Cependant les cartes auto organisatrices de kohonen [20] offrent une alternative efficace, pour ceux qui cherchent des meilleurs résultats de partitionnement [12].

- Une analyse sélective: pour le décodage sémantique des énoncés, nous nous sommes basés plutôt sur une analyse sémantique et nous avons considéré seulement les éléments significatifs pour l'application. Les mots vides sont éliminés lors de la phase du prétraitement de l'énoncé. Cette analyse est plus tolérante aux erreurs grammaticales qui caractérisent la parole spontanée. En plus, elle ne nécessite pas des connaissances linguistiques très approfondies.

- Une méthode basée sur une analyse de corpus: pour la construction de notre structure de représentation de sens SRS [4], nous avons développé une méthode basée sur une analyse de corpus pour l'extraction des mots significatifs, des mots de référence et des classes sémantiques de l'application, et sur une coopération homme/machine pour l'interprétation des mots. Les mots vides sont éliminés en utilisant un filtre lexical. Selon cette méthode le rôle de l'utilisateur est de définir et d'attribuer l'ensemble des TSE et de TSY aux mots. Et le rôle de la machine est de satisfaire les contraintes d'intégrité afin d'aboutir à une SRS non ambiguë et cohérente. Notre système se base en tout sur une dizaine de contraintes. Un exemple de contrainte à vérifier est que: Deux mots différents ne peuvent pas être décrits par un même ensemble de TSE sauf dans le cas où ils sont considérés comme synonymiques ou possédant un même rôle sémantique. Cette méthode nous a permis de faciliter la tâche d'interprétation des mots, ainsi que de la tâche de construction de la SRS et de la maintenance de sa cohérence. La figure 1 suivante, présente l'interface homme/machine utilisée pour interpréter chaque mot via TSE et TSY.

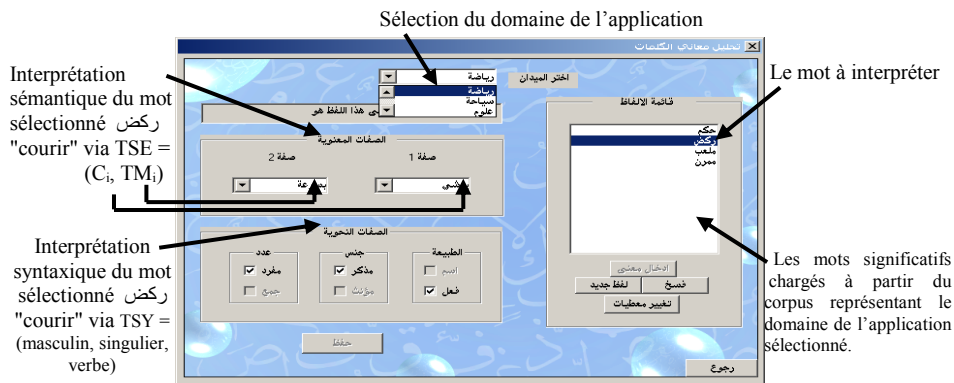


Figure 1: L'interface homme/machine d'interprétation des mots significatifs d'une application.

- Une grammaire probabiliste: cette grammaire participe au choix des TSE adéquats à la description des mots constituant l'énoncé à interpréter. Cette grammaire permet

de tenir compte de plusieurs informations contextuelles en même temps. En plus elle ne considère que les TSE pertinents déjà utilisés pour la prédiction du TSE correspondant à un mot non encore interprété. Ce modèle permet de contraindre l'analyse sémantique des énoncés reconnus, en réduisant l'espace de recherche du décodage sémantique des énoncés. Ceci est réalisé en reposant sur une estimation des probabilités d'interprétation d'un mot donné, sur les mots qui agissent sur son sens (en utilisant la notion d'information mutuelle moyenne), et sur l'utilisation de modèles de type POS tagging pour la détermination de chacun des traits de l'ensemble TSE. L'utilisation des mots agissant sémantiquement sur le mot à interpréter dans le processus d'analyse, permet de surmonter les problèmes de l'oral spontané. Ceci a été prouvé à travers le formalisme *des grammaires de cas de Fillmore*. Quand aux modèles POS tagging, leur performance a été démontrée dans le domaine de l'analyse syntaxique. Ci-dessous l'équation exprimant la probabilité d'interprétation d'un mot M_i par le couple (C_i, TM_i) en tenant compte du type de l'énoncé. Nous remarquons que dans cette formule nous n'avons pas considéré le domaine de l'application puisqu'il est prédéfini à l'avance. Dans notre cas, il s'agit du domaine des renseignements ferroviaires (voir paragraphe 4). Les approximations et les assomptions que nous avons considérées pour l'obtention de ce modèle (décrit par la formule 1) sont détaillées dans [21].

$$P((C_i, TM_i) / M_i, NT_j) = P(NT_j / Mr_k) \times P(C_i / NT_j, M_{i-1}, CP_{i-1}, CP_{i-2}) \times P(TM_i / C_i, TSP_{i-1}) \quad (1)$$

Nous remarquons bien que cette probabilité est calculée en fonction du produit de trois probabilités conditionnelles. La première probabilité $P(NT_j / Mr_k)$ permet d'identifier le type de l'énoncé, s'il s'agit par exemple d'une demande de réservation, d'annulation de billet, etc. Ceci en tenant compte des mots de références Mr_k présents dans l'énoncé du locuteur. Les mots de références sont des uni-grammes, ou des bi-grammes, ou des tri-grammes (qui peuvent être distants) dont les probabilités d'occurrence sont égales à un. Par exemple le bi-gramme $أريد حجز$ qui correspond au tri-gramme "je veux réserver" en français est un mot de référence indiquant qu'il s'agit d'une demande de réservation. La deuxième probabilité $P(C_i / NT_j, M_{i-1}, CP_{i-1}, CP_{i-2})$ permet de déterminer la classe sémantique C_i à laquelle appartient le mot à interpréter M_i , en tenant compte du type de l'énoncé et des deux classes sémantiques pertinentes précédentes CP_{i-1} et CP_{i-2} . Et la troisième probabilité $P(TM_i / C_i, TSP_{i-1})$ permet de déterminer le trait micro sémantique TM_i à attribuer à M_i , en tenant compte de la classe qui a été attribuée à ce mot et du TSE pertinent précédent TSP_{i-1} (voir paragraphe 3.3 suivant en ce qui concerne la méthode d'extraction des TSE pertinents).

3.3. Amélioration du modèle.

En remarquant que ce n'est pas obligatoirement les mots précédant immédiatement le mot à interpréter qui ont une influence sémantique sur ce dernier, nous avons décidé d'employer lors de la phase de décodage du sens d'un mot que les TSE des deux mots possédant la plus grande affinité sémantique avec celui-ci (la manière de calcul de la plus grande affinité sémantique est décrite par la formule 3 ci-dessous). Pour atteindre cet objectif, nous nous sommes basés sur la notion d'information mutuelle moyenne de Rosenfeld [22], qui permet de calculer le degré de corrélation ou de co-occurrence de deux mots donnés. Cette méthode nous a permis de ne plus utiliser systématiquement les TSE des deux mots qui précèdent immédiatement le mot à décoder. La formule de l'information mutuelle moyenne *IMM* est la suivante:

$$\begin{aligned} IMM(M_i, MED_j) = & P(M_i, MED_j) \times \text{Log} [P(M_i / MED_j) / P(M_i).P(MED_j)] + \\ & P(\overline{M_i}, \overline{MED_j}) \times \text{Log} [P(\overline{M_i} / \overline{MED_j}) / P(\overline{M_i}).P(\overline{MED_j})] + P(\overline{M_i}, MED_j) \times \text{Log} [P(\overline{M_i} / \\ & MED_j) / P(\overline{M_i}).P(MED_j)] + P(M_i, \overline{MED_j}) \times \text{Log} [P(M_i / \overline{MED_j}) / P(M_i).P(\overline{MED_j})]; \text{ avec} \\ & 1 \leq j \leq i-1 \end{aligned} \quad (2)$$

Où, les MED_j sont les mots qui appartiennent au contexte droit du mot M_i dans l'énoncé, c'est-à-dire les mots qui précèdent M_i . Nous avons préféré utiliser l'*IMM* (équation 2) plutôt que l'information mutuelle classique, car l'*IMM* permet de calculer en plus l'impact de l'absence d'un mot sur l'apparition de l'autre [12]. L'affinité sémantique la plus forte ou maximale *AffM* que possède le mot M_i avec son contexte droit est déterminée alors à partir de la formule (3) suivante:

$$AffM(M_i, MEd) = \max_{1 \leq j \leq i-1} IMM(M_i, MED_j) \quad (3)$$

3.4. Principe du décodage sémantique

Nous entendons par décodage sémantique d'un énoncé, l'étiquetage de chacun de ses mots significatifs via un ensemble de traits sémantiques TSE. Comme le montre la figure 2 ci-dessous, le décodage sémantique de l'énoncé prétraité repose sur le modèle probabiliste décrit précédemment et un lexique sémantique. Le modèle probabiliste contribue à la sélection des TSE à affecter aux mots de l'énoncé à interpréter, et le lexique sémantique décrit le sens de chaque mot via un ensemble de TSE et un ensemble de TSY. A partir de l'énoncé décodé est déduit son sens. Ceci en remplissant les attributs du formulaire sémantique identifié avec les valeurs correspondantes.

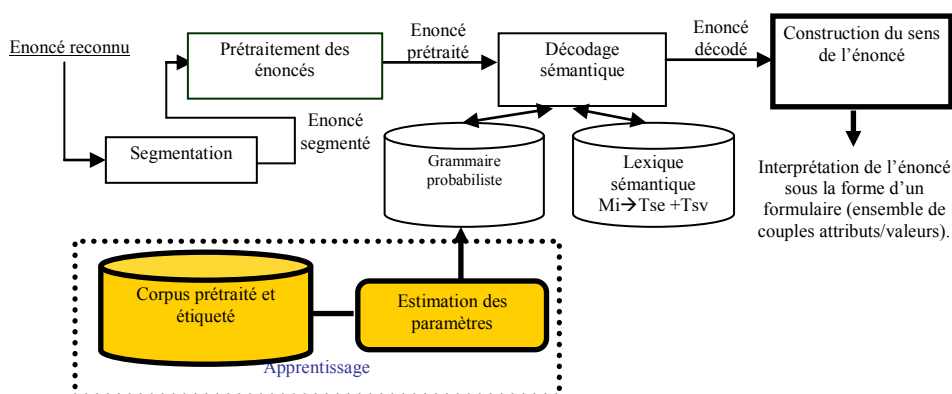


Figure 2: Architecture de l'analyseur sémantique

Au cours de la phase d'apprentissage, nous avons considéré un corpus étiqueté avec les ensembles TSE et prétraité pour l'estimation des paramètres du modèle numérique. Le prétraitement du corpus représentant l'application, nous a permis de simplifier la complexité et réduire la taille du modèle probabiliste. Ce prétraitement consiste comme pour le prétraitement de l'énoncé à éliminer par exemple les mots vides, à regrouper certains mots en une seule entrée, etc.

4. L'application finalisée considérée

4.1 Le domaine de l'application

Pour tester et estimer les paramètres du modèle proposé, nous avons utilisé un corpus représentant le domaine des renseignements ferroviaires. La principale raison de ce choix est la taille statistiquement représentative du corpus d'apprentissage dont nous disposons (voir tables 1 et 2 ci-dessous). Ce corpus a été collecté en demandant à mille personnes différentes de formuler des énoncés relatifs aux renseignements ferroviaires. Donc c'est un corpus simulé et non pas réel.

Domaine	Taille (Mo)	Nombre d'énoncés	Nombre de mots	Nombre de locuteurs
Renseignements ferroviaires	3,4	10000	85900	1000

Table 1: Caractéristiques du corpus de point de vue volume.

Nature de la tâche	Renseignements sur les:				Réservations	autres
	horaires	trajets	tarifs	durées		
Taux de sa représentation					10,41 %	40,64%
	28,7 %	9,37 %	16,66 %	3,12 %		

Table 2: Caractéristiques du corpus de point de vue contenu.

4.2. Les ensembles de traits sémantiques

Pour l'application considérée (décrite ci-dessus), nous avons utilisé en tout comme le montre la table 3 suivante, 12 classes sémantiques *Ci* différentes.

Classes sémantiques <i>Ci</i>	Exemples d'instanciations
طلب (demande)	متى (quand)- بكم (combien)- أحب (je veux)- يوجد (existe)- etc.
حركة (mouvement)	يصل (arrive)- اذهب (je vais)- اذهب (qui va)- etc.
مؤشر حركة (Indice mouvement)	من (de)- عبر (à travers)- نحو (à)- إلى (à)- etc.
مؤشر توقيت (Indice horaire)	الساعة (l'heure)- بتاريخ (à la date)- etc.
رمز (référence)	هذه (ce)- هاته (cette)- etc.
مدينة (ville)	تونس (Tunis)- سوسة (sousse)- etc.
ربط (liason)	و (et)- etc.
عدد تذاكر (nombre billets)	تذكرة (biellet)- مكان (place)- تذكرتين (deux billets)- etc.
حس (bruit)	أن (que)- نهاركم (journée)- etc.
نوع التذكرة (type billet)	مسترسل - ذهاب - وإياب - للصغار - للطلبة - etc.
شرط (condition)	لا تتجاوز أعمارهم (qui ne dépassent pas l'age)- etc.
عدد (nombre)	2 - 1 etc.

Table 3: Les classes sémantiques de l'application considérée.

En ce qui concerne le troisième trait de l'ensemble TSE, nous avons utilisé en tout 20 traits micro sémantiques *TMi*, soit presque le double des *Ci*. Ces traits sont les suivants : طلب_عام (demande_générale) - طلب_ثمن (demande_prix) - طلب_توقيت (demande_horaire) - وجهة (destination) - انطلاق (départ) - عبور (correspondance) - لحظة (moment) - ساعة (heure) - تاريخ (date) - يوم (jour) - درجة (classe), etc.

5. Application du modèle et résultats

Pour tester la performance du modèle numérique défini, 95% du corpus collecté a été consacré à l'apprentissage et 5% aux tests (soit 500 énoncés). Nous avons utilisé comme mesures de performances:

- Le nombre total de mauvaises interprétations N_f défini comme suit : $N_f = N_C + N_{TM}$, où N_C et N_{TM} sont respectivement les nombres de C_i et TM_i incorrectement attribués par le système aux mots de l'énoncé.
- Le taux d'erreur du décodage sémantique: $Taux_{erreur} = N_f / N$; où N est le nombre total de traits C_i et TM_i attribués à l'énoncé à interpréter.
- Le taux de précision est: $Taux_{précision} = N_v / N$; où N_v est le nombre des traits C_i et TM_i correctement attribués.

La figure 3 suivante, présente les taux d'erreur et de précision, trouvés en utilisant le modèle décrit précédemment. Ces taux sont répartis selon le type de renseignement demandé par l'utilisateur: demande de réservation (DR), ou de renseignements sur le trajet (DT), l'horaire (DH), le prix (DP), ou la durée du voyage (DD). Nous pouvons toujours aussi relever le taux d'erreurs des énoncés incorrectement décodés sémantiquement, en considérant le rapport entre les énoncés mal interprétés et le nombre total d'énoncés considérés dans le test (ici 500).

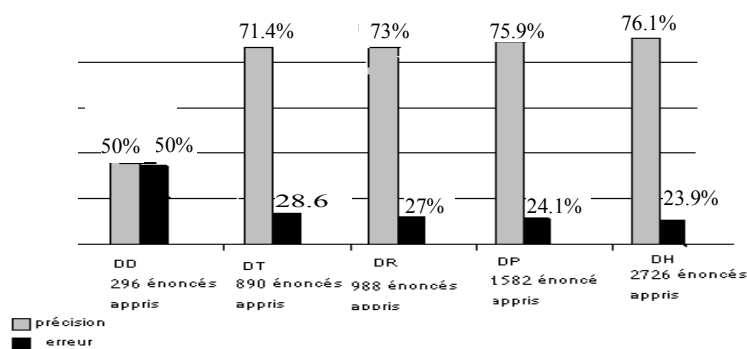


Figure 3: Taux d'erreur et de précision selon le type de la demande de l'utilisateur et le nombre d'énoncés appris.

Le taux d'erreur réellement trouvé lors de la mesure de la performance du modèle est de l'ordre de 30,7%. En analysant davantage les résultats, nous avons conclu qu'un mauvais décodage est obtenu chaque fois qu'il y a un manque de données d'apprentissage. La figure 3 ci-dessus illustre bien ceci. En effet, d'après cette figure, nous remarquons que les résultats de décodage sont bons dans presque tous les types de renseignements demandés par l'utilisateur (DT, DR, DP et DH). Le plus mauvais décodage correspond aux énoncés de type DD. Ceci est dû au fait, que le nombre des

énoncés DD considérés (3,12% du corpus) lors de la phase d'apprentissage du modèle est insuffisant. En effet, nous avons constaté que seulement à partir de 1600 énoncés appris que le modèle numérique devienne performant. A partir de ce seuil, le taux d'erreur est inférieur à 24%. Au dessous de la barre de 500 énoncés, les résultats deviennent inacceptables. Le taux d'erreurs atteint 50% pour 296 énoncés appris, alors qu'il se restreint à 23,9% pour 2726 énoncés appris (voir figure 3 ci-dessus). Donc, une mauvaise interprétation modèle est due essentiellement à un manque de données d'apprentissage, et non pas au type ou à la topologie du modèle numérique défini. En effet, en éliminant les énoncés de type DD, la performance de ce dernier passe de 30,7 à 25.8%. Ces taux sont déjà très satisfaisants, si nous les comparons aux résultats trouvés (57% et 48,6%, voir [5]), en utilisant une approche standard (qui permet généralement de déduire le sens d'un mot en analysant le contexte qui le précède ou/et qui le succède immédiatement).

En analysant notre corpus de test, nous avons remarqué que les taux d'erreur obtenus en employant le modèle numérique présenté, sont dus principalement aux énoncés ayant une structure syntaxique très complexe. Afin de remédier ce problème, certains systèmes combinent une analyse syntaxique profonde avec une analyse sélective tel que le système TINA de [23]. D'autres systèmes utilisent les stratégies d'analyses du TAL robuste [24] et [25]. Ces systèmes sont performants dans des applications ouvertes.

6. Conclusion

Nous avons présenté dans cet article un décodeur sémantique basé sur un modèle de langage hybride, qui permet d'intégrer des données contextuelles lexicales, sémantiques ainsi qu'illocutoire en même temps. Il permet en plus de ne tenir compte que des ensembles de traits sémantiques TSE pertinents dans l'historique du mot à interpréter. Pour cela, nous avons développé une méthode permettant d'extraire automatiquement ces TSE pertinents qui décrivent les sens des mots ayant une influence sémantique sur le mot à interpréter. Ceci est atteint, en se basant sur la notion d'information mutuelle moyenne de Rosenfeld [22]. Les résultats trouvés sont satisfaisants. Dans le prochain avenir nous comptons évaluer notre modèle par rapport aux modèles dits distants ou les modèles obtenus par combinaison linéaire de modèles de langage bien connus comme le maximum d'entropie.

Références

- [17] Al-Johar, B. & Mcgregor. J.: 1997, A Logical Meaning Representation for Arabic (LMRA), *Proceedings of 15th National Computer Conference*.
- [24] Antoine, J-Y., Goulian, J.& Villaneau. J.: 2003, Quand le TAL robuste s'attaque au langage parlé: analyse incrémentale pour la compréhension de la parole spontanée. *Proceedings of TALN*, Batz-sur-Mer.
- [9] Blachère, R. & Gaudefroy-Demombynes, M.: 1975, *Grammaire de l'arabe classique*, Maisonneuve & Larose, Paris.

-
- [6] Bousquet-Vernhettes, C.: 2002, Compréhension robuste de la parole spontanée dans le dialogue oral homme-machine – Décodage conceptuel stochastique. *Thèse de l'université de Toulouse III*.
- [1] Cattoni, R. Federico, M. & Lavie. A.: 2001, Robust analysis of spoken input combining statistical and knowledge-based information sources. *Proceedings of ASRU*, Trento.
- [7] Chaabene, N. & Belguith, L.: 2003, L'étiquetage morpho-syntaxique: comment lever l'ambiguïté dans les textes arabes non voyellés?. *Proceedings des 3^{ème} journées scientifique des jeunes chercheurs en génie électrique et informatique*, Mahdia, Tunisie.
- [11] Ghazali, S., Habaili, H., Zrigui, M.: 1990, Correspondance graphème phonème pour la synthèse de la parole arabe à partir du texte, IRSIT. *Proceedings of Congrès dialogue homme machine*, Tunis.
- [15] Haddad, B., Yaseen. M.: 2005, A Compositional Approach Towards Semantic Representation and Construction of ARABIC. *Proceedings of LACL*.
- [12] Jamoussi. S.: 2004, Méthodes statistiques pour la compréhension automatique de la parole. *Thèse de doctorat de l'université Henri Poincaré*.
- [13] Knight, S., Gorell, G., Rayner, M. Milward, D., Koeling, R. & Lewin. I.: 2001, Comparing grammar-based and robust approaches to speech understanding: a case study. *Proceedings of European conference on speech communication and technology*.
- [20] Kohonen. T.: 1998 *Self-organisation and associative memory*. Springer-Verlag, Berlin.
- [8] Kurdi, M-Z.: 2003, Contribution à l'analyse du langage oral spontané, Thèse de l'Université Joseph Fourier, Grenoble.
- [2] Lefèvre, F.: 2000, Estimation de probabilité non paramétrique pour la reconnaissance markovienne de la parole. *Thèse de l'Université Pierre et Marie Curie*.
- [17] MANKAI Naanaa, C.: 1996, Compréhension automatique de la langue arabe. Application: Le système Al Biruni, *Thèse de doctorat de l'université de Tunis II*.
- [19] McQueen. J.: 1967, Some methods for classification and analysis of multivariate observations, *Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability*.
- [16] Meftouh, K., Laskri, M.T.: 2001, Generation of the Sense of a Sentence in Arabic Language with a Connectionist Approach, *Proceedings of AICCSA*
- [3] Minker, W.: 1999, *Compréhension automatique de la parole spontanée*. L'Harmattan, Paris.
- [14] Ouersighni. R.: 2001, A major offshoot of the Dinar-MBC project: AraParse, a morphosyntactic analyzer for unvowelled Arabic texts. *Proceedings of ACL/EACL*.
- [22] Rosenfeld, R.: 1994, Adaptive statistical language modelling: A maximum entropy approach, *Thèse de doctorat de l'université de Carnegie Mellon*.
- [10] Saïdane, T., Zrigui, M. & Ben Ahmed. M.: 2005, Arabic speech synthesis using a concatenation of polyphones: the results. *Proceedings of Canadian Conference on AI*, Montréal, Canada.

- [23] Seneff. S.: 1992, Robust parsing for spoken language systems, *Proceedings of ICASSP*.
- [25] Van Noord, G., Bouma, G., Koeling, R. & Nederhof, M.J.: 1999, *Robust grammatical analysis for spoken dialogue systems*, Natural Language Engineering 5(1).
- [4] Zouaghi, A., Zrigui, M. & Ben Ahmed, M.: 2004, Une structure sémantique pour l'interprétation des énoncés. *Proceedings of JEP-TALN*, Fès, Maroc.
- [21] Zouaghi, A., Zrigui, M. & Ben Ahmed. M.: 2005, Un étiqueteur sémantique des énoncés en langue arabe, *Proceedings of RECITAL*.
- [5] Zouaghi, A., Zrigui, M. & Ben Ahmed. M.: 2006, L'influence du contexte sur la compréhension de la parole arabe spontanée. *Proceedings of TALN*, Louvain, Belgique.