EXPERIMENTS IN LANGUAGE ACQUISITION
BY ARTIFICIAL SYSTEMS

Gérard Sabah, Andrei Popescu-Belis
Language and Cognition Group — LIMSI
B.P.   133
91403 ORSAY Cedex — FRANCE

Abstract

This paper presents some aspects related to natural language acquisition in our CARAMEL architecture. The CARAMEL model emphasises, at a global level, the importance of both "conscious" and "unconscious" processes for natural language understanding, drawing inspiration from theoretical work by Harth, Baars and especially Edelman. Three experiments on language grounding are described: prerequisites to language for an agent in its environment; evolution of syntactic conventions between agents; and conceptual bootstrapping for an agent exposed to language.

1. Introduction

From our point of view, taking into account consciousness and its related aspects is of particular interest for natural language understanding. This leads us to propose a general model of reasoning and intelligence that should not only apply to natural language processing but also to reasoning and learning (since from our point of view, true understanding cannot be isolated from acquisition); this model should explain how authentic semantics, or symbol grounding, can appear in a given mind.

Our model is named CARAMEL — in French: *Conscience, Automatismes, Réflexivité et Apprentissage pour un Modèle de l'Esprit et du Langage*, or in English: *Consciousness, Automatic processes, Reflectivity and Learning for a Model of Mind and Language* (Sabah 1995, 1997a, Sabah and Briffault 1993). It shows how reflectivity and distributed artificial intelligence allow computer programs to represent their behaviour and reason about these representations in a dynamic way.

Non-controlled processes appeared also to be necessary in this kind of program for computer efficiency reasons as well as for cognitive ones. Therefore, we proposed a blackboard extension — the *sketchboard* (Sabah 1997b) — which uses a different kind of relation between processes: it allows for reactive feedback loops at different levels between processes that do not otherwise know each other.

The CARAMEL model advocates the idea that consciousness has a central role to play for the integration of these two kinds of processes. The model draws inspiration from many sources, the main ones being summarised below.

From Baars' (1988) "economical" conception of consciousness, and his psychological point of view, we retain three main components:
* a blackboard as a workspace where conscious data is written;
* the hierarchy of interpretative contexts, conceived as hierarchies of goals, and the handling of interruptions;

• the competition between several unconscious processes, providing a model of voluntarily control and attention.

Harth (1993) is opposed to Cartesian dualism, as well as to the more recent radical pluralism — as illustrated by (Minsky 1985): "a million witless agents instead of one clever homunculus". His model accounts for the fact that mental images are not replicas of world objects, they are combined with previous knowledge. Top-down pathways allow higher knowledge to modify messages that come from the senses and to inject into them additional information. This process becomes active as soon as sensorial input begins, and not at the end, as one would assume if it were an advanced function of the brain. Therefore, there is no homunculus scrutinising the state of the brain: the brain itself acts as an observer of its first input levels and influences them in order to maximise the recognition; the brain analyses, recreates and analyses again its own productions, in a truly *"creative loop"*.
From Harth's theory, we took :
   • the idea of feedback between unconscious processes;
   • the *a priori* evaluations of unconscious processes; and
   • consciousness acting already at the initial processing levels rather than only at the end.

Edelman's conception of consciousness (Edelman 1989, 1992) is based on a theory of the brain functions, in its turn based on a theory of their evolution and development. The core of Edelman's approach is the Theory of Neuronal Group Selection, also referred to as "Neural Darwinism". The TNGS is based on three principles: ontogenetical selection; secondary synaptic reinforcements or decay; interactions among cerebral repertoires by a bi-directional re-entry.

According to the TNGS, the cerebral cortex is structured in repertoires of neuronal groups which act toward perceptive input as selective systems. As repertoires can also categorise other repertoires' activity, the perceptual input is thus categorised in a more and more abstract way. Some repertoires are connected through reentrant pathways, enabling elementary associative learning. The TNGS thus identifies the neurobiological functions that have allowed the emergence and evolution of elaborated capacities of the human mind; these characteristics also provide a ground for consciousness.

In our view, several characteristics enabling high-level faculties and, inherently, consciousness, can be drawn from Edelman's work:
   1) neural specialisation allowing the distinction of internal signals from world signals,
   2) perceptual categorisation,
   3) memory as a process of continuous re-categorisation with the possibility of representing the activation order,
   4) learning, i.e., links between the categories and the essential values of the individual,
   5) concept acquisition, i.e. categorisation of the brain activity itself through global maps,
   6) primary consciousness, allowing to connect internal states resulting from previous perceptual categorisations to present perceptions — what Edelman calls : *the remembered present*,
   7) an ordering capability which results in *presyntax* and the basis for symbols,
   8) language and
   9) higher order consciousness.
In the realisations described in his paper, we have been mainly inspired by:
   • the definition of unconscious processes as basically producing correlations,

- the definition of semantics as correlations between concepts, sensory input and symbols [1],
- the memory model as a categorisation of processes rather that as a zone for storing representations, and
- the role of language for symbol manipulations.

The PhD theses of Andrei Popescu-Belis and Jean-Pierre Gruselle focus on the study and implementation of capacities which could enable a system to acquire some rudiments of language — lexicon, syntax and semantics — through interaction with a simulated environment. The system is viewed here as an "animat", or simulated autonomous "living" robot. The rest of the paper describes three aspects of our work: evolution of an "Edelmanian animat" (§ 2), emergence/acquisition of syntax for the communication code of a group of animats (§ 3), concept acquisition from exposure to language (§ 4).

2. An Adaptive Multi-Agent System Based on "Neural Darwinism"

Proposals for models of the TSGN

Several computer models of the TNGS (Reeke, et al. 1990) bring convincing justifications to the theory. However, despite the theory's ambitions to account for higher cognitive functions as language and consciousness, the models seem difficult to extend, as they rely heavily on very specific architectures, using fined tuned scalar connections between simulated neurons ("integrative units"). Although the TNGS defends the idea of *neuronal groups* — whose genesis has been simulated in (Reeke, et al. 1990, p. 616 - 627) — they aren't used in the other models.

Our first proposal for the implementation of such control architectures is to use *actors* instead of "integrative units" (actors are very simple active objects). Conversely, numeric synaptic inputs are replaced with symbolic messages between actors, connections with acquaintances, gradual activation states with discrete ones. Second, we propose here an explicit implementation of *values* or needs of our agent. Indeed, for the original models of the TNGS (Reeke, et al. 1990), behaviour is implicitly defined by some sensori-motor connections, used to check the proper execution of categorisation operations. However, as Edelman himself argued, value is a central element of cognitive functions explanation.

Description of the agent

Our simple agent or "animat" implements very basic features of an adaptive being: positional perception of the environment, internal needs or values, and motor capacities. The environment is linear: a segment, an infinite line, or a circle. Two areas of the environment provide two different resources, named N and B ("nutriment" and "beverage").

Mainly inspired by the TNGS notion of neuronal group, the control structure consists of "actors" or modules. Each actor monitors the patterns of a repertoire of other actors or modules, and generates a message depending on this pattern and on its own internal state.

The visual cells of the linear retina provide a topographic representation of the environment (cf. Figure 1). Actors from three repertoires monitor this input area, to detect specific features

---

1. The last word is here used in Newell and Simon's meaning (the basis of "physical symbol systems"). However, our own hypothesis differs from them: we follow Edelman by saying that such system semantics cannot be defined on fomal bases only.

(N position, B position, movement sense). Then, R-of-R actors (second order recognisers according to Edelman) categorise combinations of messages from the first repertoires. There are 16 R-of-Rs, corresponding to combinations of two or three messages.
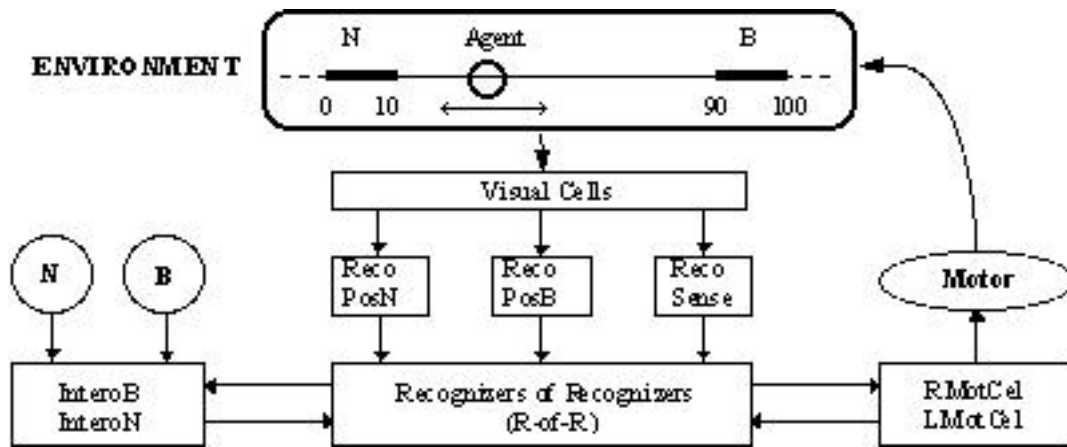

Figure 1. The agent's architecture

"Reentrant" protocols are implemented between interoceptive actors (InteroB, InteroN) and R-of-Rs as well as between R-of-Rs and motor actors (MotCel). These protocols provide the basis for adaptive learning: simultaneous activation of two actors establishes an acquaintance link between them, which is not immutable, but subject to progressive decline. Conversely, activation of an actor induces activation of its acquaintances which are not active.

There is no separation between the learning and the functioning phase, as in most biological systems. Each protocol works both ways: for learning (the MotCel → R-of-R and R-of-R → Intero senses) and application of what was learnt (Intero → R-of-R → MotCel). The experimental evidence confirms the stability of these *reentrant* protocols; see below and (Popescu-Belis 1997).

Operating principle

Once created, the animat has to acquire motor control, then spot the regions where its needs are satisfied. First, the observed coordination between perception and action (random motor activity) leads to internal sensori-motor linkage between MotCel and R-of-R repertoires. Further, the agent recognises the utility of N and B regions only after its needs have been fulfilled — which needs external intervention to bring the agent on N, resp. B, for the first time. Indeed, these regions behave initially only as perceptual landmarks, and are associated with the corresponding value only after learning.

Experimental Results

Computing the activation once for all actors makes a *cycle*. In our experiments, the agent gets initially a full load of N and B. Then, instruction by external intervention is permitted up to 1000 cycles; the agent's lifetime is recorded afterwards, until either N or B are finished. In one set of trials, 12 instances were given the same numeric parameters and similar external interventions; the average lifetime was 1,550,000 cycles. Six instances were stopped after

1,000,000 cycles – another one being left up to 6 million (6.106) cycles. Oscillation between N and B is a stable behaviour, which may be altered by changes in the environment (N/B areas).

These results are very encouraging and we are currently trying to enhance the model, with a more complex animat: a cylinder shape simulated robot, with a circular one-dimension retina, two motor wheels, pressure sensors and an olfactory gradient sensor. Current research topics include kinaesthetic capabilities, perceptual–kinesthetic coordination and learning of motor sequences.

3. Communicating Agents: Study of Emergent Syntax

We examine now how such agents may use a language-like communication code. One of the limitations of natural language understanding by computers is that programs do not generally ground semantics using perception of the real world, actions on it, and internal values. Instead, only a formal definition of the language is given, e.g. formal syntax, formal ontology. Our approach aims at implementing syntactic and semantic properties on situated agents (previous experiment).

However, we use for the moment a formal implementation to study syntactic properties of a communication code at the collective level, which is dynamic and non-deterministic. The model presented here validates the idea of representing syntactic conventions using s simplified Tree Adjoining Grammar (TAG) — see also (Allexandre and Popescu-Belis 1998). The agents are supposed already capable of categorising their environment, which contains geometrical shapes with features and relational properties. Conceptual representation of situations is given to agents in the form of *TAG derivation trees*, which are considered to be a simple representation of meaning in TAG. TAG being lexicalised, to each concept corresponds an *elementary tree*; order of the branches of these trees is subject to collective convention, as well as the overt name of the concept.

The agents' goal is to establish conventions for naming the concepts they perceive and for ordering these names in the output message using a parse tree representation, or *derived tree*. This is obtained from the derivation tree through the combination of elementary trees for each concept. Agents randomly engage in dialogs (Figure 2) and are rewarded if the message sent can be parsed by the receiver and matched with the observed situation or part of it. Initially, the agents have no words and no elementary parse trees associated to any concept, but they are allowed to create new ones randomly. Also, when an agent tries to understand a message, it is allowed to guess one or more words, thus enriching its knowledge.
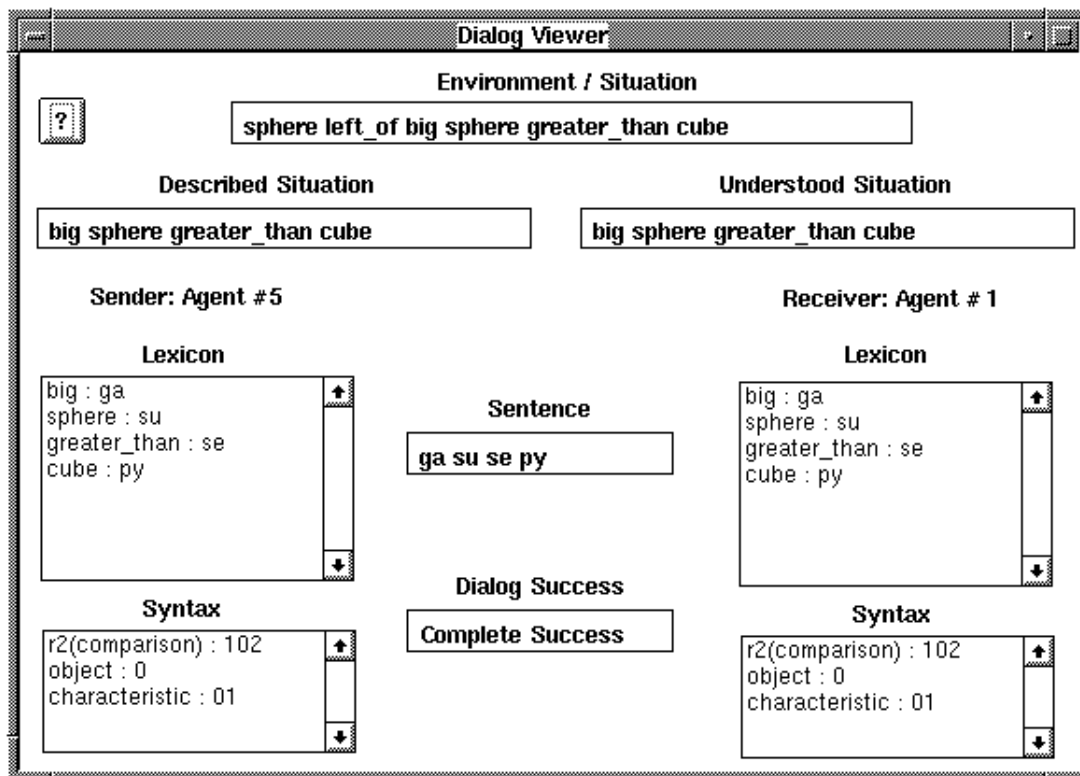
Figure 2. A sample dialogue between two agents

Convergence of the concept names and elementary parse trees among the population is proven through a learning protocol. For instance, an average of 20,000 dialogs is necessary for 5 agents to establish conventions, when learning is incremental, i.e., the situations' complexity is progressively increased. Other situations have been tested: inserting a new agent, allowing partial descriptions (these converge too), non-incremental learning, population mixing (these fail to converge in a reasonable period).

Further work concerns better command of the learning process, in order to direct communication conventions towards natural (French) ones, using one or more instructor agents to provide sample sentences.

4. Building a network of semantic proximity from simulations of verbal experiences

Instead of giving the agents symbols for concepts, we have studied how meaning can be learned from perceptual experiences. Here, we want to show how a semantic network can be built from simulation of verbal experiences —(Gruselle 1998). An individual agent observes, without further interaction, a series of natural language discourses corresponding to various "situations". The agent's task is to build a semantic network from the texts, which stores both the relationships between words, and between words and situations. The network is divided in two levels, one corresponding to the situations (situation-nodes) and the other to the words used to describe these situations (word-nodes) as in Figure 3. The network is built in two recurrent phases.
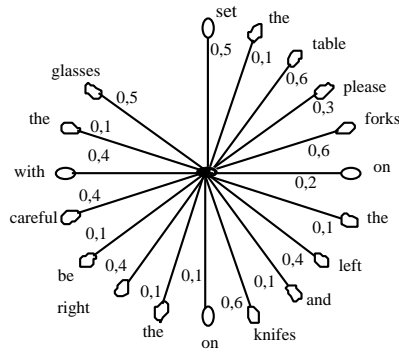
Figure 3. Situation node at the experience level

In the *acquisition* phase, the system collects data from the series of discourses, segmented in *situations.* In each situation, each (open class) word is assigned a weight representing its importance in the situation. The system further operates on these *situation representations*, which form the network's first level, or the *experience level* (cf. Figure 4). Links between nodes have a weight, which stands for the facility to spread activation between the nodes. At this level, situation-nodes and word-nodes alternate; they are assigned an activity level which represents their *pregnancy* when a new situation is processed.
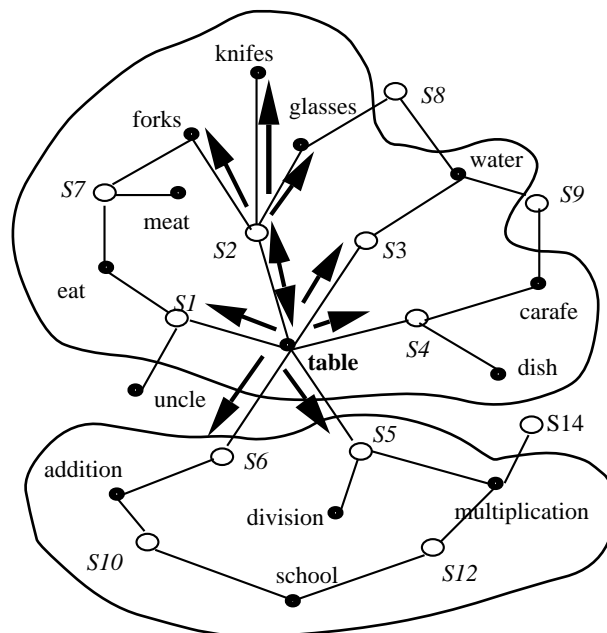

Figure 4. Two conceptual clusters at the experience level

The system selects only a few words among the current situation before adding it to the network. This filter is the computational counterpart of the working memory with a limited capacity described by psychologists. In order to select these words the system uses the weighted links between existing word-nodes and several other criteria. For instance, the familiarity criterion increases the weights of links corresponding to words which are already in the network. Another criterion (the *curiosity*) allows the system to store new words in its short term memory, in such a way that not too much of them be present (since a too new situation cannot be related to previous knowledge, and therefore cannot be understood). Then, the new nodes (situation- and word-nodes) are added to the network and activity is spread from the new situation-node, which modifies the level of activity of other nodes. Thus, the system computes new weights between these nodes using a Hebbian law.

During the *structuring phase*, the system builds the network's *conceptual level* from the level of experience. This phase occurs once enough situations were added to the network. The system spreads activity from each couple <word-node, situation-node> around a word-node. This allows the system to determine a pattern of nodes and weights representing the different meanings of the word. For example, in Figure 4, six initially distinct sets of points (one around each situation) will be clustered in two general sets (S1, 2, 3, 4 and S5, 6). This allows the system to disambiguate contextually the meaning of a word, using the activities of the other words belonging to the same cluster in the constructed network.

Then the system then creates the corresponding *conceptual nodes* at the conceptual level, and links to it each node of the activation pattern, also using weighted links. The links between conceptual nodes represent the semantic proximity between meanings (Figure 5).
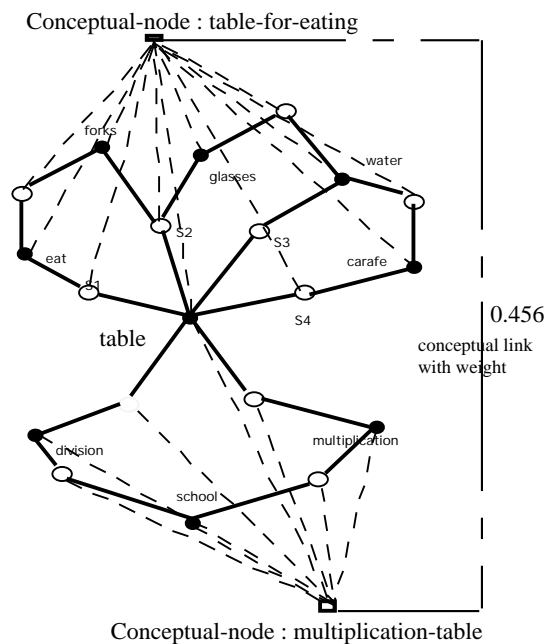

Figure 5. The emergence of two conceptual nodes

In order to implement the system, the most adequate propagation laws and their application conditions had to be determined experimentally, as well as the thresholds for the clustering of words and concepts. Various simulations have shown the feasibility of the system, using a French newspaper corpus.

The model thus elaborated constitutes a "long term memory", and may explain some of its relations to short term memory. Depending on their activity level, nodes may be considered as belonging only to long term memory (low level), or, when adequately activated, they enter the short term memory too. This provides a convenient articulation, within the CARAMEL model, with the attention states and the automatic/reflective processes, as described elsewhere (Sabah 1995, 1997a).

5. Perspective: towards a synthesis

In this paper, we have described three levels of experimentation for the automatic acquisition of (pre)linguistic capacities. We have first shown how categorisation, values and motor control could be implemented in a very simple agent. We then studied the emergence of simple syntactic conventions in a group of agents which are capable of describing situations to each other. Further work aims to replace their built-in categorisation with a more realistic

mechanism inspired from both the first and the third model. Indeed, the last experiment has shown how meaning can be learned from perceptual experiences (restricted here to discourses), and how the symbolic aspects of words may be acquired.

There are at least three models of language emergence in nature: the evolution of language, the acquisition of language, and creolisation (Bickerton 1990). Even if the first and second provide inspiration to our modelling, it is the third that will probably lead us to an integrative model. Indeed, in Bickerton's view, creole children probably acquire lexical semantics by repetition and association (our third model), but they *develop* syntax despite their parents' inability to use it. In terms of agents, instructor agents could thus lack syntactic abilities, and provide the communication code with a lexicon, acquired by "infant agents", which further establish syntactic conventions using "innate" syntactic devices.

Clearly a lot of work remains to be done to integrate this in a coherent system within which genuine semantics can emerge, but we hope to have been convincing enough to show with experimental evidence that this is not as far from our modelling capabilities as some could claim.

6. References

Allexandre Christophe and Andrei Popescu-Belis 1998, Emergence of Grammatical Conventions in an Agent Population Using a Simplified Tree Adjoining Grammar, Proceedings Third International Conference on Multi-Agent Systems (ICMAS'98), IEEE Computer Society, Paris, volume 1/1, p. 383-384.

Baars Bernard 1988, *A cognitive theory of consciousness*, Cambridge University Press, Cambridge.

Bickerton Derek 1990, *Language and Species*, The University of Chicago Press, Chicago.

Edelman Gerald 1989, *The remembered present: a biological theory of consciousness*, Basic Books, New York.
Edelman Gerald 1992, *Biologie de la conscience*, Editions Odile Jacob, Paris.

Gruselle Jean-Pierre 1998, A Cognitive Sciences System for Symbol Grounding, Proceedings Fifth International Conference of the International Society for Knowledge Organisation, ISKO'5, 25-28 Août, Lille.
Harth Erich 1993, *The creative loop; how the brain makes a mind*, Addison-Wesley, New York.

Minsky Marvin 1985, *The Society of Mind*, Simon and Schuster, Nex York.

Popescu-Belis Andrei 1997, An Adaptive Multi-Agent System Based on 'Neural Darwinism', Proceedings First International Conference on Autonomous Agents, ACM Press, Marina del Rey, CA, p. 484 - 485.
Reeke G. N., L. H. Finkel, O. Sporns and G. M. Edelman 1990, Synthetic neural modeling: A multilevel approach to the analysis of brain complexity, *Signal and Sense : Local and Global Order in Perceptual Maps*, Wiley-Liss, New York, p. 607 - 707.

Sabah Gérard 1995, Natural Language Understanding and Consciousness, Proceedings AISB — workshop on "Reaching for Mind", Sheffield.

Sabah Gérard 1997a, Consciousness: a Requirement for Understanding Natural Language, *Two sciences of mind*, John Benjamins, Amsterdam, p. 361-392.

Sabah Gérard 1997b, The "Sketchboard": A Dynamic Interpretative Memory and its Use for Spoken Language Understanding, Proceedings Eurospeech'97, Rhodes, volume 2/5, p. 617-620.

Sabah Gérard and Xavier Briffault 1993, Caramel : a Step towards Reflexion in Natural Language Understanding systems, Proceedings IEEE International Conference on Tools with Artificial Intelligence, Boston, p. 258-265.