INTERACTION BETWEEN LEXICON AND IMAGE: LINGUISTIC SPECIFICATIONS OF ANIMATION

Maryvonne ABRAHAM *†, Jean-Pierre DESCLÉS †

atelier "Formalisation" du GDR 957 (CNRS) Sciences Cognitives de Paris. France†

ABSTRACT

Certains textes d'une langue peuvent être considérés comme des spécifications linguistiques capables d'engendrer des animations qui simulent la compréhension des textes d'entrée. Pour réaliser un tel programme, il nous semble indispensable de construire des représentations intermédiaires compatibles, d'un côté avec les descriptions sémantiques d'unités linguistiques, et d'un autre côté avec des spécifications d'animation d'images.

Comprendre un texte relatif à des mouvements spatiaux revient: i) à représenter sémantiquement le texte; ii) à engendrer une animation mettant en jeu des images; cette animation vise à reproduire ce qui est compris lors de la lecture du texte. Ce programme ambitieux suppose qu'au préalable chaque verbe ait des représentations que nous appelons schèmes sémantico-cognitifs (SSC), correspondant aux différentes significations du verbe. Tout schéma de mouvement associé à la signification d'un verbe représente trois types de connaissances: i) des déclarations et des relations invariantes pendant le mouvement; ii) la description cinématique du mouvement qui fait passer d'une phase initiale (Sit_1) à une phase finale (Sit_2); iii) les conditions dynamiques éventuelles qui rendent possibles ou qui contrôlent le mouvement cinématique.

Nous reprenons l'approche de R.Schank ou de J.Sowa qui remet en cause une association directe mot-concept pour adopter une décomposition de la signification des unités linguistiques en éléments de "sens" plus petits. Notre formalisme a pour but d'une part, de mieux traiter les problèmes de temps, et d'autre part, de fonder notre choix de primitives, dans le cadre d'une théorie qui articule les niveaux de représentation linguistiques et cognitifs.

Introduisons quelques distinctions:
1- le verbe est une entité lexicale qui peut être polysémique. ex: *circuler*.
2- le verbe syntaxique entre dans une construction syntaxique. ex: N_1 circuler prep N_{loc}.
3- le prédicat logique n-aire est associé au verbe syntaxique. ex: CIRCULER est l'expression du prédicat 1-aire $CIRCULER_1$.
4- A un prédicat correspondent un ou plusieurs sens, désignés par /verbe^i/, représentés chacun par un SSC. Un SSC est organisé à partir de primitives cognitives; il conduit à une représentation cognitive dès qu'il est instancié.
5- l'archétype cognitif noté /verbe/, s'il existe, se présente alors comme "le sens abstrait" commun à tous les SSC d'une même entrée lexicale polysémique. La construction de l'archétype repose sur une organisation des SSC en réseau. Elle utilise une description analogue à celle de J.Picoche.

A partir de l'exemple d'un texte français qui représente des mouvements spatio-temporels, nous donnerons pour quelques verbe qui lie les SSC et les animations correspondantes. L'animation complète déclenchée par la compréhension de texte nécessite cette première phase de recherche.

Les schèmes que nous proposons sont intégrables dans un dictionnaire sémantique du français. Notre étude présentera à partir de l'exemple choisi la méthodologie appuyée par une réalisation.

mots-clés { *compréhension de texte, représentation des connaissances dynamiques, lexique verbal, méthodologie, choix de primitives sémantico-cognitives, verbes de mouvement, analyse d'un texte, dictionnaire sémantique, spécification linguistique et animation* }

*Département Mathématiques et Systèmes de Communication, LIASC, Telecom Bretagne, BP 832, 29285 Brest cedex, tel: 98 00 11 11 - FAX: 98 00 10 98 *e-mail: ABRAHAM@brest.enst-bretagne.fr*

†C.A.M.S., unité mixte CNRS - EHESS - Université Paris-Sorbonne, 96 Boulevard Raspail, 75006 Paris - *tel* (16-1) 45 48 08 04 - FAX : (16-1) 40 49 07 98 - *e-mail : isha@frciti51.bitnet*

1 understanding a text and a moving picture

Certain texts of a given langage can be considered as linguistic specifications able to produce motion pictures simulating the understanding of entry texts. To achieve such a program, we think that it is necessary to build intermediate representations which are compatible, on the one hand with the semantic description of linguistic units, and on the other hand, with specifications for the moving of pictures.
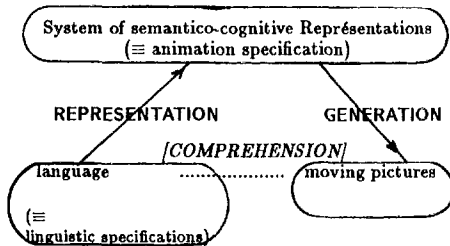

Figure 1: interaction between lexicon and image

Understanding a text relative to spatial motion consists of: i) a semantic representation of this text; ii) movement with pictures; this animation aims to reproduce what is understood during the reading of the text. As a first step, this ambitious program requires that each verb should have representations which we call semantico-cognitive schemes (SCS). These representations correspond to the different meanings of the verbs. Each scheme of motion associated with the meaning of a verb represents three kinds of knowledge: i) declarations and relations which are invariant during the motion; ii) the cinematic description of the motion from an initial phase (Sit1) to a final phase (Sit2); iii) the dynamic conditions which make or control kinetic motion.

1.1 semantico-cognitive representations

Although several formalisms of knowledge representations (e.g. semantic networks) suppose that the word is the elementary semantic unit,

other approaches contest this direct association word- concept, and start from a breakdown of the signification of the linguistic units[1], following R.Schank in the 1975's. The meaning of the words is no longer considered as primitive but can be split up into smaller signification units. More recently, Sowa has presented a formalism of representation based on graphs with two sorts of nodes representing concepts and relations. However, many researchers feel that the choice of the primitives looks arbitrary ([SAB88], pp.241-243). We propose a formalism similar to Sowa's graphs [SOW84], but we try to process time better and establish our choice of primitives[2], using a theory which arranges links between the levels of linguistic and cognitive representations.

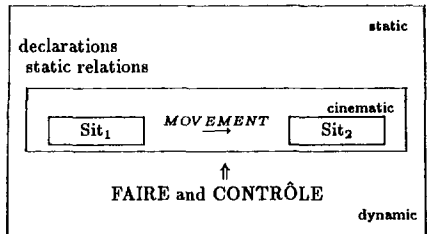## 1.2 semantico-cognitive schemes of a verb

Figure 2: formal diagram representing semantico-cognitive schemes (SCS)

The cognitive archetypes (or semantico-cognitive schemes, written SCS, see figure 2), were introduced in 1985 [DES85] in order to represent the verbal knowledge underlying the language. Our method to discover them is to collect linguistic data which allows us to characterize prototypic schemes of meanings from which an abstract invariant may appear, in some cases.

The following levels can be taken into consideration:
1- The verb, written *verb*, is a lexical entity; it can be polysemic (e.g., *sortir*).
2- The syntactic verb written $verb_i$ is a part

[1]See [KAY89] about the notion of variable depth.

[2]The Applicative and Cognitive Grammar model increases S.Schaumyan's initial model with more cognitive levels where we aim to discern operations about cognitive categories; We have described the main primitives (MOUVT, FAIRE, CONTRÔLE, ...) which are used in several publications [DES90]; see [JPD91] for time processing, and [DES91] for the theory underlying the choice of the primitives.

of a syntactic construction (e.g., N_1 sortir prep N_{loc}).

3- The n-ary predicate is associated with the syntactic verb (e.g., $SORTIR_1$)[3].

4- With a predicate, one or several meanings are associated, written /verb[i]/[4]. A SCS is organised from cognitive primitives. It produces a cognitive representation the moment it is instanciated.

5- Then, if it exists, the cognitive archetype (written /verb/) appears as the "abstract meaning" common to each of the SCS of one lexical polysemic entry. Its construction is organised from SCSs. It uses a description similar to the one in J.Picoche's method [PIC86][5].

From a sample text representing spatio-temporal motion, we give the semantico-cognitive schemes and the corresponding motions for each verb of the text. This first step is required in order for an understanding of the text to trigger complete animation.

2 A few verbs to represent a text

The schemes that we propose are integrated in a semantic dictionary of French. From a chosen example, our study presents the methodology based on demonstration. To illustrate our method, we analyse the verbs in the following French text and we give cognitive representations of them:

> *Madame Dupont circulait dans une petite rue à sens unique où stationnaient des voitures de chaque côté. Un véhicule sortant d'un parking surgit brusquement sur sa droite.*

which means:

> *Mrs Dupont was driving along a small one-way street where cars were parked on each side. A vehicle leaving a car park shot out on her left[6].*

---

[3]In agreement with the conceptual language of R.Jackendoff [JAC83]

[4]i indexes one of the prototypical meanings of the polysemes.

[5]See two examples associated with the different meanings of the French verbs *rouler* [ABR92] and *monter* (J.P.Desclés, article to be published). We need syntactic descriptions such as those of M.Gross ([GRO81]). For identifying the SCSs, we use works on semantics such as [COL83] and [MEL84].

[6]Notice that we make more than a lexical translation from French into English: cognitive laws of the highway code are "translated" too.

As we have not enough room in this paper to describe each verb of this text, we have chosen to describe the French verbs *stationner*, and *surgir*.

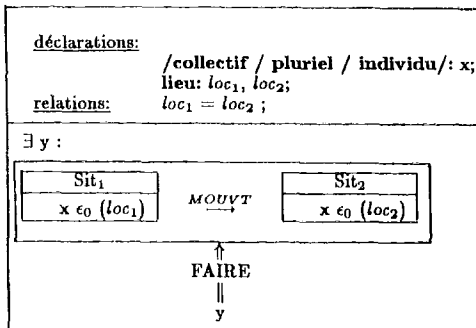### 2.1 SCS of the French verb *stationner*

Figure 3: SCS of the French verb "stationner"

In x *stationne*, we have a static positional relation. This situation is implicitly bound by a constraint (which can be more or less explicit). The signification of /stationner/ in this proposition is given by the dynamic scheme in figure 3. The quantification which binds the variable y blocks every instanciation of y. In our example, the SCS is instanciated with:

$x := $ *des voitures*

$y := $ *de chaque côté d'une petite rue à sens unique*

A constraint is needed to keep a car parked. It is expressed here using the primitive "FAIRE" which keeps the movement from Sit_1 to Sit_2 in the same place. We say that /x stationne/ if the behaviour of x implies a constraint on its own movement which keep it in the same place. This constraint is expressed by the equality between the localisations loc_1 of Sit_1 and loc_2 of Sit_2.

2.2 SCS of the French verb *surgir*

x *surgit* means a sudden transition, a movement entering the perception zone of an entity able to perceive. The perception zone is generally taken from the speaker's point of view. We have a rule ([DES90] pp.285-287, [ABR91]§2.1) concerning localisation:
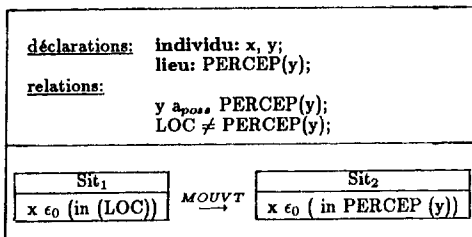
déclarations: individu: x, y;
lieu: PERCEP(y);
relations:
y a_{poss} PERCEP(y);
LOC $\neq$ PERCEP(y);

Sit_1	$\xrightarrow{MOUVT}$	Sit_2
x ϵ_0 (in (LOC))		x ϵ_0 (in PERCEP (y))

Figure 4: SCS of the French verb "surgir"

Figure 5: on Mrs Dupont's left (left in GB!)

rule: *if x is an entity typed* individu *x determines a localisation loc(x) including the entity x.*

In fact, to each x is associated a family of localisations, among which a proper localisation, a peripheral localisation depending on cultural habits, and a neighbouring localisation dependent on the context. The localisation loc (x) can be determined by taking its internal area, its external area, or its frontier. The family of localisations can be extended to the localisations perceived by the individual entity.

definition of perception: *The localisation PER-CEP(x) is defined as a result of the intersection[7] of an external area to be perceived and of the properties of perception of the person perceiving (its senses).*

x *surgit* if x moves from Sit_1 to Sit_2. In Sit_1, x is situated in any localisation; in Sit_2, x is situated in the place of perception of the witness, or, by default, in the area of perception of the speaker.

/surgir/ can be associated with a preposition, for example, *sur* (on), which partially specifies the place of arrival. The French relation ÊTRE-SUR can be represented formally by the following set of relations[8]:

definition of the relation ETRE-SUR:
i) *The one referred to is outside a domain D (closed localisation)*
ii) *The one referred is in contact with the external frontier of the domain D.*
iii) *An external action (by default, gravitation) keeps the referred one in contact with this external frontier.*

---

[7]PERCEP(x) can be considered as the result of the intersection of the area SENSE(x) (of an entity that x can perceive) and of the external domain which is examined. VIS(A) is the visibility area of A.
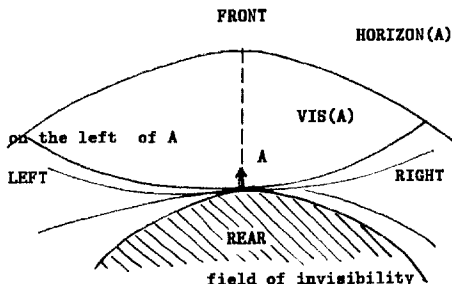
[8]For any explanation, see [ABR91], §1.2.3

Concerning visual space perception (figure 5), the notions of right and left are relative: they can be organised from the point of view either of the subject observed, or of the speaker. Let us take an orientation from the subject observed A, here, *Madame Dupont*, which has an intrinsic orientation. A defines FRONT(A) REAR(A) LEFT(A), RIGHT(A), VIS(A). The intersection of these areas defines frontiers. In our example, we can consider that /to be on the left of Mrs Dupont/ means /to be on the external frontier between the areas VIS(Mrs Dupont) and LEFT(Mrs Dupont)/.

3 Towards a semantico-cognitive lexicon of verbs

We implemented a semantico-cognitive lexicon on a work-station. The lexicon appears as a set of several windows showing:
- a list of entry points: the lexical verbs; by clicking on a verb, a new window is opened, showing:
- a list of examples of the entry verb.
- for each of the examples, a button helps obtain the SCS associated with a given verb (and the corresponding archetype, if it exists).
- Visual animation is associated with the SCS, and it is obtained from the archetype. This animation is the cognitive representation of a skeleton of motion; it may then be "clothed" with the textures specific to the moving object. Our work is limited to the cognitive specification of verbs of motion; it does not take into account the distortion specific to the objects.
- the archetype of the verb, if it exists.

Our demonstration shows the movement corresponding to the verbs in the text. The SCSs (and the corresponding archetype, if it exists) establish a specification from which: i) infor-

mation for inferences is extracted; ii) virtual animation can be generated on a screen. In parallel, we have defined the general language for computer specifications which use the semantico-cognitive primitives that we propose in building the SCSs and the cognitive archetype [ChJ91].

We thank Janet Ormrod, Telecom Bretagne, for her help in producing the English version of this text.

Références

[ABR91]   Maryvonne ABRAHAM J.P.DESCLÉS
R.AFZALI J.BERRI CH.JOUIS H.G.OH.
Représentation du mouvement verbalisé.
AFCET, Actes du 8ème congrès RFIA,
vol.1, pp.11-21, novembre 1991.

[ABR92]   Maryvonne ABRAHAM Rasek AFZALI
Christophe JOUIS. *Les archétypes cogni-
tifs sont des générateurs de représentation
des connaissances.* à paraître in T.A.Information,
1992.

[ChJ91]   B.GAUDINAT J-M.ROUSSEAU Ch.TORA
Ch.JOUIS, F.COMPAGNON. *METODAC:
une méthodologie pour l'acquisition et la
modélisation des connaissances.* AFCET,
Actes du 8ème congrès RFIA, novembre
1991.

[COL83]   COLLECTIF sous la direction de H.Lewicka
et K.Bogacki. *Dictionnaire Sémantique
et Syntaxique des Verbes Francais.* Państwowe
Wydawnictwo Naukowe, Warsovie, 1983.

[DES85]   Jean-Pierre DESCLÉS. *Représentation des
connaissances: archétypes cognitifs, schèmes
conceptuels, schémas grammaticaux.* EHESS-
CNRS, Actes Sémiotiques, VII, no 69-70,
pp. 5-51, 1985.

[DES90]   Jean-Pierre DESCLÉS. *Langages appli-
catifs, langues naturelles et cognition.* Hermès,
Paris, 1990.

[DES91]   Jean-Pierre DESCLÉS. *La prédication opérée
par les langues (ou à propos de l'interaction
entre langage et perception).* in LANGAGE,
N0 103, Larousse, Paris, sept 1991.

[GRO81]   Maurice GROSS. *Formes syntaxiques et
prédicats sémantiques.* Larousse, Langages,
n0 63, 1981.

[JAC83]   Ray JACKENDOFF. *Semantics and Cog-
nition.* MIT Press, Cambridge, 1983.

[JPD91]   J.P.DESCLÉS CH.JOUIS H.G.OH D.REPPERT
F.SEGOND. *Représentations et recherches
des valeurs sémantiques des temps de l'indicatif
du francais pour une mise en oeuvre in-
formatique.* CNRS GDR 957, Centre d'Analyse
et de Mathématiques Sociales - ISHA -
Paris, Mai 1991.

[KAY89]   Daniel KAYSER. *Adéquation inadéquation
de la logique au traitement sémantique
des langues.* Actes du colloque sémantique
et cognition, à paraitre aux éditions du
CNRS, Université Paris-Sorbonne, 1989.

[MEL84]   Igor MELCHÚCK & alii. *dictionnaire ex-
plicatif et combinatoire du français con-
temporain - recherche lexico-sémantique.*
Presses de l'Université, Montréal, Canada,
1984.

[PIC86]   Jacqueline PICOCHE. *Structures sémantiques
du lexique français.* Nathan, France, 1986.

[SAB88]   Gérard SABAH. *L'intelligence artificielle
et le langage. t1; représentation des con-
naissances.* Hermès, Paris, 1988.

[SOW84]   J.F. SOWA. *Conceptual structures: In-
formation Processing in Mind and Ma-
chine.* Addison-Wesley, London, 1984.