

Modéliser le langage

Christian Retoré

LaBRI (Université de Bordeaux & CNRS)
Responsable du thème Informatique et Linguistique
INRIA Bordeaux Sud-Ouest
Responsable de l'équipe Signes:
Signes linguistiques, grammaire et sens,
algorithmique logique de la langue
Rédacteur en chef de la revue
Traitement Automatique des Langues

29 nov 2007

Université Ouverte Cergy

1

Modéliser le langage

La **linguistique computationnelle** au carrefour
de **l'informatique**, de **la linguistique**,
des **sciences cognitives**
et des **outils pour l'Internet**.

29 nov 2007

Université Ouverte Cergy

2

Linguistique et informatique

Une connexion productive
des origines de l'informatique
à nos jours

29 nov 2007

Université Ouverte Cergy

3

Plan

- Linguistique Computationnelle: historique
- Quelques applications
- Les modules de la linguistique
- Grammaire générative principes, hiérarchie, acquisition
- Logique et sémantique richesse des questions classiques
- Un cas d'école: syntaxe de la LSF

29 nov 2007

Université Ouverte Cergy

4

Linguistique et informatique: une longue histoire (1/2)

- 1949 Machine Translation aide à la traduction,
 - The flesh is weak but the spirit is willing
 - The meat is rotten but the vodka is strong
 - La chair est faible mais l'esprit est fort.
- 1963 aux USA arrêt ou plutôt réorientation des recherches de ce style, suite à un rapport de Yehoshua Bar-Hillel.

29 nov 2007

Université Ouverte Cergy

5

Linguistique et informatique: une longue histoire (2/2)

- Aujourd'hui réparti dans les deux disciplines suivantes
 - 1960 Computational Linguistics Structuration du précédent (théories mathématiques, linguistiques)
 - 1965 Automatic / Natural Language Processing Focalisé sur les outils et plus particulièrement:
 - Analyse syntaxique
 - Méthodes statistiques
 - 1970 Natural Language Understanding (AI) approches cognitives

29 nov 2007

Université Ouverte Cergy

6

La linguistique computationnelle

vue à travers quelques outils

29 nov 2007

Université Ouverte Cergy

7

Quelques outils issus de la linguistique computationnelle (1)

- Le Graal: la traduction automatique (il faut savoir tout traiter pour y parvenir)
- Aide à la traduction:
 - domaine spécifique
 - repère les expressions idiomatiques (aller bon train)
 - propose pour chaque mot ou expression des traductions
 - les assemble avec les choix du lecteur
 - (éviter au maximum la représentation des connaissances)

29 nov 2007

Université Ouverte Cergy

8

Quelques outils issus de la linguistique computationnelle (2)

- L'interface homme/machine en langue naturelle par exemple:
 - interrogation de BD en langage naturel
Quels sont les films des années cinquante qui passent actuellement à Bordeaux?

29 nov 2007

Université Ouverte Cergy

9

Quelques outils issus de la linguistique computationnelle (3)

- Correcteurs orthographiques (pas simple):
 - Synapse Word (souligné vert: français, italien,...)
 - Quels livres crois-tu qu'il sait que je pense que tu as lus ?
- Génération automatique de bulletins météo, de comptes-rendus,..
- Résumé automatique: deux techniques contrastées

29 nov 2007

Université Ouverte Cergy

10

Quelques outils issus de la linguistique computationnelle (4)

- Recherche d'information (notamment sur Internet)
 - production laitière / production de lait
 - production minière / production de mine(s) ???
- Reconnaissance de la parole (par ex. pour sous-titrage) nécessite une analyse morpho-syntaxique pour fonctionner en temps réel

29 nov 2007

Université Ouverte Cergy

11

Linguistique computationnelle

Méthodes et objectifs

29 nov 2007

Université Ouverte Cergy

12

Un domaine interdisciplinaire

- **Mathématiques**
 - Logique et théorie des langages
 - Probabilités
- **Informatique**
 - Algorithmique
 - Génie logiciel
- **Linguistique**
 - Grammaire générative
 - Descriptions linguistiques
 - Philosophie du langage

29 nov 2007

Université Ouverte Cergy

13

Des objectifs variés

- **Réalisation d'outils de traitement des langues**
- **Formalisation des théories linguistiques**
vérification ou réfutation d'hypothèses
 - Par ex. modèles syntaxiques analysables et apprenables efficacement (en temps polynomial)
- **Développement des théories informatiques et mathématiques pour elles-mêmes,**
éventuellement pour d'autres objectifs
 - Par ex. Théorie des langages et bioinformatique

29 nov 2007

Université Ouverte Cergy

14

Un aperçu des domaines de la linguistique

Diviser l'objet d'étude en aspects plus simples

29 nov 2007

Université Ouverte Cergy

15

Voix et sons

- **Phonétique: étude des sons concrets d'une langue**
 - Acoustique
 - Système phonatoire/auditif
- **Traitement du signal / médecine**
- **Phonologie**
 - Les sons abstraits: système discret (dans un continu)
 - Bali / Paris indistincts pour un japonais
- **Théorie des langages / automates**

29 nov 2007

Université Ouverte Cergy

16

Prosodie (module transverse)

- **Structure du phrasé et de leur enchaînement: pauses, intonation**
 - "Je serai très heureux de venir parler au LaBRI, laboratoire auquel je dois ma formation initiale en informatique, par exemple sur la lambda-DRT."
 - "Je serai très heureux de venir parler au LaBRI --- laboratoire auquel je dois ma formation initiale en informatique --- par exemple sur la lambda-DRT. »
- **Systèmes d'annotation généralement superposé à d'autres informations**

29 nov 2007

Université Ouverte Cergy

17

Morphologie: structure des mots (1/2)

- **morphologie dérivationnelle: formation des mots**
 - préfixes, suffixes, nom composés, etc.
 - changement de catégorie possible
 - noble → noblesse
 - petit → petitesse
 - maison → maisonnette
 - Camion → camionnette
 - carpe → carpette ?
- **Théorie des langages / automates**

29 nov 2007

Université Ouverte Cergy

18

Morphologie: structure des mots (2/2)

- morphologie flexionnelle
déclinaisons, conjugaisons
 - en général pas de changement de catégorie (sauf exceptions, par exemple participe passé)
 - arriver → arriv[er][ons]
 - cheval → chevaux
- Théorie des langages, automates

29 nov 2007

Université Ouverte Cergy

19

Syntaxe

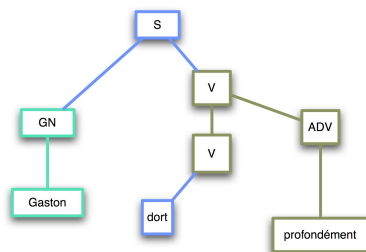
- Analyse de la structure de la phrase
 - *Je fais la réparer
 - Je la fais réparer
 - * [[Pierre [mange une]] pomme]
 - Pierre [mange [une pomme]]
- Théorie des langages de chaînes, d'arbres voire de graphes

29 nov 2007

Université Ouverte Cergy

20

Syntaxe



29 nov 2007

Université Ouverte Cergy

21

Sémantique lexicale: les sens des mots et leurs relations

- Exemple Livre:
 - livre, imprimer (objet concret),
 - lire (contenu abstrait)
 - Rôle télique: être lu, informer, cultiver,
 - Rôle constitutif: pages, couvertures etc.
 - Rôle agentif: imprimeur,...
- Logique, probabilités

29 nov 2007

Université Ouverte Cergy

22

Sémantique logique (phrase, discours, dialogue, ...)

- Deux aspects indépendants:
 - Sémantique Vériconditionnelle: (sens = formule logique et interprétations dans des mondes possibles) Le sens d'un énoncé c'est l'ensemble de ses conditions de vérité.
 - Une vache regarde le train passer.
 - Sémantique Compositionnelle (sens = formule logique ou construction abstraite) On calcule le sens d'un constituant, d'une phrase, d'un discours d'un dialogue à partir du sens de ses constituants et de sa structure (syntaxique, discursive,...)
 - J'ai oublié à l'hôtel ce livre que j'ai beaucoup aimé.
- Logique

29 nov 2007

Université Ouverte Cergy

23

Pragmatique: sens et contexte énonciatif

- Discours, dialogue, ...
- Référence des indexicaux: 1ère et 2e personnes (je, nous, vous), ici, maintenant, ce, cette, ...
 - Allons plutôt dans ce restaurant.
- En cours de formalisation, extensions des méthodes logiques en sémantique

29 nov 2007

Université Ouverte Cergy

24

La grammaire générative

Théorie incontournable
du XX^e siècle

La grammaire générative

- Théorie linguistique aussi utilisée en
 - Informatique (compilation, parallélisme)
 - Mathématiques (théorie des groupes)
 - Biologie(génomique)
- Origines
 - (Panini, Inde, V^e siècle avant J.C.)
 - Noam Chomsky 1955

Rupture avec le « behaviorisme »

- Une langue N'EST PAS l'ensemble des énoncés produits par les locuteurs
- MAIS
- Un ensemble fini de règles inconscientes qui permet de produire ces énoncés.

Rupture avec le « behaviorisme »

- Arguments:
 - Soit P la phrase la plus longue à ce jour, il croit que P est sans doute aussi une phrase.
 - Règles inconscientes:
 - le jeune enfant dit « vous faites » puis « vous faisez » puis « vous faites ».
Le « vous faisez » ne peut provenir que de règles (surgénéralisation).
 - Il a aimé trois des livres qu'Echenoz a écrit.
Il ≠ Echenoz.
 - Le chien de Paul pense qu'il ne l'aime pas.
Il ≠ tout le reste est possible.

Deux principes

- Les phrases sont analysables (compréhensibles) en temps raisonnable (traduction informatique: en temps polynomial)
- Il existe un bon algorithmes d'apprentissage de la grammaires à partir d'exemples positifs en nombre relativement faible.

Compétence / Performance

- Les règles / Nos limites (mémoire)
 - Le loup a dévoré la chèvre.
 - La chèvre que le loup a dévoré avait mangé le chou.
 - ? Le chou que la chèvre que le loup a dévoré avait mangé appartenait au passeur.
 - ?? Le passeur auquel le chou que la chèvre que le loup a dévoré avait mangé appartenait possédait plusieurs bateaux.
 - ??? Les bateaux que le passeur auquel le chou que la chèvre que le loup a dévoré avait mangé appartenait possédait sont des barges.
- (néanmoins correct, en prenant son temps et un crayon)

Quels langages formels pour la syntaxe du langage naturel

Seulement le principe de complexité de l'analyse et d'adéquation descriptive

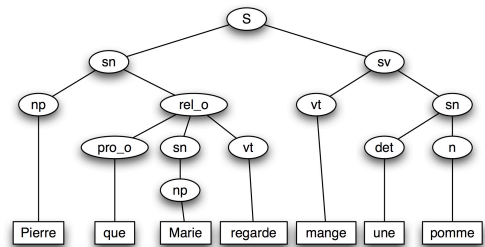
Quelles règles modélisent la compétence?

- T terminaux (mots), N non terminaux
- Règles $W \rightarrow W'$ (W : au moins un N)
 - $W=W_1 Z W_2$ and $W'=W_1 W'' W_2$ contextuelles
 - $|W'| \geq |W|$ croissantes
 - $|W|=1$ non contextuelles
 - $|W|=1$ et $W'=mZ$ régulières

Exemple de grammaire hors-contexte

- $s \rightarrow sn \ sv$
- $sn \rightarrow det \ n \ | \ np \ | \ det \ n \ rel_s \ | \ det \ n \ rel_o \ | \ rel_s \ | \ n \ p \ rel_o$ np
- $rel_s \rightarrow pro_s \ sv$
- $rel_o \rightarrow pro_o \ sn \ vt$
- $sv \rightarrow vi \ | \ vt \ sn$
- $pro_o \rightarrow que$
- $pro_s \rightarrow qui$
- $vt \rightarrow regard \ e \ | \ regardent \ | \ mange \ | \ mangent$
- $vi \rightarrow dort \ | \ dorment \ | \ tombe \ | \ tombent$
- $det \rightarrow une \ | \ un \ | \ la \ | \ le \ | \ des \ | \ les$
- $n \rightarrow pommes \ | \ pomme \ | \ femme \ | \ femmes$
- $np \rightarrow pierre \ | \ marie$

Exemple de dérivation



Quel type de règles?

- les langages réguliers ne suffisent pas:
 - (ex. précédent relative avec « que »)
 - Sujet1 Sujet2 Sujet3 ...
Verbe3 Verbe2 Verbe1

Quel type de règles?

- les langages hors-contexte non plus:
 - (complétives NL)
Sujet1 Sujet2 Sujet3 ...
Verbe1 Verbe2 Verbe3
 - ...dat ik₁ Henk₂ haar₃ de nijlpaarden₃ zag₁
helpen₂ voeren₃
... que je₁ vois₁ Henk₂ l₃ aider₂ à nourrir₃
les hippopotames

Quel type de règles?

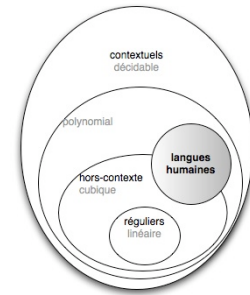
- Un peu plus complexes que hors-contexte, mais avec analyse polynomiale :
 - TAG ou grammaires hors-contexte avec mouvements

29 nov 2007

Université Ouverte Cergy

37

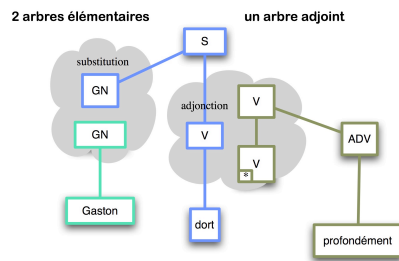
Hierarchie des langages formels



29 nov 2007

38

Grammaires d'arbres adjoints

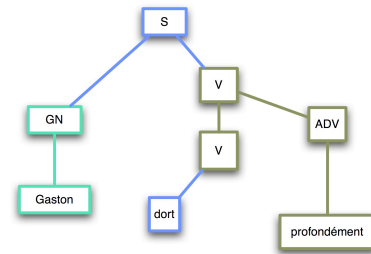


29 nov 2007

Université Ouverte Cergy

39

Grammaires d'arbres adjoints

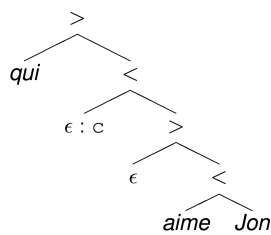


29 nov 2007

Université Ouverte Cergy

40

Arbres minimalistes



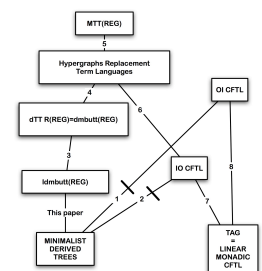
29 nov 2007

Université Ouverte Cergy

41

Hierarchie des langages d'arbres

- Sujet de recherche actuel (Los-Angeles, Berlin, Bordeaux, ...)



29 nov 2007

Université Ouverte Cergy

42

Gros plan sur un analyseur à large échelle GRAIL (Richard Moot)

Extraction automatique de la grammaire
Association mots entrées lexicales
Analyses des séquences les plus probables

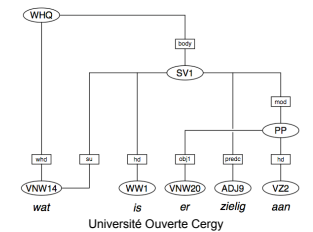
29 nov 2007

Université Ouverte Cergy

43

Un gros plan: Grail (Moot) 1/4

- Corpus de néerlandais parlé, transcrit et annoté (dépendances, constituants) [origine: NWO]



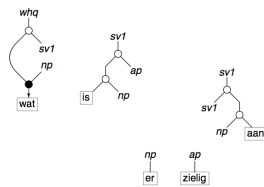
29 nov 2007

Université Ouverte Cergy

44

Un gros plan: Grail (Moot) 2/4

- Extraction automatique d'arbres lexicaux (techniquement: formules de la logique multimodale non associative)



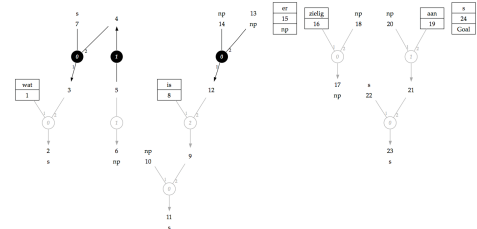
29 nov 2007

Université Ouverte Cergy

45

Un gros plan: Grail (Moot) 3/4

- Supertagging: étiquetage le plus probable de la suite de mots par des arbres (environ 100 arbres par mot...)



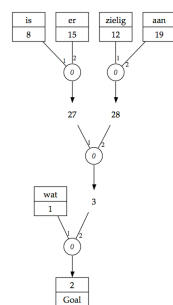
29 nov 2007

Université Ouverte Cergy

46

Un gros plan: Grail (Moot) 4/4

- Analyse de la phrase
 - Analyse avec chacune des n suites d'arbres les plus probables
 - Minimisation de la somme des distances des liens qui établissent la consommation des traits grammaticaux.



29 nov 2007

Université Ouverte Cergy

47

Etat de l'art en pratique

- Richard Moot MMCG: extraction, parsing
 - NWO Dutch Spoken Corpus (conversation spontanée, transcription annotée)
 - 1.002.098 mots
 - 114.801 phrases (7,6 mots par phrase)
 - 44.306 mots différents
 - Multi-Modal Categorical Grammar, extraite automatiquement (en moyenne 100 arbres pas mot!)
 - Supertagging (les n plus probables suites d'assignations d'arbres aux mots de la phrase)
 - Résultats sur le corpus de test 19.237 phrases, 146.497 mots (supertagging >> parsing):
 - 1 supertag 2'53" 40% d'analyses correctes (9 ms/phraser 1.18 ms/mot)
 - 10 meilleures supertags 48'34" 70% d'analyses correctes (151ms/phraser, 20ms/mot)

29 nov 2007

Université Ouverte Cergy

48

Etat de l'art en pratique

- Benoît Sagot, Eric de la Clergerie LFG parsing
 - Corpus EASy (Evaluation des Analyseurs Syntactiques)
Journaux, web, mail, discours politiques, littérature, ...
 - 87177 mots
 - 4322 phrases (20,2 mots par phrase)
 - Grammaire LFG écrite "à la main"
 - Choisit une analyse par phrase
 - temps d'analyse: total 152s, 35ms/phrased 1,7ms/mot
 - Tronçons (chunks) corrects: 86%
 - Relations correctes entre tronçons (chunks): 49%
 - Analyses correctes en nombre inconnu (pas de corpus de test)

29 nov 2007

Université Ouverte Cergy

49

Analyseurs difficilement comparables

- | | |
|---|---|
| 1. Ecrit | 1. Parlé |
| 2. Phrases assez longues, une vingtaine de mots | 2. Phrases courtes mais tordues |
| 3. Annotations élémentaires | 3. Corpus bien annoté |
| 4. Grammaire écrite | 4. Grammaire acquise automatiquement |
| 5. Lexical Functional Grammar | 5. MultiModal Categorical Grammar |
| 6. Mesure de correction: % tronçons et relation | 6. Mesure de correction: % analyses complètes correctes |

29 nov 2007

Université Ouverte Cergy

50

Acquisition de la syntaxe et grammaire universelle

Après l'efficacité de l'analyse,
Deuxième guide de la grammaire
généralive.

29 nov 2007

Université Ouverte Cergy

51

Faits connus sur le jeune enfant 18-36 mois

- Exemples positifs seulement
- Exemples assez peu nombreux par opposition à complexité d'une langue naturelle
- Sens des mots connu au préalable
- Intonation utile

29 nov 2007

Université Ouverte Cergy

52

L'hypothèse de la grammaire universelle

- Pas une grammaire au sens usuel mais des contraintes sur la forme des grammaires des langues humaines
- Avec cette hypothèse le processus d'acquisition devient explicable
- Apprentissage par choix de paramètres
- Exemple bête SVO ou SOV?
un exemple suffit:
 - maman conduit la voiture

29 nov 2007

Université Ouverte Cergy

53

Quelques principes de la grammaire universelle

- Tout groupe nominal doit recevoir un cas,
- et seul un verbe conjugué donne un cas.
 - Il semble que l'été arrive.
 - L'été semble arriver.
 - * Il semble (que) l'été arriver.
- Un pronom doit être gouverné par son antécédent (position relative dans l'arbre d'analyse)
 - * Il_i a aimé deux livres que Chomsky_j a écrit.
 - Combien de livres que Chomsky_i a écrit a-t-il_i aimés?

29 nov 2007

Université Ouverte Cergy

54

Modèle de Gold: classe apprenable

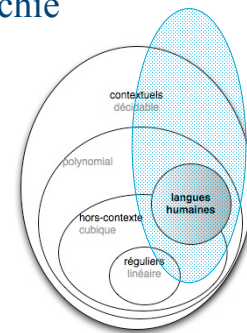
- Algorithme
Phrase₁... Phrase_n → Grammaire G_n
G_n engendre Phrase₁... Phrase_n
- Si la totalité des Phrase₁.....
énumèrent un langage L de la classe
- ALORS à partir d'un nombre fini n
d'exemples, l'hypothèse faite G_n
ne varie plus et le langage engendré par G_n
est L.

29 nov 2007

Université Ouverte Cergy

55

Langages apprenables et hiérarchie



29 nov 2007

56

Limites de l'approche à la Gold

- (Pratique) Nécessite des exemples
avec beaucoup d'informations, très
structurés: données disponibles?
- (Théorique) Les algorithmes, dits par
généralisation, font grossir le langage
jusqu'à la cible. (Sauf I. Tellier U. Lille3),
à partir de représentations sémantiques:
apprentissage par spécialisation)

29 nov 2007

Université Ouverte Cergy

57

Linguistique et logique

Une longue tradition
Aujourd'hui: l'analyse du sens
devient (en partie) calculable.

29 nov 2007

Université Ouverte Cergy

58

Comprendre les phrases analysées

- Arbres, structures calculées
automatiquement, mais ensuite,
comment leur associer des formules
manipulables par une machine?
- Quels problèmes rencontre-t-on?
Ils sont très classiques
... mais pas faciles.

29 nov 2007

Université Ouverte Cergy

59

Logique et grammaire: un lien naturel et traditionnel

- Depuis l'antiquité (Aristote, Denis de Thrace)
- puis au Moyen-Âge (scholastique),
- au 18e (Port-Royal)...
- La phrase a une structure logique,
importante en pratique.
 - Les enfants prendront une pizza
 - Chaque en prend une pizza pour lui?
 - Ils partagent la même?

29 nov 2007

Université Ouverte Cergy

60

Les langues naturelles sont logiquement (trop?) riches

- Un, des certains, Tous, tous, les chaque
- + d'autres quantificateurs: *la plupart, les, un grand nombre de, un petit nombre de,*
 - La plupart des politiciens ont lu un livre d'économie.
- Les nombres sont aussi des sortes de quantificateurs:
 - Mettre huit gouttes dans trois cuillères à soupe d'eau.
 - $3 \times 8 = 24$ gouttes?
 - 8 gouttes?

29 nov 2007

Université Ouverte Cergy

61

Problèmes de portées, suite

- Lectures de re et de dicto
 - James Bond croit que l'un des chercheurs du laboratoire est un espion.
 - James Bond pense que Blofeld est un espion.
 - Il existe un espion x et JB croit que x fait partie du laboratoire.
 - James Bond a trouvé un microfilm dans le laboratoire.
 - JB croit qu'il existe un espion dans le laboratoire, mais il ne soupçonne personne en particulier.

29 nov 2007

Université Ouverte Cergy

62

Interprétation, mondes possibles, intentionnalité

- Sens d'un énoncé (vériconditionnel) classe des mondes possibles dans lesquels il est vrai.
 - Cet étudiant croit que Chomsky est informaticien.
 - Dans tous les mondes possibles compatibles avec les croyances de cet étudiant, Chomsky est informaticien.

29 nov 2007

Université Ouverte Cergy

63

Une logique d'ordre supérieur

- Propriétés de propriétés
 - Une fraise rouge vermillon
 - Rouge propriété, vermillon propriété de propriétés
 - tous les médecins sont des conducteurs
 - (donc) tous les médecins bordelais sont des conducteurs bordelais
 - *(donc) tous les bons médecins sont des bons conducteurs
 - Bordelais, médecin: propriétés
 - Bon: propriété de propriété, transformateur de propriété

29 nov 2007

Université Ouverte Cergy

64

Compositionnalité

- Frege le sens du tout est construit à partir du sens des parties.
 - Les étudiants reçus sont partis fêter ça.
- Limites de la compositionnalité:
 - Si un paysan possède un âne, alors il (=le paysan) le (=l'âne) bât.

29 nov 2007

Université Ouverte Cergy

65

Attention à ne pas s'écarter de ce qui est dit

- J'avais trois trombones dans ma poche, je les ai tous perdus sauf un.
Je le range dans un tiroir.
- J'avais trois trombones dans ma poche, j'en ai perdu deux.
* Je le range dans un tiroir.
- Pourtant, d'un point de vue purement logique, la situation est identique.

29 nov 2007

Université Ouverte Cergy

66

Lien avec la syntaxe

- Aspects sémantiques des catégories syntaxiques
- Les catégories ou parties du discours ont une contre partie logique.

29 nov 2007

Université Ouverte Cergy

67

Lien avec la syntaxe

- Groupes nominaux: individus (individus ou variables d'individus quantifiables) E
- Verbes, groupes verbaux: prédicats
dort: fonction de E dans T $E \rightarrow T$
regarde de E dans E dans T: $E \rightarrow E \rightarrow T$
- Adjectifs partage les caractères avec les noms (accord, déclinaisons) et avec les verbes (expriment un prédicat) plutôt $E \rightarrow T$
- Groupes prépositionnels :
ni des prédicats, ni des individus
– Modificateur de prédicat:
sur un banc ($E \rightarrow T$) \rightarrow ($E \rightarrow T$)
dort \rightarrow dort sur un banc

29 nov 2007

Université Ouverte Cergy

68

Grammaires catégorielles

Mot	Catégorie Syntaxique	Type Sémantique
Pierre	SN	E
Dort	SNIS	$E \rightarrow T$

Pierre dort	S	T	Dort(Pierre)
-------------	---	---	--------------

29 nov 2007

Université Ouverte Cergy

69

Langue des Signes Française LSF et organisation syntaxique Emilie VOISIN, Patrick HENRY

Etude de terrain et formalisation
Projet Région Aquitaine

29 nov 2007

Université Ouverte Cergy

70

Problématique

- Étude des unités discursives :
 - Ordre des constituants
 - Notion d'accord et de flexion verbale: le verbe

29 nov 2007

Université Ouverte Cergy

71

Cheminement de l'étude

- Recueil de corpus
 - locuteurs sourds signants ;
 - 5 locuteurs différents ;
 - matériel vidéo HD de la région Aquitaine
- Préalable significatif: différence LSF / Français Signé
- Examen des contraintes liées à l'ordre des signes

29 nov 2007

Université Ouverte Cergy

72

Ordre des signes

- Types de phrases étudiées :
 - Phrases simples (dans un premier temps) ;
 - Verbes soumis à la variation ;
 - Possibilité de prise de rôle (transferts personnel et situationnel)

29 nov 2007

Université Ouverte Cergy

73

Première étape

- Identifier ce qui est susceptible de faire varier l'ordre de la phrase simple

29 nov 2007

Université Ouverte Cergy

74

Ordre des signes et influence des transferts

- Transfert situationnel (TS) :
 - La situation est ramenée sur le plan de l'espace de signation
- Transfert personnel (TP)
 - Prise de rôle : le signeur devient l'agent qui fait l'action

29 nov 2007

Université Ouverte Cergy

75

Exemples de transferts

- Transfert situationnel (TS)

CHAT (loc A) SOURIS (loc B)
(loc A)-MANGER-(loc B-class_animaux)

- Transfert personnel (TP)

SOURIS (loc A) CHAT (TP)
(je)-MANGER-(loc A)

29 nov 2007

Université Ouverte Cergy

76

Ordre des signes et principe d'économie

- Mise en évidence d'un principe d'économie :
- Reflète l'iconicité de l'énoncé grâce à la fluidité entre les concepts lorsque la proforme (configuration manuelle reprenant la forme choisie pour représenter un objet donné) intégrée au verbe a la même forme que le nom
- Le/un garçon mange une pomme:
- + garçon pomme mange
 - pomme garçon mange

Conséquences : reprise de la proforme
→ Remarque : ce principe contraint fortement l'ordre des signes

29 nov 2007

Université Ouverte Cergy

77

Deuxième étape :

- Quelques remarques sur le fonctionnement du verbe

29 nov 2007

Université Ouverte Cergy

78

La notion d'accord en LSF

- Définition générale :
 - Le verbe porte des indices actanciels pouvant être apparentés à des affixes ;
 - Coréférence de l'affixe avec le terme nominal ;
 - Le verbe s'accorde avec ce terme nominal

29 nov 2007

Université Ouverte Cergy

79

L'accord en proforme

- Proforme : une proforme est un classificateur (configuration manuelle utilisée pour reprendre la forme de l'objet - classificateur de forme)
 - Ballon (lancer)
 - Ballon (écusson dessus)
- Variation de la configuration

VOITURE (conf main plate)
SE_DÉPLACER (conf main plate)

29 nov 2007

Université Ouverte Cergy

80

L'accord en loci

- Variation des lieux de réalisation (dans l'espace de signation) :

(1) FILLE-point(loc A) JEAN-point(loc B) TP
DESIRER(loc B)

29 nov 2007

Université Ouverte Cergy

81

« Double » accord

- Accord en proforme et en loci

ZOE(loc A) TOM(loc B) LIVRE(conf-bco)
(loc A)-DONNER(conf-bco)-(loc B)

29 nov 2007

Université Ouverte Cergy

82

Troisième étape :

- Comment caractériser cet accord verbal?

29 nov 2007

Université Ouverte Cergy

83

La coalescence(1)

- Fusion plus ou moins poussée du verbe et d'un nom ;
- Peut aller jusqu'à l'incorporation (intégration du nom à l'intérieur même de la forme verbale)

29 nov 2007

Université Ouverte Cergy

84

La coalescence (2)

- Le nom perd sa fonction actancielle ;
- La relation sémantique entre le verbe et le nom résulte de la confrontation de leur contenu respectif (possibilité d'incorporation de termes à valeur instrumentale par exemple: écrire)

29 nov 2007

Université Ouverte Cergy

85

Perspectives

- Examen des formes verbales fléchies : incorporation nominale?
- Comparaison avec des langues présentant ce type de phénomènes (notamment langues amérindiennes)

29 nov 2007

Université Ouverte Cergy

86

Incorporation nominale

- Mithun - à propos de la distribution de l'incorporation nominale (1984) :
 - « *If a language incorporates nouns of only one semantic case, they will be patients of transitive verbs (...) if a language incorporates only two types of arguments, they will be patient of transitive and intransitive verbs (...) many languages additionally incorporate **instrument** and/or location* »
- *Mais en LSF verbe et nom incorporé peuvent être séparés par un déterminant démonstratif, ce qui n'est pas décrit dans ces langues.*

29 nov 2007

Université Ouverte Cergy

87

Perspectives et hypothèses

→ Que dire de l'idée selon laquelle,
(par exemple) :

MANGER_POMME ; MANGER_SANDWICH
ou
MANGER_générique ; MANGER_OISEAU (> PICORER)

seraient des verbes totalement différents les uns des autres...?

29 nov 2007

Université Ouverte Cergy

88

Conclusion

de cet aperçu morcelé

29 nov 2007

Université Ouverte Cergy

89

Bilan

- Point de vue théorique, cognitif: des principes régissant l'usage le langage.
- Modèles mathématiques, algorithmes adaptés: réalisations informatiques possibles
- Les connexions entre aspects du langages sont plus difficiles à comprendre et à modéliser.

29 nov 2007

Université Ouverte Cergy

90



Bilan

- Les modèles utilisés aident l'étude de terrain des langues moins connues (LSF).
- Réciproquement l'étude de langues a priori éloignées permet d'enrichir les modèles (d'informations discursive, spatiales,...)
- On manque de ressources linguistiques corpus grammaires dictionnaires: données chères et guère uniformisées.

29 nov 2007

Université Ouverte Cergy

91



Perspectives

- Pratiques: constitution et interopérabilité des ressources, formalismes de haut-niveau
- Théoriques et pratiques: développement des aspects sémantiques
- Théorique: que cela dit-il de nos capacités cognitives (par ex. Quelle logique utilisons nous et comment?)

29 nov 2007

Université Ouverte Cergy

92



Bibliographie

- The language Instinct de Steven Pinker traduit aux éditions Odile Jacob L'instinct de langage
- Cori M et Léon J, "La constitution du TAL. Etude historique des dénominations et des concepts", *Traitement Automatique des Langues*, n° 43-3 :21-55. 2002.

29 nov 2007

Université Ouverte Cergy

93



Bibliographie

- Syntaxe et Traitement Automatique des Langues. 2005 in F. Lambert et H. Nolke (éds) *La syntaxe au cœur de la grammaire, mélanges offerts à Claude Muller*, pages 271-286. Presses universitaires de Rennes
- Les mathématiques de la linguistique computationnelle. Premier volet: la théorie des langages. A paraître dans *la Gazette des mathématiciens* (Société Mathématique de France).
- Modèles algorithmiques de l'acquisition de la syntaxe : concepts et méthodes, résultats et problèmes (avec D. Bechet, R. Bonato, A. Dikovsky, A. Foret, Y. Le Nir, E. Moreau, I. Tellier), *Recherches Linguistiques de Vincennes*. volume 36, pages 123--152. Octobre 2007.

29 nov 2007

Université Ouverte Cergy

94