

Compréhension Automatique de la Parole combinant syntaxe locale et sémantique globale pour une CHM portant sur des tâches relativement complexes

Jérôme Goulian, Jean-Yves Antoine
VALORIA, EA 2593 - Université de Bretagne Sud
Site de Tohannic, rue Yves Mainguy, 56000 Vannes, France
{jerome.goulian,jean-yves.antoine}@univ-ubs.fr

Résumé - Abstract

Nous présentons dans cet article un système de Compréhension Automatique de la Parole (CAP) tentant de concilier les contraintes antinomiques de robustesse et d'analyse détaillée de la parole spontanée. Dans une première partie, nous montrons l'importance de la mise en œuvre d'une CAP fine dans l'optique d'une Communication Homme-Machine (CHM) sur des tâches moyennement complexes. Nous présentons ensuite l'architecture de notre système qui repose sur une analyse en deux étapes : une première étape d'analyse syntaxique de surface (*Shallow Parsing*) générique suivie d'une seconde étape d'analyse sémantico-pragmatique – dépendante du domaine d'application – de la structure profonde de l'énoncé complet.

This paper presents a spoken french understanding system which aims at providing a detailed linguistic analysis as well as preserving the robustness of standard methods. The first part focuses on the importance of a richer analysis for applications that are not dedicated to a very restricted task. We then present our two-level architecture system : a robust independant-domain shallow parsing step followed by a task based analysis of the whole structure of the utterance.

Mots clefs - Keywords : Communication Homme-Machine, Compréhension Automatique de la Parole, robustesse, analyse syntaxique partielle, grammaires de dépendances.

Man Machine Dialogue, speech understanding, robustness, shallow parsing, dependency grammars.

1 Introduction

Au cours de ces dernières années, la Communication orale Homme-Machine a atteint une maturité dont témoigne l'apparition récente de systèmes grands publics opérationnels¹. Ces réussites reposent principalement sur les progrès significatifs des modules de reconnaissance de la parole. Situés en entrée de la chaîne de traitement, leur rôle est de transformer le signal vocal en une (souvent les n meilleures) séquence(s) de mots tenant lieu d'énoncé reconnu. Nous nous intéressons dans cet article aux modules de Compréhension Automatique de la Parole (CAP par la suite) qui se situent en aval de la reconnaissance. Le but de ces systèmes est de fournir une représentation sémantique de l'énoncé exploitable par la suite par les modules de gestion du dialogue. Notons qu'il s'agit ici d'une compréhension hors-contexte, la résolution des références

¹Citons par exemple le système de réservation par téléphone des chemins de fer néerlandais faisant suite au projet européen ARISE (den Os et al., 1999).

en contexte étant traitée par le gestionnaire du dialogue.

La parole spontanée induit l'apparition d'inattendus structurels (hésitations, répétitions, corrections, inachèvements) (Blanche-Benveniste et al., 1990) qui cassent la régularité syntaxique des énoncés. La compréhension de la parole ne peut donc s'envisager aisément sous la forme d'une analyse linguistique détaillée, du moins avec les modèles traditionnels développés pour le traitement de l'écrit. C'est la raison pour laquelle la plupart des systèmes de dialogues développés à l'heure actuelle ne concernent que des domaines applicatifs très finalisés (celui du renseignement aérien –ATIS²– ou ferroviaire). Ces tâches très spécialisées ont en effet permis la mise en œuvre d'approches très pragmatiques ne reposant sur aucune analyse détaillée des énoncés oraux. En particulier, les approches dites *sélectives* (Minker et al., 1999) consistent à ne détecter et ne considérer dans l'énoncé que certains segments-clés nécessaires à l'élaboration d'une requête d'interrogation d'une base de données (sens dit "utile" de l'énoncé). Si ces méthodes se sont révélées robustes face au traitement de la parole spontanée, rien ne garantit cependant qu'elles se révéleraient toujours aussi efficaces sur des domaines d'application plus riches. Ainsi, la question de la généralité des méthodes utilisées constitue une interrogation centrale du domaine (Hirschman, 1998).

Dans cet article, nous nous plaçons précisément dans un cadre applicatif qui, s'il reste finalisé, est néanmoins relativement plus riche que les domaines étudiés classiquement : celui du renseignement touristique. Dans une première partie, nous évoquons les besoins que semble nécessiter une interaction Homme-Machine qui ne soit plus limitée à une recherche d'information très finalisée. Nous présentons ensuite un système de compréhension s'appuyant sur des techniques d'analyses robustes et des formalismes développés pour le TAL. Son objectif est de mener une analyse fine de l'énoncé tout en respectant la contrainte de robustesse imposée par le caractère spontané des productions orales.

2 CAP et tâches moyennement complexes : quelles difficultés ?

En raison du caractère très finalisé³ des domaines d'application classiquement étudiés par les systèmes de compréhension, les énoncés oraux sur lesquels ils travaillent présentent une ambiguïté lexicale très limitée⁴. Les méthodes de CAP sélectives peuvent ainsi se limiter à la recherche de séquences clés (*îlots* ou *segments conceptuels*) dans l'énoncé. Ces segments conceptuels (groupes de mots exprimant un contenu sémantique "utile" à l'application considérée) sont utilisés pour remplir des structures types (ou *schémas*) prédéfinies qui tiennent lieu de représentation sémantique de l'énoncé.

Or la généralisation du dialogue oral à des cadres applicatifs plus riches se traduit par une augmentation sensible de l'ambiguïté lexicale. On peut noter par exemple l'augmentation très sensible de la perplexité des modèles de langage utilisés en reconnaissance de parole entre le domaine ATIS (renseignement aérien) et l'application *Broadcast News* (information générale) (Roukos, 1995). De même, l'étude du corpus Pariscorp (Bonneau-Maynard, Devillers, 1998) que nous avons menée (Goulian, 2000) nous renseigne sur le nombre de schémas (requêtes types) et le nombre de concepts (objets de l'application sur lesquels peuvent porter les requêtes, propriétés de ces objets, etc.) nécessaires dans ce cadre applicatif. Nous comparons ces résultats dans le tableau 1 avec le nombre de schémas et de concepts répertoriés dans deux domaines applicatifs limités (ATIS et MASK) (Minker et al., 1999).

²Air Transport Information System.

³Ceux-ci se caractérisent ainsi par un vocabulaire de généralement moins de 10000 formes fléchies.

⁴Nous ne considérons pas ici l'ambiguïté due aux résultats issus de la reconnaissance de parole.

	rens. aérien(ATIS)	rens. SNCF(MASK)	rens. tourist. (PARISCORP)
schémas	5	8	18
concepts	38	44	113

Table 1: Nombre de schémas et de concepts en fonction du type d'application

On constate clairement une nette augmentation pour le domaine du renseignement touristique. Pour se rendre compte de l'importance de cet accroissement pour la compréhension de la parole, prenons par exemple le concept *horaire*. Ce concept se limite, dans le domaine des renseignements ferroviaires, aux horaires des trains. Dans notre domaine d'étude, ce concept peut se décliner non seulement en horaires de train (ou d'avion) mais aussi en horaires *d'ouverture* et/ou *de fermeture* de *commerces*, de *monuments*, etc. Ainsi, dans l'exemple (1),

(1) "Je crois qu'il y a une *brasserie* dans la salle Méditerranée vous savez celle des *TGV* pour Marseille c'est quoi **ses horaires** au juste"

la brasserie et le TGV sont tous les deux des objets potentiels d'une requête de type horaire, ambiguïté à laquelle risque de se heurter une méthode uniquement orientée par la tâche.

De même, le calcul de la référence hors contexte, qui prend une importance plus grande dans de tels domaines, pose problème. Enfin, (Pierrel, 2000) insiste sur l'importance d'une détection plus fine des intentions de l'utilisateur pour un dialogue réellement coopératif. La prise en compte de domaines applicatifs plus riches risque ainsi de requérir une analyse linguistique détaillée de la structure de l'énoncé (Van Noord et al., 1998).

Si la nécessité d'une analyse détaillée des énoncés semble s'imposer, se pose néanmoins le problème du traitement des inattendus de l'oral (répétitions, hésitations, etc.). Quel type d'analyse envisager ? Une première réponse peut nous être donnée par l'étude linguistique de corpus oraux. Les études de (Blanche-Benveniste et al., 1990) attestent par exemple que ces inattendus présentent des régularités sur lesquelles une analyse automatique peut utilement se reposer. On peut noter entre autres que les répétitions ou les reprises s'effectuent toujours au début des syntagmes enrichis ou avortés. De ce point de vue, une analyse syntaxique de surface garde toute sa pertinence en permettant d'exploiter localement la structure interne des *syntagmes*. De tels outils d'analyse syntaxique robustes ont été développés pour le TAL (Chanod, 2000), certaines méthodes récentes d'analyse structurelle de surface (*Shallow Parsing*) ayant démontré leur capacité à analyser de manière robuste des textes portant sur des domaines relativement larges (Abney, 1996) (Aït-Mokhtar, Chanod, 1997).

Si ces méthodes nous permettent d'envisager une analyse détaillée de surface des énoncés oraux, reste le problème de l'analyse de la structure profonde de l'énoncé. La question de la variabilité de l'ordre des mots en français parlé est à ce niveau importante. Ainsi, on ne pourra pas s'affranchir des problèmes posés par le traitement de la variabilité faible (i.e. mouvements n'entraînant pas de discontinuités (Holan et al., 2000)), problèmes déjà relevés par des systèmes sélectifs utilisés dans un cadre très finalisé (Minker et al., 1999). Pour traiter ces phénomènes, nous envisageons une analyse globale basée majoritairement sur des critères sémantico-pragmatiques⁵. Le traitement des structures discontinues ne semblant pas en revanche être central dans le cadre du dialogue oral homme-machine (Antoine, Goulian, 2001), cette analyse pourra reposer sur un formalisme projectif.

⁵Le terme "pragmatique" fait ici référence au domaine de l'application et non à la prise en compte du dialogue.

3 Un système combinant syntaxe locale et sémantique globale

Nous présentons dans ce paragraphe un système de compréhension ayant pour objectif de permettre la génération d'une représentation sémantique fine de l'énoncé tout en respectant la contrainte de robustesse imposée par le caractère spontané du français parlé. Nous proposons d'adopter une approche proche des outils d'analyse syntaxique robustes développés pour le TAL. Ces systèmes sont conçus pour pouvoir marquer (ou extraire) des structures syntaxiques prédéfinies et légitimées. (Ejerhed, 1993) précise qu'ils s'articulent en général sur deux étapes : une première étape repérant des structures minimales et une seconde étape calculant des structures ou relations plus complexes. La compréhension des énoncés est précisément réalisée suivant une approche comparable en deux étapes (figure 1). L'architecture correspondante est présentée brièvement ici. Nous détaillons chacune des étapes dans les paragraphes suivants.

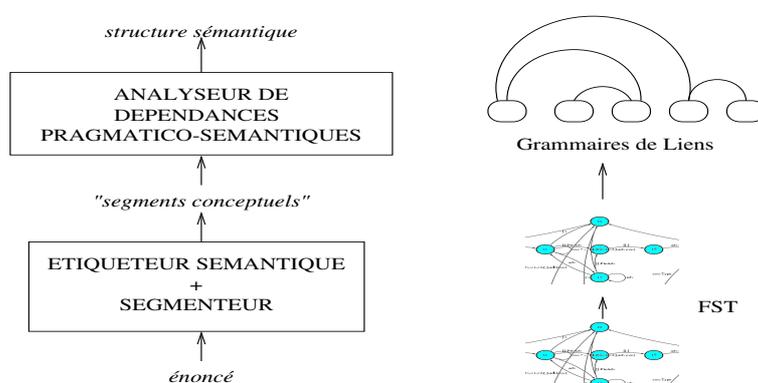


Figure 1: Architecture du système

1. La première étape a pour but de fournir une *analyse robuste et détaillée au niveau du syntagme*. Il s'agit d'une :

- **segmentation syntaxique partielle** de l'énoncé,
- fournissant une structuration partielle de celui-ci en constituants minimaux non récursifs (*Chunks*) (Abney, 1991). Ces groupes correspondent aux unités de reprise en cas d'interruption de la production orale (Blanche-Benveniste et al., 1990),
- suivie d'un **étiquetage sémantique**. Cet étiquetage a pour but d'identifier les dépendances internes à ces constituants existantes autour de leur tête lexicale.

Cette première analyse a été implémentée sous forme de cascades de transducteurs à états finis, compilés à partir d'expressions régulières (utilitaire Fsa (Van Noord, 1997)). Elle présente les deux propriétés principales suivantes :

- **robustesse** : il s'agit d'une analyse partielle ne cherchant pas à porter de jugement syntaxique global sur la totalité de l'énoncé.
- **généricité** : il s'agit d'une analyse syntaxique donc indépendante du domaine d'application; la construction des dépendances locales (l'étiquetage sémantique) réalisé est lui-même générique (cf. infra). Cette caractéristique est importante comme nous l'avons déjà mentionnée (cf. §2).

2. La seconde étape a pour but d'extraire la représentation sémantique complète de l'énoncé par la recherche des dépendances globales entre les têtes lexicales associées aux segments détectés⁶. Il s'agit d'une analyse :

⁶Les dépendances locales internes aux syntagmes ayant déjà été construites dans l'étape précédente.

- majoritairement dépendante de l'application, la recherche des dépendances s'effectuant à un niveau **sémantico-pragmatique**. Cette analyse est proche en ce sens des méthodes sélectives mais récupère néanmoins les structures issues de l'analyse détaillée réalisée lors de la première étape.
- utilisant le formalisme des grammaires de liens (Sleator, Temperley, 1991). Ce formalisme lexicalisé, inspiré des grammaires de dépendances, repose sur le postulat de la **projectivité** du langage étudié (cf. §2). Les travaux de (Grinberg et al., 1995) ont montré la capacité des grammaires de liens à analyser de façon robuste des corpus de dialogues oraux. On notera que dans un tel formalisme, les dépendances syntaxiques sous-tendent les dépendances sémantiques (Mel'cuk, 1988). Ceci, ajouté au fait que notre analyse – partielle – reste finalisée, devrait faciliter l'adaptation de ce formalisme, conçu pour l'analyse de la syntaxe, à notre problématique.

3.1 Segmentation syntagmatique de l'énoncé et étiquetage sémantique

Cette étape se veut minimale et superficielle et ne cherche pas à lever toute l'ambiguïté structurale. En particulier, les rattachements (syntagmes prépositionnels, portée des conjonctions de coordinations, etc.) sont laissés à l'étape suivante et s'effectuent, dans notre cas, sur des critères sémantico-pragmatiques. Plusieurs actions sont réalisées ; nous les détaillons ci-dessous en prenant l'exemple de l'analyse de l'énoncé (2).

(2) *je voudrais connaître les tarifs du petit du petit restaurant chinois près de la gare non près de l'hôtel*

1. **Pré-étiquetage.** Cette première sous-étape a pour objectif le repérage des locutions et le *pré-étiquetage syntaxique* des mots constituant l'énoncé. Nous avons choisi un nombre réduit d'étiquettes syntaxiques qui nous paraissent suffisantes pour caractériser les relations de sous-catégorisation nécessaires à la modélisation du syntagme. En effet, notre domaine d'application restant relativement finalisé⁷, l'utilisation d'informations morphologiques ne nous est pas utile. Notre étiqueteur repose sur des règles contextuelles⁸ écrites sous forme d'expressions régulières. En cas d'ambiguïté, nous conservons les différentes séquences, traitées en parallèle par les étapes suivantes.

pr_pers[je] mod[voudrais] v_inf[connaître] art[les] n[tarifs] prep[du] adj[petit] prep[du] adj[petit] n[restaurant] adj[chinois] prep[près-de] art[la] n[gare] neg[non] prep[près-de] art[l] n[hôtel]

2. **Segmentation syntaxique en syntagmes.** Nous proposons un découpage en constituants minimaux à l'image d'(Abney, 1991) que nous adaptons au cadre du dialogue oral. Le tableau 2 répertorie quelques uns des groupes syntagmatiques que nous considérons. Trois catégories syntagmatiques sont ainsi envisagées. La première correspond aux grands groupes syntaxiques classiques de l'écrit. S'agissant de constituants minimaux, les groupes verbaux par exemple n'incluent jamais leurs arguments. En revanche, leurs modalités (présence de négation, d'un modal) sont identifiées. La seconde correspond aux expressions langagières indépendantes du domaine d'application (date, heure, adresse). Enfin, la dernière catégorie est spécifique au dialogue oral. Elle permet de repérer et de délimiter les marqueurs des inattendus dans la production orale. Conserver ces marqueurs nous semble judicieux pour aider le traitement de ces phénomènes dans la phase suivante. Des approches de prétraitements (patterns d'extragrammaticalité par exemple (Bear et al., 1992)), non implantées à l'heure actuelle dans notre système, pourraient néanmoins être utilisées.

⁷Ambiguïté lexicale relativement limitée.

⁸On notera à ce sujet qu'à l'heure actuelle, les approches à bases de règles et les approches stochastiques présentent des performances équivalentes que ce soit pour l'étiquetage morpho-syntaxique ou la CAP.

Catégorie	Groupe	Exemples
syntaxe	GVerbaux GAdjectivaux GNominaux GPrépositionnels GAdverbiaux Coordinations Pronoms relatifs	GVmod[<i>veux réserver</i>],GV[<i>coûte</i>],GVinf[<i>aller</i>]... Gadj{[<i>chinois</i>],[<i>très cher</i>]... GN{[<i>l'hôtel Caumartin</i>],[<i>deux places</i>], ...} GPNom[<i>du quartier</i>],GPLoc[<i>près de la gare</i>],GP[<i>avec restaurant</i>]... Gadv[<i>essentiellement</i>]... Coo[<i>ou</i>]... PrRel[<i>qui</i>]...
expressions langagières	Date Heure Adresse	Date{[<i>mardi 6 mars</i>],[<i>mardi prochain</i>]... Heure{[<i>trois heures du matin</i>],[<i>cinq heures moins le quart</i>]... Adresse[<i>rue de la paix</i>]...
spécificités de l'oral	Hésitations Corrections	Hes[<i>euh</i>]... Cor{[<i>non pardon</i>],[<i>non en fait</i>]...}

Table 2: Groupes syntagmatiques utilisés pour la segmentation

En pratique, cette analyse résulte de l'application d'une séquence finie et ordonnée de transducteurs. Chaque transducteur est utilisé pour introduire, dans l'énoncé étiqueté syntaxiquement, des marqueurs de délimitation autour des instances d'un groupe particulier. Chacun de ces groupes est décrit par un ensemble de règles exprimées au moyen d'expressions régulières. L'ambiguïté de segmentation est gérée par l'heuristique suivante : nous privilégions comme (Abney, 1991) la détection des groupes les plus longs. Ce choix permet la capture, dès ce niveau de l'analyse, d'un maximum de dépendances locales sur lesquelles l'analyse suivante n'aura pas à revenir. Pour ce faire, les expressions intègrent un opérateur spécifique exprimant une relation de remplacement contrainte par la direction gauche-droite de l'analyse et favorisant cette détection *maximale* des instances considérées (Karttunen et al., 1996) (Karttunen, 1996). Nous donnons ci-après un exemple (simplifié) d'une telle règle pour la délimitation des instances maximales des groupes nominaux, formées par la succession d'un article optionnel, d'un ou plusieurs groupes adjectivaux, d'un nom puis d'un ou plusieurs groupes adjectivaux⁹, suivi du résultat de l'analyse de l'énoncé (2).

```
replace(('GN' --> [c(art)^ ,c('GAdj')*,c(n),c('GAdj')*]))
```

```
GN[pr_pers(je)] GVmod[mod(voudrais) v_inf(connaître)] GN[art(les) n(tarifs)] prep[du] Gadj[adj(petit)]  
GPNom[prep(du) GN[Gadj(adj(petit)) n(restaurant) Gadj(adj(chinois))] GPLoc[prep(près-de) GN(art(la)  
n(gare))] Cor[neg(non)] GPLoc[prep(près-de) GN(art(l) n(hôtel))]
```

Notre segmenteur a fait l'objet d'une évaluation préliminaire sur 300 énoncés dans le cadre de la campagne d'évaluation "par défi" des systèmes de CAP initiée par le Groupe de Travail 5.1 du GDR-PRC-I3 (Antoine, 2001). Le tableau 3 donne le pourcentage d'erreur en *substitution* (étiquette erronée), en *scission* (un chunk attendu segmenté en plusieurs chunks), et en *regroupement* (plusieurs chunks attendus regroupés en un seul).

	%Erreur		%Précision	%Décision
<i>Substitution</i>	<i>Scission</i>	<i>Regroupement</i>		
3.38%	6.4%	2.1%	88.12%	100%

Table 3: Evaluation préliminaire du segmenteur

Notons que cette précision de 88%, obtenue sur cette évaluation préliminaire, pourra être améliorée. Notre décision à 100% résulte quant à elle du choix de notre stratégie

⁹replace est l'opérateur défini dans (Karttunen, 1996), --> introduit les marqueurs [...] autour des instances détectées, c(X) repère une catégorie X et transforme les marqueurs [...] en marqueurs (...), cm(X, m) est identique à c(X) mais valable uniquement pour le mot m ; les autres opérateurs étant ceux des expressions régulières.

d'analyse.

3. **Étiquetage sémantique.** Il s'agit d'un remplacement, réalisé au fur et à mesure de la segmentation, des étiquettes syntaxiques par des étiquettes sémantiques "génériques" articulées autour de la *tête lexicale* des groupes détectés. Ces étiquettes sont issues de l'étude du corpus Pariscorp. Elles ont été validées sur 1000 énoncés mais restent indépendantes du domaine d'application. Elles correspondent en effet à de grands types sémantiques : la classe générique des *objets* de l'univers quel que soit celui-ci, la classe des *propriétés* portant sur ces objets, la classe des *relations* sur des objets, etc. Chaque groupe reçoit la classe sémantique correspondant à sa tête lexicale (étiquetée 'T' pour 'Type'). Le tableau 4 donne un aperçu de cette classification (exemples du tableau 2).

Syntagmes	Classes sémantiques	Exemples
GV Gadj GNominaux GP	Acte PteObj Obj Pte	ActeMod[mod(<i>veux</i>) T(<i>réserver</i>)], Acte[T(<i>coûte</i>)], Action[T(<i>aller</i>)] PteObj{[T(<i>chinois</i>)], [PteModifieur(<i>très</i>)T(<i>cher</i>)]}... ObjDéfini[PteNb(1) T(<i>hôtel</i>) Référence(<i>Caumartin</i>)], Obj[PteNb(2) T(<i>places</i>)], ... PteRéférence[T(<i>quartier</i>)], Pte[T(<i>avec</i>) Obj(<i>restaurant</i>)], PteLoc[T(<i>près-de</i>) ObjDéfini(PteNb(1) T(<i>gare</i>))]... PteModifieur[T(<i>essentiellement</i>)]...
Gadv Coo PrRel	PteModifieur Coo ObjRéférent	Coo[T(<i>ou</i>)]... ObjRéférent[T(<i>qui</i>)]...
Date Heure Adresse	Date Heure Adresse	Date{[Nomjour(<i>mardi</i>) Numjour(6) Nommois(<i>mars</i>)], Heure{[Nheure(3) PteMomentjournée(<i>matin</i>)], [NHeure(5) ModifieurMinutes(<i>moins</i>) NMinutes(15)]}... Adresse[T(<i>rue</i>) PteRéférence(<i>de-la-paix</i>)]
Hes Cor	Hes Cor	Hes[T(<i>eah</i>)]... Cor[T(<i>non</i>) Pte(<i>pardon</i>)], ...

Table 4: Correspondances catégories syntagmatiques / classes sémantiques utilisées

Le groupe nominal "un restaurant chinois" de tête lexicale "restaurant" correspond ainsi à un *Objet* de type *restaurant* possédant deux propriétés : une spécialité : *chinois* (PteObj) et le fait d'être indéfini : article *un* (PteDéfini(non)). Ces propriétés correspondent à des *dépendances locales* autour de la tête lexicale, internes à ces segments. Cet étiquetage, basé sur des règles génériques de correspondances catégories syntagmatiques / classes sémantiques et réalisé en partie en contexte, servira de base au rattachement (dépendances globales) effectué en seconde partie¹⁰. La figure 2 donne le résultat de l'analyse de (2).

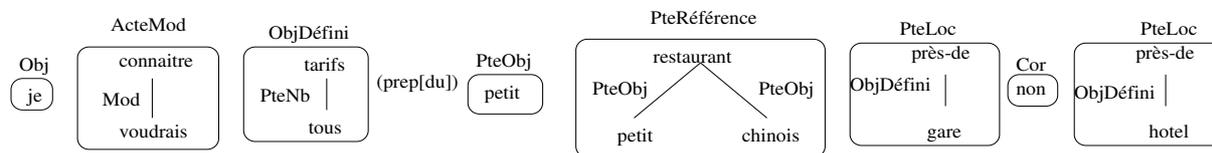


Figure 2: Segmentation de l'exemple (2) et construction des dépendances locales sémantiques

3.2 Analyse de la structure de l'énoncé par rattachements pragmatiques

Cette deuxième étape a pour but d'établir les *relations de dépendances globales* existantes entre les segments identifiés. Nous utilisons pour cela un formalisme lexicalisé : les Grammaires de Liens (*Link Grammars*)(Sleator, Temperley, 1991). Inspiré des *grammaires de dépendances* (Tesnière, 1959), ce formalisme envisage traditionnellement la structure de la phrase non plus

¹⁰Un étiquetage lexicalisé dépendant de l'application est tout de même réalisé (transducteurs) avant la seconde phase d'analyse. Il permet par exemple de remplacer l'étiquette *PteObj* de l'objet *restaurant* en *Spécialité*.

en termes de constituants mais uniquement selon les *relations syntaxiques* que les mots entretiennent entre eux. Chaque unité lexicale possède des connecteurs indiquant les liens pouvant être établis avec les autres mots de la phrase. Les liens sont établis en respectant cette contrainte locale sur l'ordre des mots de la phrase. L'analyse repose également sur deux contraintes globales : la connectivité (tous les mots doivent être reliés entre eux) et la planarité (les liens ne doivent pas se croiser – non-projectivité – cf. §2). Nous proposons deux adaptations de ce formalisme :

1. Les liens ne s'effectuent plus sur les mots de l'énoncé mais uniquement sur les têtes lexicales des segments identifiés. Deux observations doivent être faites :
 - les dépendances locales identifiées précédemment n'interviennent donc plus ici.
 - les mots de l'énoncé n'ayant pu être insérés dans un syntagme (pour (2) c'est le cas de la préposition *du*) sont exclus de cette analyse. Le caractère partiel de l'analyse n'est lié qu'à l'étape précédente, la contrainte de connectivité étant respectée ici.
2. Les relations envisagées sont ici essentiellement pragmatiques i.e. dépendantes du domaine d'application. Elles ont été caractérisées manuellement sur le corpus Pariscorp (Goulian, 2000). On distingue deux types de relations :
 - Celles caractérisant la nature de l'énoncé. Pour une requête de type *Sélection* (demande d'une liste d'objets ayant certaines propriétés), on retiendra par exemple deux types de relations. La première (*ObjSélection*) fait le lien entre la nature de la requête et la nature de l'objet à sélectionner; la seconde (*CritèreSélection*) indique le ou les critères souhaité(s) sur les objets (critère de proximité, de tarifs, etc.)
 - Celles caractérisant les propriétés des objets. Les relations expriment alors pour chaque type d'objet du domaine les propriétés valides qu'il peut supporter. Pour un *hôtel* par exemple, il s'agira de relations de *Catégorie* (1 étoile), de *Services* proposés, etc.

Par ailleurs, des relations particulières permettent de traiter les inattendus structuraux tels que les reprises marquées, les conjonctions de coordination et les relatives (cf. §3.2.1).

3.2.1 Représentation des relations

Nous expliquons dans ce paragraphe comment les relations évoquées ci-dessus sont représentées en grammaires de Liens. Chaque relation est exprimée au moyen de formules logiques. En pratique, nous utilisons deux types de règles :

des règles lexicalisées simples, permettant d'exprimer des relations dépendantes de la tête lexicale du segment considéré. Nous donnons ci-dessous deux exemples (très simplifiés pour les besoins de l'explication) de telles règles nécessaires au traitement de l'énoncé (2).

- (1) $PteRéférence(T(\textit{restaurant})) : \textit{Objet} - \textit{Tarif}^* \& \{ \textit{Taille}^- \} \& \{ \textit{Spécialité}^+ \} @$
- (2) $ObjDéfini(T(\textit{tarif})) : \textit{Objet} - \textit{Tarif}^*$

La règle (1) concerne uniquement le mot *restaurant* correspondant dans la phrase à la catégorie *PteRéférence* et indique qu'il doit être *objet d'une requête* (dans cet exemple simplifié *Objet-Tarif*) placée avant ou après lui dans l'énoncé (opérateur *) et possédant éventuellement (entre { }) une relation de type *Taille* située à sa gauche et une ou plusieurs relations (opérateur @) de type *Spécialité* située(s) à sa droite dans l'énoncé (opérateur +). De la même manière, la règle (2) indique que le mot-clé *tarif* doit faire l'objet d'une relation *Objet-Tarif*.

Ces règles permettent par ailleurs de traiter les inattendus non marqués (règle (1) en ce qui concerne la répétition non marquée présente dans l'énoncé (2)).

des méta règles, correspondant à un ensemble factorisé des règles précédentes et permettant d’exprimer des relations valables pour toutes ou un sous-ensemble des têtes lexicales possibles pour un segment donné (chaque règle particulière en est automatiquement déduite). C’est le cas par exemple de la règle (3) qui indique que tous les segments étiquetés *PteLoc*, quelle que soit leur tête lexicale (symbole ***) (préposition *près-de*, *loin-de*, etc.), peuvent conduire à un lien de type *Localisation-Objet*. De même la règle (4) indique la possibilité d’une relation *LocalisationObjet* entre des objets *PteRéférence* dont la tête lexicale appartient à l’ensemble *ObjLocalisables*.

- (3) $PteLoc(T(*)) : LocalisationObjet^-$
- (4) $PteRéférence(T(ObjLocalisables)) : LocalisationObjet^+$

De telles méta règles permettent entre autres de gérer les coordinations ainsi que les inattendus marqués (notamment les corrections) qui sont traités de manière identique¹¹ : les liens se font sur des segments de même classe sémantique ou incluant cette classe. La règle (5) exprime par exemple que la coordination doit être reliée à droite *et* à gauche à des segments de même classe sémantique.

- (5) $Coo(Type(*)) : (PteLoc^- \& PteLoc^+) | \dots$

3.2.2 Obtention de la représentation sémantique finale

La représentation sémantique finale de l’énoncé est obtenue par la structure établie à l’issue de cette analyse (figure 3), elle-même transformée en requête d’interrogation de base de données.

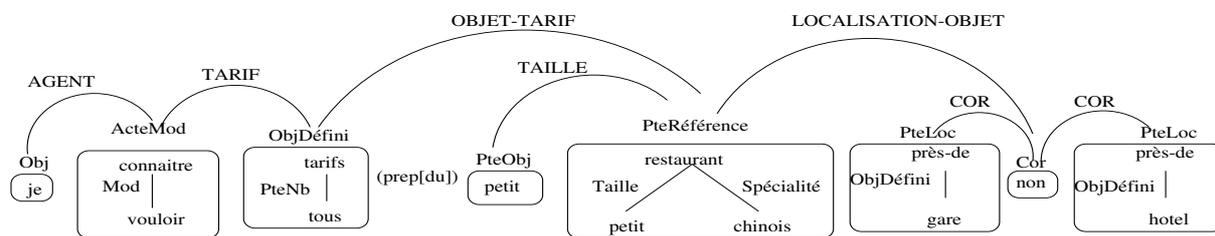


Figure 3: Construction des dépendances globales sur l’exemple (2)

L’algorithme utilisé repose sur la contrainte de connectivité. Cette contrainte ne permet pas en l’état de gérer les inachèvements ou les incises. Nous envisageons d’adapter notre algorithme en relâchant cette contrainte en cas d’échec de l’analyse.

4 Conclusion

Le système de compréhension présenté résulte d’une mise en œuvre conjointe de méthodes issues du TAL robuste et de la CAP sélective. Son objectif est de permettre une analyse robuste et détaillée d’énoncés oraux dans le contexte d’une CHM portant sur des tâches relativement complexes. L’analyseur des dépendances globales est à l’heure actuelle en cours de développement¹². Une première phase d’évaluation portera ensuite sur 1000 énoncés du corpus Pariscorp et permettra d’apprécier la qualité de l’analyse linguistique détaillée réalisée et la robustesse de l’ensemble du système. Cette évaluation se poursuivra dans le cadre de la campagne du G.T. 5.1 du GDR-PRC I3 (Antoine, 2001).

¹¹Le traitement des corrections demande encore à être évalué en détail. Il reste toutefois que leur présence ne perturbe pas l’analyse et que les marqueurs, s’ils existent, sont conservés pour d’éventuels pré-traitements ou post-traitements à la dernière étape de l’analyse.

¹²Activités de recherche financées par le conseil régional de Bretagne.

Références

- Abney S. (1991), *Parsing by Chunks*, Kluwer, in Berwick, A. & T. (ed.), “Principle Based Parsing”.
- Abney S. (1996), Partial parsing via finite-state cascades, Actes de *W. Robust Parsing, ESSLLI'96*, 8-15.
- Aït-Mokhtar S., Chanod J.P. (1997), Incremental finite-state parsing, Actes de *ANLP'97*, 72-79.
- Antoine J.Y., Goulian J. (2001), Etude des phénomènes d'extraction en français parlé sur deux corpus de dialogue oral finalisé, application à la CHM orale. *T.A.L.*, Vol. 42.1, (à paraître).
- Antoine J.Y. (2001), Evaluation des systèmes de CAP, Campagne d'évaluation “par défi”, Rapport technique, *GDR-PRC-I3, Pôle Parole, G.T. 5.1.*, http://www.univ-ubs.fr/valoria/antoine/Gt51/Eval_defi.html
- Bear J., Dowding J., Shriberg E. (1992), Integrating multiple knowledge sources for detection and correction of repairs in Human-Computer dialogue, Actes de *ACL'92*, 56-63.
- Blanche-Benveniste C., Bilger M., Rouget C., van den Eynde K. (1990), *Le Français parlé : études grammaticales*, Paris, CNRS Editions.
- Bonneau-Maynard H., Devillers L. (1998), Acquisition, Transcription et Annotation du Corpus Pariscorp, Rapport interne, *Action de Recherche Concertée “dialogue oral” de l'AUF*
- Chanod J.P. (2000), *Robust Parsing and Beyond*, Kluwer, in Van Noord & Junqua (ed.), “Robustness in Language Technology”.
- Ejerhed E. (1993), Nouveaux courants en analyse syntaxique, *T.A.L.*, Vol. 34.1, 61-82.
- Goulian J. (2000), Représentations sémantiques pour la CAP dans le cadre d'une application de demande de renseignements touristiques, Rapport interne, laboratoire VALORIA, Université de Bretagne-Sud.
- Grinberg D., Lafferty J., Sleator D. (1995), A robust parsing algorithm for Link Grammar, Rapport de recherche CMU-CS-TR-95-125, School of Computer Science, Carnegie Mellon University.
- Hirschman L. (1998), Language understanding evaluations : lessons learned from MUC and ATIS, Actes de *LREC'98*, Grenade, Espagne, 117-122.
- Holan T., Kubon V., Oliva K., Plátek M. (2000), On complexity of word order, *T.A.L.*, Vol. 41.1, 273-300.
- Karttunen L. (1996), Directed Replacement, Actes de *ACL'96*, Santa-Cruz.
- Karttunen L., Chanod J.P., Grefenstette G., Schiller A. (1996), Regular expressions for language engineering, *Natural language Engineering*, Vol. 2(4), 305-328.
- Mel'cuk I.A. (1988), *Dependency Syntax, Theory and practice*, State University of New York Press, Alb.
- Minker W., Waibel A., Mariani J. (1999), *Stochastically based semantic analysis*, Amsterdam, Kluwer.
- den Os E., Boves L., Lamel L., Baggia P. (1999), Overview of the ARISE project, Actes de *Eurospeech'99*, Budapest, Hongrie, 1527-1530.
- Pierrel J.M., Romary L. (2000), *Dialogue Homme-Machine*, in Pierrel (ed.) “Ingénierie des langues”, Hermès, 331-349.
- Roukos S. (1995), *Survey of the state of the art in Human Language Technology*, Chapter language representation, in Cole R.A. et al. (ed.), NSF & DG XIII de la communauté Européenne, CSLU, 35-42.
- Sleator D., Temperley D. (1991), Parsing English with a Link Grammar, Rapport de recherche CMU-CS-91-196, School of Computer Science, Carnegie Mellon University.
- Tesnière L. (1959), *Eléments de syntaxe structurale*, Paris, Klincksiek.
- Van Noord G. (1997), *Fsa utilities : a toolbox to manipulate finite-state automata*, Springer Verlag, in Raymond, Wood & Yu (ed.), “Automata Implementation”, 87-108.
- Van Noord G., Bouma G., Koeling R., Nederhof M.J. (1998), Robust grammatical analysis for spoken dialogue systems, *Natural language Engineering*, Vol. 1, 1-48.