

Introduction aux communications palliatives

2 la linguistique

plan

- Écrire les mots : les alphabets
- Ranger les mots : Les dictionnaires
- Construire des phrases : la grammaire
- Décire les mots : la Sémantique

Des phrases de la langue ou bien un puzzle



Alphabet

- Une base de signes et une loi de composition permettant de créer des mots
- Que représentent les signes de la base?
 - des sons?
 - Des concepts?
 - Des idées?
 - Des choses?

Codage des caractères à partir de l'alphabet latin

L'alphabet : encoder les sons

Abcdefghijklmnopqrstuvwxyz
 ABCDEFGHIJKLMNOPQRSTUVWXYZ
 1234567890
 ? , ; : !
 - Diacritiques

- Cet ensemble ne convient pas à toutes les langues

Alphabet arabe : encoder les consonnes

Encoder / représenter les choses

figurer/ symboliser

10	mdj	C n
100	ch(r)l	e
1.000	kha	s
10.000	qjba	l
100.000	h'm	h
1.000.000	h'h'	h'
10.000.000	?	o
1/2	gs	A n
1/3	r(a), wy	o x
1/4	h'sb	o x
1/8		o x
1/16		o x
1/32		o x
1/64		o x
3/4	khw-(a)	o x

Le système d'écriture Bliss

encoder des concepts

Les pictogrammes

Figuratifs?
Symboliques?
Conceptuels?

1-févr.-05

Abraham/LaLICC Brest

10

Les pictogrammes

- Un mot \leftrightarrow une chose ?
- Mais ...
- Les mots ne désignent pas que des choses; il y a
 - Les actions
 - Les déterminations
 - Les déictiques
 - Les adverbes
- comment représenter tout cela?

1-févr.-05

Abraham/LaLICC Brest

11

plan

- Écrire les mots : les alphabets
- Décrire les mots : la Sémantique
- Ranger les mots : Les dictionnaires
- Construire des phrases : la grammaire

1-févr.-05

Abraham/LaLICC Brest

12

Les mots

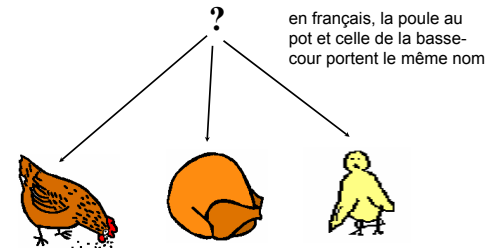
- Que représentent-ils?
- Un mot, une chose? – la polysémie
- Comment les catégoriser?
- Catégories syntaxiques
- Catégories sémantiques
- Comment les utiliser?

1-févr.-05

Abraham/LaLICC Brest

13

le lexique : la polysémie nominale



1-févr.-05

Abraham/LaLICC Brest

14

quel est le statut des pictogrammes?

la suite



[Noël] [manger] [le bonbon]

devra produire la phrase « Noël mange le bonbon »,
et non

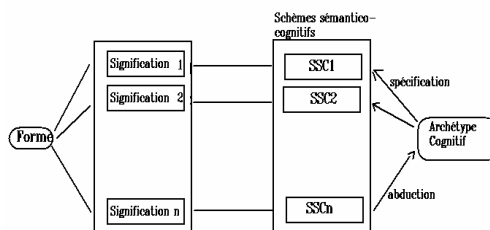
« (A) Noël, (je) mange le bonbon ».

1-févr.-05

Abraham/LaLICC Brest

15

Polysémie verbale



Polysémie verbale

A une même forme d'entrée dans le dictionnaire, correspondent plusieurs significations que l'on décrit par des SSC

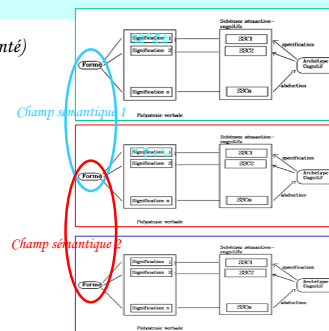
1-févr.-05

Abraham/LaLICC Brest

16

Polysémie, synonymie

Champ du déplacement (orienté)
Monter, grimper, gravir



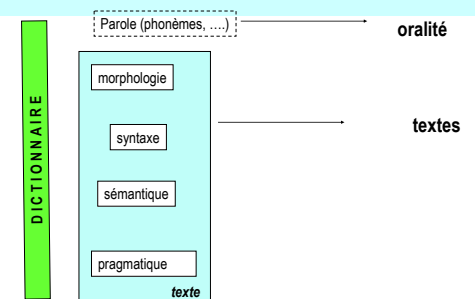
Champ de la construction
Assembler, construire, bâtir, édifier,

1-févr.-05

Abraham/LaLICC Brest

17

Modèle



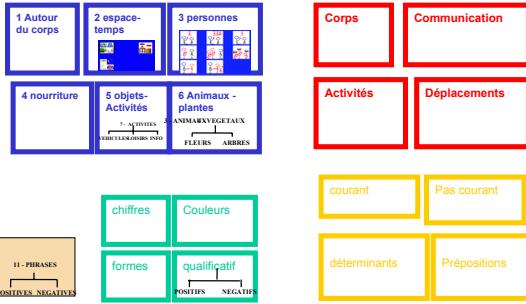
1-févr.-05

Abraham/LaLICC Brest

18

le DICTIONNAIRE

sélection par balayage



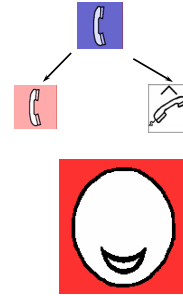
1-févr.-05

Abraham/LaLICC Brest

19

lesMDS : générativité de la langue

- Quelques entités concrètes
- Téléphone -> téléphoner
- Verbes de perception : les parties du corps servent de primitives dans les verbes de perception, ex:
- Bouche -> parler
- Gueule ->



1-févr.-05

Abraham/LaLICC Brest

20

syntaxe d'un langage

- c'est l'ensemble des règles de construction de ce langage
- analyse syntaxique:
 - l'ordinateur lit dans un certain format la suite de caractères composant une phrase; celle-ci doit d'abord être découpée en une suite de mots du langage : ce travail est effectué par un analyseur morpho-lexical
 - la suite des mots accompagnés de leur catégorie est ensuite fournie à l'analyseur syntaxique qui va tenter de regrouper ces mots en phrases, conformément aux règles syntaxiques

1-févr.-05

Abraham/LaLICC Brest

21

morpho-syntaxe

- règles morphologiques et règles syntaxiques sont exprimées par une grammaire (algébrique)
- les analyseurs morpho-syntaxiques effectuent la reconnaissance d'une suite de symboles sans se préoccuper de la ou des significations possibles de ces symboles (*le piano écoute*)

1-févr.-05

Abraham/LaLICC Brest

22

La phrase et la linguistique générative

Chomsky publie «structures syntaxiques» en 1957 "tout est syntaxique "

- idée : on apprend une langue, non pas en mémorisant toutes les phrases possibles entendues, mais en ayant une sorte de connaissance innée de la structure de la langue.
- on s'en aperçoit par les erreurs que font les enfants en formant trop de régularités ex : *il a pleuvu, quarante-dix*
- Chomsky tente de décrire cette aptitude à la langue en développant une série de grammaires génératives
- le type de grammaire que Chomsky décrit ne définit pas ce qui est correct ou incorrect dans l'usage de la langue, mais plutôt ce qui produit des relations structurelles acceptables
- (Winograd, p.12) la linguistique générative voit le langage comme un objet mathématique, et construit des théories qui sont tout à fait semblables aux ensembles d'axiomes et aux règles d'inférence en mathématique

1-févr.-05

Abraham/LaLICC Brest

23

grammaires et automates

- **les grammaires sont des systèmes formels générateurs**: les expressions bien formées, ou phrases du langage sont obtenues (engendrées) à partir d'un symbole initial en appliquant un ensemble de productions ou règles de formation.
- inversement **les automates sont des systèmes formels accepteurs** : Ils partent d'une phrase, et les règles vont vérifier si cette phrase appartient ou non à un langage donné.
- **Un automate est alors équivalent à une grammaire s'il accepte exactement le langage engendré par la grammaire.**

1-févr.-05

Abraham/LaLICC Brest

24

grammaires formelles:

mise en évidence de la structure syntaxique

- une grammaire formelle est un quadruplet $G = \{V_t, V_n, P, S\}$
- - le vocabulaire: $V = V_n \cup V_t$
- - réécriture P ensemble des règles de production

$S \rightarrow GN GV$

.....

- S est l'axiome ; P l'ensemble de règles de production; V_n le vocabulaire non terminal; V_t le vocabulaire terminal

1-févr.-05

Abraham/LaLICC Brest

25

A*

- Soit A un ensemble quelconque nommé alphabet
- L'ensemble A peut être constitué de caractères ou de mots pris dans un dictionnaire donné ou encore d'objets plus abstraits comme des classes grammaticales. On appellera lettres les éléments de A indépendamment de l'interprétation; on forme des suites finies $a_1 a_2 a_3 \dots$ an d'éléments que l'on appelle des mots
- On note A^* l'ensemble de tous les mots sur l'alphabet A, e la suite vide
- On nomme langage formel tout ensemble de mots sur l'alphabet A
- Un automate sur l'alphabet A ou A-automate est un graphe dont les flèches sont étiquetées par des mots sur l'alphabet A. les sommets du graphe s'appellent des états de l'automate
- On se donne de plus 2 ensembles, distingués de l'ensemble Q des états: un ensemble I d'états initiaux et un ensemble T d'états terminaux.
- Un mot est dit accepté ou reconnu par l'automate s'il est l'étiquette d'une chemin allant de I à T.

1-févr.-05

Abraham/LaLICC Brest

26

les grammaires formelles régulières : type 3

- $a^m b^p$
- grammaire régulière (à gauche ou à droite), ou d'état fini (engendrée par AEF), ex:
 - $X \rightarrow a Y$ ou $X \rightarrow a$ (gauche)
 - $X \rightarrow Y a$ ou $X \rightarrow a$ (droite)
- markovien: un mot dépend uniquement du précédent.
 - 1 $X \rightarrow a X$
 - 1' $X \rightarrow a Y$
 - 2 $Y \rightarrow b Y$
 - 2' $Y \rightarrow b$
- ex: $a^m b^p$ est engendré par $x \rightarrow_1 a X \rightarrow_1 a^2 X \rightarrow_1 a^3 Y \rightarrow_1 a^3 b Y \rightarrow_1 a^3 b^2$

1-févr.-05

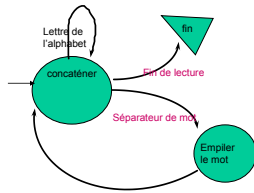
Abraham/LaLICC Brest

27

exemple : lecture d'un mot (AEF)

Cas :

- <lettre de l'alphabet> alors concaténer
- <séparateur de mot> alors empiler le mot
- <Fin de lecture> alors arrêt



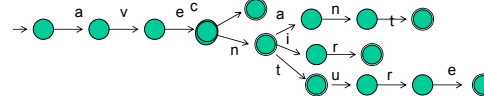
1-févr.-05

Abraham/LaLICC Brest

28

Langages formels

- Un langage formel X est dit reconnaissable par automate fini s'il existe un automate fini qui reconnaît les mots de X et eux seulement
- Un automate fini est déterministe si
 - Il a un seul état initial
 - L'étiquette de chaque flèche est une lettre
 - De chaque état sort au plus une flèche d'étiquette donnée
- Tout automate fini peut être remplacé par un automate déterministe



1-févr.-05

Abraham/LaLICC Brest

29

la phrase : fonction syntaxique?

- Le verbe attend des arguments qui vont jouer un rôle dans la sémantique de la phrase
 - sujet, verbe, ... sujet, attribut, cod, ccirc, ..
 - ou : groupe sujet, groupe verbal ... (fonction syntaxique attribuées à un groupe nominal, groupe verbal, ...)
- Le rôle interprétatif d'un paramètre est inséré dans un contexte, et non pas dans une ontologie
 - Une même entité peut avoir les rôles d'agent, de patient, d'instrument ...

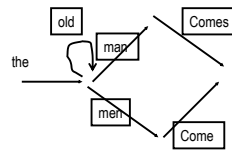
1-févr.-05

Abraham/LaLICC Brest

30

Engendrer la langue par un AEF?

- AEF qui produit 2 phrases :
- The man comes, ..
- On augmente :
- The old man....
- The old old man



- Il est impossible de construire une machine du type AEF qui produirait uniquement les phrases correctes de l'anglais

1-févr.-05

Abraham/LaLICC Brest

31

ces grammaires sont insuffisantes pour la langue naturelle:

- AEF: au cours de la génération, la seule information nécessaire pour terminer la phrase correctement est la connaissance de l'état dans lequel on se trouve.
- la connaissance de tout ce qui précède est inutile pour trouver une continuation correcte de la phrase.
- Cette caractéristique montre clairement que ce type de grammaire est insuffisant pour traiter les divers aspects des langues, en particulier les structures imbriquées ou parenthésées: (.....)

1-févr.-05

Abraham/LaLICC Brest

32

les grammaires formelles non-contextuelles: type 2

- non-contextuelle, hors contexte, (context-free) $a^n b^n$
- côté gauche de la règle: un non-terminal
- pas de contrainte sur le membre droit.
- structure d'emboîtement récursif
- définition formelle
 - Une grammaire context-free est un quadruplet
 - $G = \{V_t, V_n, P, S\}$ où
 - V_t est le vocabulaire terminal,
 - V_n est le vocabulaire non terminal,
 - P est l'ensemble des règles de dérivation,
 - $S \in V_n$ est l'axiome.

1-févr.-05

Abraham/LaLICC Brest

33

grammaires de type 2

- Les règles de dérivation sont de la forme:

$A \rightarrow \alpha$ où

$\square \alpha, A \in V_n$

- α est une suite non vide d'éléments de $V_t \cup V_n$

ex : $S \rightarrow X$

$X \rightarrow a X b \mid a b$

on ne sait pas décider si les langues naturelles sont ou non context-free; ces grammaires sont malgré tout suffisantes pour des applications restreintes

1-févr.-05

Abraham/LaLICC Brest

34

Gram. type 2: analyse syntaxique

grammaire

génération :

$S \rightarrow N1 V$

$N1 \rightarrow \text{Art nom}$

$V \rightarrow \text{manger}$

$\text{Art} \rightarrow \text{le}$

$\text{nom} \rightarrow \text{chat}$

automate

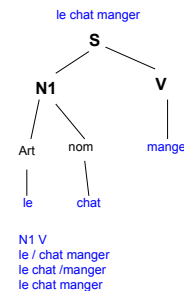
reconnaissance

$\text{Art} \rightarrow \text{le}$

$\text{nom} \rightarrow \text{chat}$

$V \rightarrow \text{manger}$

$S \rightarrow N1 V$



$N1 V$
le / chat manger
le chat / manger
le chat manger

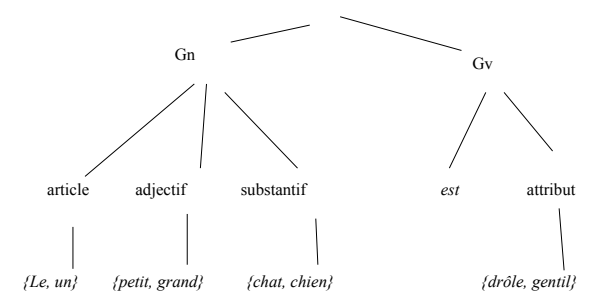
1-févr.-05

Abraham/LaLICC Brest

35

exemple gr.c-free

PHRASE



1-févr.-05

Abraham/LaLICC Brest

36

analyseur syntaxique

- Phrase -> Gn Gv
- Gv -> est attribut
- Gn -> article adjectif substantif
- article -> le | un
- adjectif -> petit | grand
- attribut -> drôle | gentil

base de règles:
 PHRASE (x,y) -> GN(x) GV(y);
 GV("est" y) -> attribut(y);
 GN(x,y,z,nil) -> article(x) adjectif(y) substantif(z);

base de données:
 article("le") ->;
 article("un") ->;
 adjectif("petit") ->;
 adjectif("grand") ->;
 substantif("chat") ->;
 substantif("chien") ->;
 attribut("drôle") ->;
 attribut("gentil") ->;

1-févr.-05

Abraham/LaLICC Brest

37

les grammaires formelles : type 1

- sensible au contexte
- $a^n b^n c^n$
- contrainte de chevauchement, ex J, P et M sont respectivement les voisins de Je, Pe et Me
- plus d'un symbole à gauche et plus de symboles à droite qu'à gauche.
- $U X V \rightarrow U Y V$, ex :
 $P \rightarrow a B C / a P B C$
 $BC \rightarrow bC$
 $C B \rightarrow B C$
 $b B \rightarrow bb$
 $c C \rightarrow cc$
 $a B \rightarrow a b$
 $B c \rightarrow b c$
 $C \rightarrow c$

1-févr.-05

Abraham/LaLICC Brest

38

définition d'un langage décidable:

- s'il existe un algorithme tel que pour tout élément x de V^* on peut savoir au bout d'un temps fini si x appartient ou non au langage.
- V^* monoïde libre construit sur V ensemble des chaînes finies

1-févr.-05

Abraham/LaLICC Brest

39

définition d'un langage positivement semi-décidable:

- s'il existe un algorithme tel que pour tout élément x de V^* qui appartient au langage l'algorithme l'indique au bout d'un temps fini;
- si x n'appartient pas, l'algorithme ne conclut pas.

1-févr.-05

Abraham/LaLICC Brest

40

une grammaire de type 1 engendre un langage décidable

- la forme des règles implique que le nombre de symboles de la chaîne engendrée ne peut que croître.
- Etant donnée une chaîne de longueur n, si on ne l'a pas rencontrée après avoir engendré toutes les chaînes de longueur n, on est sûr qu'on ne la rencontrera plus et donc que la chaîne n'appartient pas au langage.

1-févr.-05

Abraham/LaLICC Brest

41

les grammaires formelles : type 0

- aucune contrainte sur la forme des règles de réécriture
- langage récursivement énumérable, reconnu par une machine de Turing

1-févr.-05

Abraham/LaLICC Brest

42

Qu'est-ce qu'une Machine de Turing ?

- Une machine de Turing se résume à une tête de lecture comportant un nombre fini d'états internes et à un ruban. La puissance de l'analyse de Turing (1912-1954) tient au fait que sa tête de lecture ne lit qu'un seul symbole à la fois, et que cette lecture, associée à la table d'états adéquate, suffit à effectuer toutes les opérations possibles.
- La Machine de Turing est toutefois une machine imaginaire, abstraite, et idéale. Elle n'a pu être construite. Son fonctionnement implique en effet d'avoir un ruban extensible à volonté donc infini! La combinaison de cette mémoire infinie et d'un nombre d'états fini a cependant apporté la preuve que, si un problème est calculable, alors il existe une machine pour le résoudre. Dès lors, une MACHINE, reprenant les règles de toutes les autres, est UNIVERSELLE.
- Grâce à cette idéalisation mécanique Turing a montré la faisabilité d'un automate programmable capable de calculer toute fonction calculable. Ainsi, les ordinateurs sont tous des Machines de Turing: chaque table d'états de cette machine correspondant à un de leurs programmes, et le ruban à leur mémoire.

1-févr.-05

Abraham/LaLICC Brest

43

grammaires formelles et automates

classification de la complexité syntaxique : *un automate est équivalent à une grammaire s'il accepte exactement le langage engendré par la grammaire*

	type	langage	automate	remarque
$a^m b^p$	3	régulier (markovien) 	AEF	un mot dépend uniquement du précédent
$a^n b^n$	2	context-free 	à pile	structure d'emboîtement récursif
$a^n b^n c^n$	1	context-sensitif 	linéairement borné	contrainte de chevauchement
	0	récursif	(m. Turing)	calculabilité

1-févr.-05

Abraham/LaLICC Brest

44

«gr.sémantiques»

- A.Bonnet : au prix d'une compréhension plus pauvre mais suffisante pour le but recherché, il est possible d'éviter une analyse syntaxique pour comprendre des phrases dans un domaine technique où le langage est précis.
- inconvenient : " les grammaires sémantiques ne permettent pas d'obtenir une analyse complète de la phrase. La structure produite par l'analyseur ne précise pas quel est le sujet, le complément direct d'objet ou quelle est exactement la partie du groupe nominal qui est modifiée par le groupe prépositionnel, etc..
- En conséquence, certaines nuances ne peuvent être prises en compte par l'analyse.

1-févr.-05

Abraham/LaLICC Brest

45

grammaires «sémantiques»

la complexité des phrases acceptables est limitée mais ceci n'est pas très contraignant car la tendance des spécialistes d'un domaine est d'utiliser un jargon précis et concis, voire non grammatical ex : langage médical, petites annonces,...

Nodule volumineux de 3cm de large poumon droit..

- Les classes d'information que l'on fait apparaître dans la grammaire sont simples : {maladie}, {taille}, {localisation}, ..
- "la grammaire est partiellement transportable à d'autres domaine : la connaissance peut se représenter sous forme de triplets associatifs "objets-attributs-valeurs" ceci inclut la plupart des bases de données développées jusqu'ici.

1-févr.-05

Abraham/LaLICC Brest

46

grammaires formelles les grammaires attribuées (Knuth, 1968) 1

- C'est un formalisme dont le but est d'intégrer la spécification sémantique d'un langage à sa spécification syntaxique par l'attachement d'objets sémantiques appelés attributs aux noeuds des arbres de dérivation des grammaires hors-contexte
- La syntaxe et la sémantique sont définies de façon modulaire et déclarative dans les grammaires attribuées;
- l'une est définie par une grammaire hors-contexte, l'autre par des attributs.

1-févr.-05

Abraham/LaLICC Brest

47

grammaires formelles les grammaires attribuées (Knuth, 1968) 2

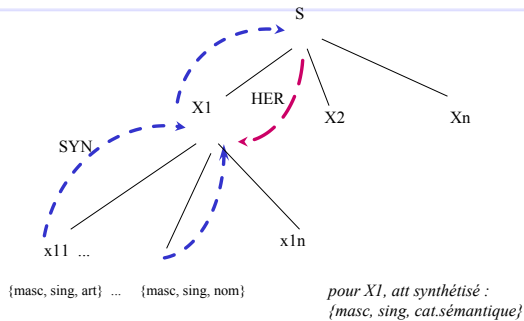
- Soit une grammaire hors-contexte $G_0 = \{Vt, Vn, P, R\}$
 - Vt vocabulaire terminal
 - Vn vocabulaire non-terminal
 - R ensemble des règles de production
 - P est l'axiome.
- chaque règle de production peut s'écrire $r : X_0 \rightarrow X_1 X_2 \dots X_i \dots X_n$.
- Une grammaire attribuée G de support G_0 consiste en un système attribué associé à une grammaire hors-contexte sous-jacente.
- C'est un quintuplet $\{Vt, Vn, Attr, R', P\}$ tel que
- La fonction Attr associée à chaque élément $X \in \{Vt \cup Vn\}$ un ensemble fini de symboles appelés attributs.
- Pour chaque symbole non-terminal $X \in N$, on associe 2 ensembles finis de symboles, Her(X) et Syn(X), les attributs hérités et synthétisés.

1-févr.-05

Abraham/LaLICC Brest

48

gr. attribuées



La méthode attache la sémantique aux syntagmes

1-févr.-05

Abraham/LaLICC Brest

49

grammaires formelles les grammaires attribuées (Knuth, 1968) 3

- Ces ensembles vérifient
- $Her(S) = \emptyset$ (l'axiome n'a pas d'attribut hérité)
- pour tout $X \in Vt$, $Syn(X) = \emptyset$ (les terminaux n'ont pas d'attribut synthétisé)
- Chaque attribut $a \in Attr(X)$ possède un domaine de valeur éventuellement infini noté $VAL(a)$
- P' est un ensemble de règles de production tel que toute production $p' \in P'$ est une production $p \in P$, à laquelle ont été ajoutées des règles sémantiques spécifiant sous forme d'équation le calcul des attributs synthétisés pour X_0 et des attributs hérités de $X_1 \dots X_n$.

1-févr.-05

Abraham/LaLICC Brest

50

Questions sur cette analyse

La méthode attache la sémantique aux syntagmes

Les syntagmes sont concaténés

Est ce juste?

→ Vers un système d'opérations appliquées à des opérands

1-févr.-05

Abraham/LaLICC Brest

51

syntaxe et sémantique

- dans le TAL, est-il licite de considérer que les structures sémantiques sont en bijection avec les structures syntaxiques?
 - Actants (agent, patient)
 - paramètre
 - places
 - rôles interprétatifs
 - → économie de l'encodage linguistique
- problèmes de représentation de la sémantique (cf schéma de compilation)
 - comment construire la sémantique du tout à partir de la sémantiques des parties?
 - combinaison plutôt que concaténation

1-févr.-05

Abraham/LaLICC Brest

52

Grammaires catégorielles

- « les unités linguistiques sont considérées comme des opérateurs qui s'appliquent à des opérands adéquats pour construire des opérateurs ou dans certains cas, des opérands absolus comme les unités de type 'phrase' ».
- Dans une GC, la concaténation syntagmatique d'unités linguistiques élémentaires n'est pas l'opération fondamentale comme dans les grammaires de réécriture de Chomsky mais c'est l'application d'un opérateur à un opérande qui est l'opération constitutive des unités linguistiques plus complexes, comme les prédicats verbaux les prépositions les adjectifs, les adverbes...

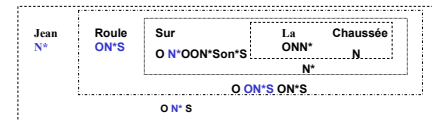
1-févr.-05

Abraham/LaLICC Brest

53

Grammaire catégorielle : exemple

OXY attend une unité de type X pour construire une unité de type Y



Catégorie syntaxique	Type syntaxique	Exemples
Noms	N	
Syntagmes nominaux clos	N°	
phrases	S	
Verbes intransitif	ON°S	courir
Verbes transitifs	ON°ON°S	sortir quelque chose
Articles	ON°	le, la, les, un
Adjectif	ONN	petit, blanc, ...
Adverbe (déterminant de verbe)	OO°N°ON°S	lentement
Adverbe (déterminant de phrase)	OSS	malheureusement
Conjonction de phrases	OSOSS	et, parce que
Prépositions (transpositions)	ON°OO°N°ON°S	sur, vers, dans

1-févr.-05

Abraham/LaLICC Brest

54

Cas syntaxiques, rôles syntaxiques

- latin nom acc gén dat abl
- russe nom gén dat acc instrumental prépositionnel
- allemand nom acc dat gen
- on n'a pas le même nombre de cas suivant les langues, et pas les mêmes usages de ces cas
- les cas syntaxiques ont un lien avec les rôles syntaxiques mais nous sommes loin d'une simple relation point à point
- *un même cas est utilisé pour une variété de significations et souvent des notions qui se ressemblent sont exprimées par des cas différents*

1-févr.-05

Abraham/LaLICC Brest

55

le processus de compilation

c'est un ensemble hiérarchisé de programmes qui assure automatiquement les traductions:

- entre les expressions externes accessibles aux utilisateurs humains,
- et les représentations internes, directement compatibles avec les structures électroniques d'une machine à traiter des informations

1-févr.-05

Abraham/LaLICC Brest

56

intérêt de la compilation

- il est en pratique possible de manipuler des formes symboliques hautement structurées qui représentent des objets complexes (chaînes de caractère, tableaux, arbres, piles, listes, graphes, fenêtres, images..) en utilisant des machines dont la structure des composants de base reste relativement simple :
- pour cela il suffit de changer, par étapes successives, les représentations de façon à se rapprocher du code des représentations supportées par la machine

1-févr.-05

Abraham/LaLICC Brest

57

la théorie de la compilation

elle nous a appris que l'on pouvait

- d'une part, concevoir abstraitement des systèmes de représentations symboliques, plus ou moins déclaratifs et ayant des structures complexes qui tendent à être indépendantes des structures des supports physiques de la machine
- d'autre part, ramener ces représentations à des représentations immédiatement compatibles avec les structures des supports
- assurer le passage par un programme de compilation qui engendre des *représentations internes et hiérarchisées* qui sont les intermédiaires entre les représentations externes accessibles aux utilisateurs du langage, et les représentations numérisées accessibles aux machines

1-févr.-05

Abraham/LaLICC Brest

58

cas, actants, rôles

- cas syntaxiques / cas sémantiques
- idée : proposer une structure casuelle, où un cas n'aurait qu'une réalisation ;
- ce n'est pas ce qui se passe dans les langues
- quel est le critère pour décider des cas ?
 - Le modèle syntaxique est insuffisant
 - Il encode une signification, mais comment la trouver?
- il n'y a pas bi-univocité entre syntaxe et sémantique (*le médecin et le malade guérissent*)
- On peut proposer une structure actancielle, avec différents rôles sémantiques
- reconnaître des rôles sémantiques et voir comment ils sont encodés par la grammaire

1-févr.-05

Abraham/LaLICC Brest

59

grammaire des cas Fillmore

ex : *Jean casse la branche avec une pierre* ->

la branche casse

la pierre casse la branche

On propose les cas suivants :

- agent
- instrument
- datif : animé affecté par l'action
- factitif objet résultant
- lien
- objet

1-févr.-05

Abraham/LaLICC Brest

60

Schank dépendances conceptuelles

- les cas sémantiques relient des entités conceptuelles (et non des éléments syntaxiques comme un GN, un GP) à une des onze actions conceptuelles de base (et non aux verbes de surface)
- un seul cas sémantique peut lier des entités différentes (le cas «bénéficiaire», par exemple fait intervenir le donneur et le receveur)
- les cas liés à une action conceptuelle de base sont tous obligatoires (s'il ne sont pas réalisés dans une phrase, il faut faire des inférences pour les retrouver)

1-févr.-05

Abraham/LaLICC Brest

61

Ambiguïtés (!?) de la LN

- Polysémie des mots (créneau, rouler, blanc)
- Transformations syntaxiques (paraphrases)
 - Marie lave le linge
 - La lessive lave le linge
 - * Marie et la lessive...
 - Jean cuit le mouton -> la cuisson du mouton par Jean
 - Jean aime le mouton -> l'amour du mouton?
- Structures de surface, structures profondes
- Systèmes casuels (un cas n'encode pas une seule valeur)

1-févr.-05

Abraham/LaLICC Brest

62

Universaux Ou Tendances?

- un fait sollicite la réflexion: il est universellement possible de traduire => il faut bien que les langues aient de sérieuses homologies pour pouvoir être ainsi converties les unes dans les autres. Mais... [À quel niveau se situent ces homologies?](#)
- Les langues sont isomorphes ; des principes communs sont sous-jacents à leur structure => approche structuraliste: [y a-t-il des traits universels dans la structure des langues? À quel niveau se situent ces universaux?](#)
- typologie et universaux en linguistique
- la structure des langues est l'ensemble des principes d'organisation qui en assurent le fonctionnement sur le plan des sons de la grammaire et du lexique
- on peut l'étudier selon différents points de vue: la typologie, la recherche d'universaux qui s'attache aux particularités qui sont supposées propres à la totalité des langues, par exemple : [systèmes de sons](#) [sujet aspect genre ordre des mots](#) [deixis](#) [thématisation relatives](#)
- c'est toujours la relation entre forme et sens dans la chaîne et dans le système qui sont étudiés

1-févr.-05

Abraham/LaLICC Brest

63

Lexique : cat.syntaxique : les parties du discours

- Les parties du discours ont des schémas de variation syntaxiques différents; les classifications sémantiques ne traversent pas les parties du discours
- huit classes : **nom, pronom, verbe, participe (adjectif?), conjonction, adverbe, préposition, interjection**
- remarques : la théorie des parties du discours serait valide si ses catégories étaient représentées dans toutes les langues; de plus elles ne sont pas homogènes (ex : adverbe)
- si on se limite à la description d'une langue donnée, **ce n'est pas universel**
- **Elles ne sont pas homologues d'une langue à l'autre**

1-févr.-05

Abraham/LaLICC Brest

64

sémantique d'un langage: du mot à la phrase

on souhaite associer **une** signification à chaque *élément* syntaxiquement correct du langage

- le mot
 - Catégoriser les mots (catégories syntaxiques : parties du discours)
 - la sémantique : relation terme <-> concept
 - dans les langages informatiques / mathématiques
 - en terminologie (normes)
 - dans les langues naturelles
- La phrase
 - la valeur dénotée par une phrase du langage est appelée la sémantique de cette phrase ou la signification ou l'interprétation de cette phrase
 - Les structures syntaxique et sémantique sont-elles homologues?

1-févr.-05

Abraham/LaLICC Brest

65

S34/ C2- : LN et TAL

- Du lexique à la phrase:
 - Rôle des parties du discours : comportement morpho-grammatical
 - Syntagmes à rôles syntaxiques (sujet, objet) porteurs de rôles interprétatifs (agent, patient, ...)
- syntaxe et sémantique
- grammaires formelles
 - quelques difficultés de ces grammaires formelles
- grammaires cognitives

1-févr.-05

Abraham/LaLICC Brest

66

Catégories lexicales

- Catégories syntaxiques : les parties du discours
 - Noms,
 - Verbes,
 - Adjectifs,
 - Pronoms
 - Prépositions
 - conjonctions Adverbes
- Classement sémantique interne à chaque partie du discours
 - Noms :
 - Verbes,
 - Adjectifs : couleurs, formes, qualités (jugement) ...
 - Pronoms ??? Je, me, qui,
 - Prépositions locatives ?
 - conjonctions Adverbes
- Ontologies ??? (organisation du monde)

1-févr.-05

Abraham/LaLICC Brest

67

Du lexique à la phrase

- le lexique : utilisation des parties du discours?
classement des divers éléments du lexique selon leur catégorie grammaticale
 - nom, verbe, adjectif, pronom, relatifs, prépositions, conjonctions,
- la phrase : fonction syntaxique?
 - sujet, verbe,... sujet, attribut, cod, ccirc, ... (un verbe peut être placé en position de sujet)
 - ou : groupe sujet, groupe verbal ... (fonction syntaxique attribuées à un syntagme nominal, syntagme verbal,...)
- **Le rôle interprétatif d'un paramètre est inséré dans un contexte, et non pas dans une ontologie**
 - Une même entité peut avoir les rôles d'agent, de patient, d'instrument ...

1-févr.-05

Abraham/LaLICC Brest

68

fonctions syntaxiques 1

- en informatique, la syntaxe permet de construire (vérifier) la bonne formation de la phrase
- dans la langue, la syntaxe est une partie de la grammaire qui décrit les règles par lesquelles les unités linguistiques se combinent en phrases
- la fonction d'un élément n'est pas directement déterminée par sa nature syntaxique :
 - deux éléments de nature différente peuvent avoir même fonction (ex deux mots appartenant à des parties du discours différentes peuvent jouer un même rôle : un substantif et un adjectif peuvent être attribués)
 - ou bien, un même élément peut se trouver dans deux fonctions : le lapin mange/ je mange le lapin

1-févr.-05

Abraham/LaLICC Brest

69

plan

- Écrire les mots : les alphabets
- Ranger les mots : Les dictionnaires
- Construire des phrases : la grammaire
- Décrire les mots : la Sémantique

1-févr.-05

Abraham/LaLICC Brest

70

Grammaires cognitives

1-févr.-05

Abraham/LaLICC Brest

71

Schank

- cherche à identifier **des notions primitives indépendantes de la langue de référence**
- système de dépendances conceptuelles : relations conceptuelles (indépendantes de la langue) entre des objets et des actions
- acteur animé qui produit l'action
- objet ce sur quoi porte l'action
- bénéficiaire celui qui reçoit un objet
- direction point de départ et d'arrivée d'un objet
- état valeur d'un attribut
- instrument élément qui contribue à l'action

1-févr.-05

Abraham/LaLICC Brest

72

Schank

les concepts qui interviennent dans les représentations construites appartiennent à diverses classes :

- les objets physiques
- les actions
- les attributs d'objets qui précisent des états sous la forme attribut-valeur
- les attributs d'action
- deux autres classes servent à définir le contexte :
- le lieu loc
- le temps représenté par un point sur un axe

1-févr.-05

Abraham/LaLICC Brest

73

Schanck les onze concepts de base

- propel appliquer une force à qqch *actions physiques*
- move déplacer une partie du corps
- grasp attraper un objet
- ingest introduire qqch à l'intérieur d'un objet animé
- expel rejeter qqch à l'extérieur d'un objet animé
- ptrans changer la position physique d'un objet *action elle-même*
- atrans changer une relation abstraite d'un objet (comme la possession)
- speak produire un son *instruments d'autres actions*
- attend diriger un organe d'un sens vers un stimulus
- mtrans transfert d'informations d'un individu à un autre ou entre deux éléments de la mémoire d'un même individu
- mbuild création de nouvelles pensées

1-févr.-05

Abraham/LaLICC Brest

74

Schanck

- la description des «actes mentaux» exige aussi un modèle de fonctionnement de la mémoire
- on la considère comme composée de :
- un processeur conceptuel (PC), là où arrivent toutes les informations conscientes et qui ne traitent qu'un item à la fois
- d'une mémoire intermédiaire (MI), les lieux où sont stockés les connaissances servant aux opérations en cours
- et d'une mémoire à long terme MLT qui contient toutes les connaissances
- Schanck ne montre pas comment l'on passe de l'encodage linguistique aux structures de dépendance conceptuelles

1-févr.-05

Abraham/LaLICC Brest

75

La sémantique décomposer les mots du lexique

• ,

1-févr.-05

Abraham/LaLICC Brest

76

Description du lexique verbal : principe et méthodologie

Principe d'ancrage :

les catégorisations linguistiques sont ancrées sur les catégorisations opérées par la perception et l'action (cf. notion de « situation saillante du point de vue de la perception visuelle »)

Méthodologie :

- Décrire les différentes significations d'un verbe :
 - à l'aide de primitives (types, relateurs, opérateurs)
 - organisées dans des schèmes (les SSC)
- Relier ces significations entre elles (dans un réseau)

1-févr.-05

Abraham/LaLICC Brest

77

Trois sortes de primitives sémantico-cognitives (dans GA&C)

- **Invariants formels : (structurants) modes de composition et de construction des schèmes;**
- **Invariants grammaticalisables**
 - Statiques
 - Cinématiques
 - Dynamiques
- **Invariants expérientiels (perception empirique)**

Les Primitives

Types (catégorisations perceptives de base)

<u>I</u>	entités individualisables	ex. : Jean, pierre, voiture
<u>L</u>	lieux	ex. : maison, virage, route
<u>C</u>	entités collectives	ex. : foule, file (de voitures)
<u>M</u>	entités massives	ex. : beurre, eau

Opérateurs (topologiques) :

EX	extérieur
IN	intérieur
FR	frontière
FE	fermeture



Relateurs (statiques, cinématiques ou dynamiques) :

ε	repérage entre un objet (ou un lieu) et un lieu qui fonctionne comme repère
MOUVT	mouvement spatio-temporel d'une entité passant d'un lieu à un autre
CHANGT	changement affectant la propriété d'une entité, d'un mouvement ou d'une situation complète
CONTR	contrôle d'une situation de modification affectant une entité par une entité
FAIRE	effectuaiton d'une situation de modification par une entité
TELEO	téléonomie d'une situation visée comme un but par une entité

1-févr.-05

Abraham/LaLICC Brest

79

Nature des primitives ?

Prédicats lexicaux, prédicats profonds, opérateurs abstraits ?

- Dans quelle architecture s'insèrent-ils ?
- Ancrage dans d'autres activités cognitives ?
- Dans quel système formel sont-elles formulées, manipulées, composées ?
- Comment met-on en place les primitives ?

Primitives dans la GA&C (1)

(i) Les primitives sont des **opérateurs abstraits** déterminés par des propriétés formelles (combinatoires et topologiques).

(ii) Les primitives sont **ancrées** sur la perception et l'action (hypothèse interactionniste)

(iii) Les primitives sont des **constituants des schèmes**; ces derniers **engendrent**, par composition, **des représentations sémantico-cognitives**; primitives, schèmes et représentations font partie d'un niveau de représentation cognitive **avec une structure non prédicative**.

Les primitives dans la GA&C (2)

(iv) Les primitives sont des opérateurs insérés dans le formalisme de la logique combinatoire avec types (de Curry) Où des opérateurs abstraits – des combinateurs – composent les primitives entre elles de façon à construire des schémas et des représentations de situations.

(v) Les primitives sont mises en place à partir d'actes abductifs : à partir d'un problème et de données constituées, on formule une hypothèse plausible (existence de primitives et de schémas reliés aux observables par des relations formelles et calculées).

L'ordinateur intervient comme instrument de simulation expérimentale.

1-févr.-05

Abraham/LaLICC Brest

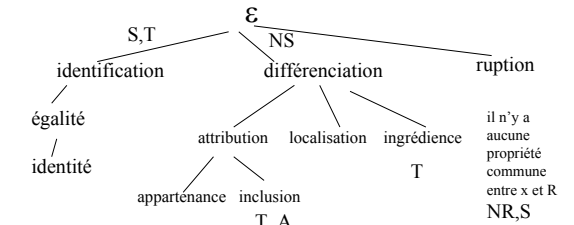
83

plan

- Écrire les mots : les alphabets
- Décrire les mots : la Sémantique
- Ranger les mots : Les dictionnaires
- Construire des phrases : la grammaire

le repérage : $x \in R$

- schéma général de relations :
- x est repéré par rapport à un repère R



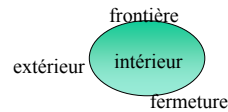
1-févr.-05

Abraham/LaLICC Brest

84

la relation de localisation

- être localisé par rapport à un repère
 - des spécification supplémentaires peuvent être apportées par des relateurs topologiques, qui spécifient :
 - l'intérieur $x \in_0$ IN loc
 - l'extérieur $x \in_0$ EX loc
 - la frontière $x \in_0$ FR loc
 - la fermeture $x \in_0$ FE loc
- d'un lieu loc



1-févr.-05

Abraham/LaLICC Brest

85

Exemple : description d'une entité

- une voiture est un véhicule **voiture \in {véhicules}**
- constitué de 4 roues **roue \in_{ingr} voiture**
- un volant **volant \in_{ingr} voiture**
- 4 portes **porte \in_{ingr} voiture**
- des sièges **siège \in_{ingr} voiture**
- elle peut engendrer un lieu dans lequel sont placés le chauffeur et les passagers **chauffeur \in_0 IN voiture**
- elle engendre un système évolutif qui peut se trouver dans différents états (en marche, à l'arrêt,...) **SYS(voiture) : <E1,E2,...>**

1-févr.-05

Abraham/LaLICC Brest

86

Description du lexique verbal

Représentation des verbes de mouvement à l'aide du formalisme des Schémas Sémantico-Cognitifs (SSC)

- Description du lexique verbal : principe et méthodologie
- Opérations réalisées par les prépositions (spatiales)

1-févr.-05

Abraham/LaLICC Brest

87

Importance du verbe dans la phrase



1-févr.-05

Abraham/LaLICC Brest

88

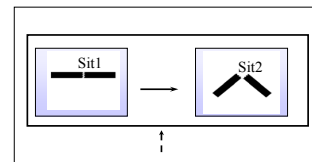
exemple : représenter /casser/



en BLISS: on verbalise une rupture



communimage



dans la GAC: /casser/ se représente par un schéma de transformation entre deux situations référentielles Sit1 et Sit2

1-févr.-05

Abraham/LaLICC Brest

89

Schémas sémantico-cognitifs (SSC)

Décrire une signification verbale à l'aide d'un agencement structuré de primitives

- les Situations statiques
catégorie de l'équilibre, du stable et du stabilisé
- les Situations évolutives
catégorie des modifications

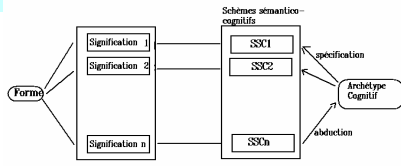
1-févr.-05

Abraham/LaLICC Brest

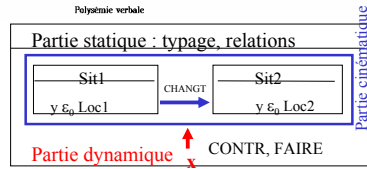
90

Le verbe : Polysémie, Sens et significations

A une même forme d'entrée dans le dictionnaire, correspondent plusieurs significations que l'on décrit à l'aide de primitives et de relateurs, dont la combinaison forme des SSC

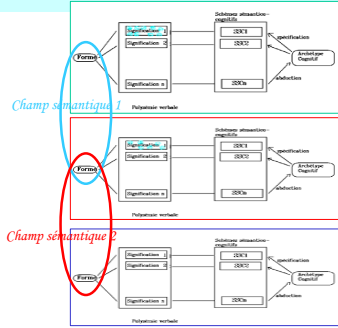


Schémas de transformation



Polysémie, synonymie

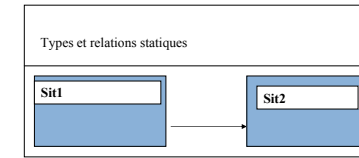
Champ du déplacement (orienté)
Monter, grimper, gravir



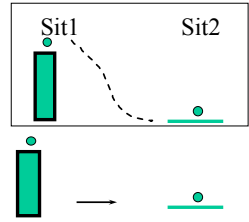
Champ de la construction
Assembler, construire, bâtir, édifier,

Catégorie des SSC évolutifs

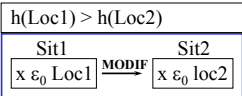
- les SSC cinématiques : expriment une modification spatio-temporelle sans agent qui l'effectue et la contrôle



Exemple : la bille tombe

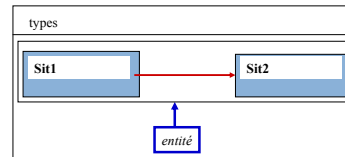


il se passe une modification spatio-temporelle entre une situation de référence Sit1 et une situation de référence Sit2: en Sit1, la bille est localisée par rapport à un lieu loc1; en Sit2, elle est localisée par rapport à un lieu loc2; de plus, il existe une relation entre les hauteurs des lieux loc1 et loc2

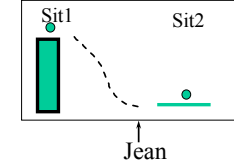


Catégorie des SSC évolutifs

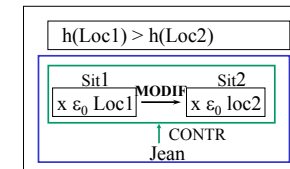
- les SSC dynamiques : expriment une modification qui est sous la dépendance d'une entité externe (un agent qui effectue et contrôle un mouvement, un instrument qui rend possible un changement, sous la dépendance d'un agent ou non, ...)



Exemple : Jean fait tomber la bille



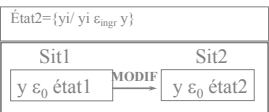
il faut de plus indiquer que Jean exerce une action pour que la bille tombe; c'est bien ce qui est encodé par la phrase



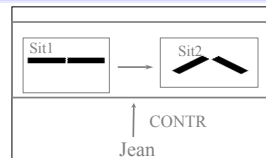
Cinématique : la branche casse



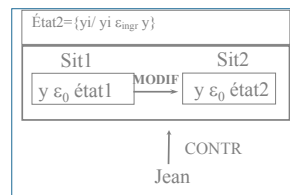
Casser, c'est constituer des ingrédients à partir d'une entité y. On peut retenir deux situations saillantes de référence: il se passe une modification spatio-temporelle entre une situation de référence Sit1 et une situation de référence Sit2: en Sit1, la branche est dans un état état1; en Sit2, elle est dans un état état2; Ce changement d'état peut être précisé par les relations : y1 ε_{ingr} y; ... y_n ε_{ingr} y;



Dynamique : Jean casse la branche



il faut de plus indiquer que Jean exerce une action pour que la branche casse; c'est bien ce qui est encodé par la phrase



la modification est dynamique

SSC : sémantique du schéma verbal

- un verbe encode une **combinaison de relations**
 - statiques, cinématiques, dynamiques
- un même item verbal peut encoder des situations statiques cinématiques dynamiques
- la partie statique décrit les relations constantes pendant la transformation spatio-temporelle
- la partie cinématique décrit la transformation qui peut affecter l'entité, son état, ses propriétés, ses attributs, son type sémantique
- la partie dynamique, qui n'est pas grammaticalisée en français (ex : guérir) encode un agent ou un instrument responsables de la transformation
- au niveau prédicatif, l'intégration de ces relations par la logique combinatoire construit un prédicat décrivant la sémantique du schème

archétype cognitif d'une lexie

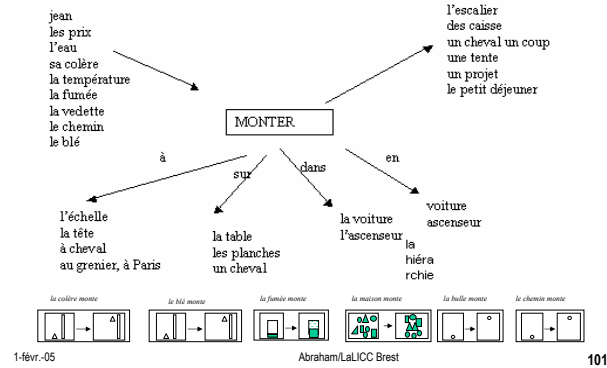
- invariant, mais pas en langue, de nature *noémique*, *ie*, *conceptuelle*
- il correspond à une *forme lexicale*
- Il est obtenu par abduction, à partir des différentes réalisations linguistiques
- Un des significations s'obtient par des spécifications dans les domaines du spatial, du notionnel (ou de la cause)
- Remarque : les primitives « empiriques » spécifient l'archétype

1-févr.-05

Abraham/LaLICC Brest

100

Exemple : monter

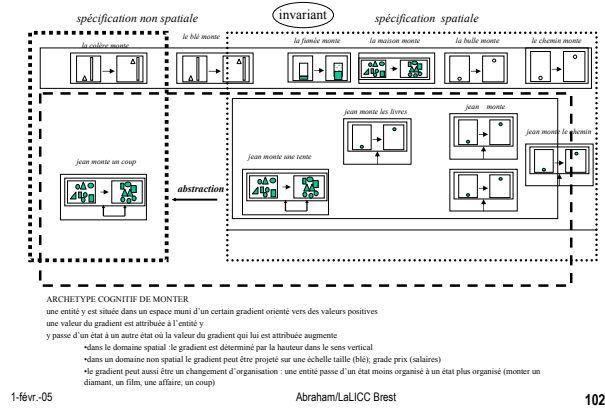


1-févr.-05

Abraham/LaLICC Brest

101

structuration des SSC de /monter/



1-févr.-05

Abraham/LaLICC Brest

102

Archétype cognitif de monter

- une entité est située dans un espace muni d'un certain gradient orienté vers des valeurs positives ; une valeur du gradient est attribuée à l'entité y ; cette entité y passe d'un état à un autre état où la valeur du gradient attribué à y augmente ».
- Cet archétype peut être réalisé dans le **domaine spatial** ; dans ce cas, le gradient est déterminé par la hauteur verticale: une entité passe d'une position à une autre position plus élevée.
- Le même archétype peut être réalisé dans un **domaine non spatial**. Dans ce cas, le gradient peut être projeté sur une échelle (augmentation de taille: *le blé monte*; du grade: *monter en grade*; du prix: *les salaires montent* ; ...).
- Le gradient peut être aussi un **changement d'organisation** : une entité passe d'un état moins organisé à un état plus organisé (*monter une bague*, *monter un film*, *monter une affaire*, *monter un coup*)

1-févr.-05

Abraham/LaLICC Brest

103