Frédéric Landragin

CNRS – LaTTICe Laboratory
1 rue Maurice Arnoux
F-92120 Montrouge
France

frederic.landragin@linguist.jussieu.fr
www.linguist.jussieu.fr/~flandragin/

1 Research Interests

My research interests lie in the area of **formal semantics and pragmatics of dialog**, with the aim to provide formal and cognitive models for speech understanding and dialog management, which can be exploited when designing speech-based dialog systems. A first aspect is the study of **referring phenomena**, particularly when a visual scene is displayed on a touch screen and allows the user to employ deictic gestures together with speech. This research topic includes linguistic and multimodal concerns, and is linked to **multimodal fusion**. A second aspect is the study of **speech acts in dialog**, with a special focus on complex speech acts, i.e., **indirect speech acts** (meaning something instead of something else) and **composite acts** (meaning several things simultaneously). One objective is to bring together complex acts processing and collaborativeness in task-oriented dialog. Another objective is to extend this work to gesture, with the notion of **gesture act** and the **fusion of dialog acts**.

1.1 Referring in multimodal dialog

My study of referring phenomena follows a linguistic method, with the analysis of the use of various referring expressions (*"the N"*, *"this N"*, *"that"*, *"an N"*) and their possible associations with gestures. Cognitive aspects are taken into account, in particular when introducing the way the user sees the visual context (*how many perceptual groups does he perceive?*). The two notions I proposed and detailed are the notion of salience (Landragin *et al.*, 2001) and the notion of multimodal reference domain (Landragin, 2006). A reference domain is a kind of sub-context (or focus space) where the referring action takes place, and is exploited as a pattern during multimodal fusion. This notion and the related mechanisms have been applied to haptic interaction (Landragin *et al.*, 2002b). The notion of salience has also an interest for anaphora resolution, and, with multimodal concerns, for exophora resolution (Landragin, 2007). My main results take the form of a unified formal and cognitive model for reference resolution.

1.2 Complex speech acts in multimodal dialog

My study of speech acts in dialog is guided by Relevance Theory (see Landragin *et al.*, 2002a) as well as by multimodal concerns. How can a system react to an indirect speech act, to a composite one, or to a combination of a speech act and a gesture act? With the analysis of examples such as "*how long does it take to go by this way which seems to be the shortest one?*" I proposed some arguments and parameters for acts decomposition and treatment in a collaborative manner (Landragin, 2005). When confronting speech to illustrative and iconic gestures, the problem lies in the fusion of a speech semantic item accompanied by its pragmatic force, with a gestural semantic item that can own another pragmatic force or another semantic direction. The dialog domain and task are here crucial to fuse semantic contents and pragmatic forces.

1.3 Design of multimodal dialog systems

Some semantic and pragmatic aspects of my work have been integrated to several development projects. As a first participation to the design of a multimodal dialog system, my work in ACTS COVEN project (*COllaborative Virtual ENvironments*, 1995-1998) was limited to the improvement of the reference resolution module, with the incorporation of an automatic detection of visual salience. The results proved the importance of managing a salience model in speech-based systems. Within IST MIAMM project (*Multidimensional Information Access using Multiple Modalities*, 2001-2004), I focused on the importance of properly managing reference domains (Landragin and Romary, 2004). Then, in the framework of IST OZONE project (O$_3$, *Offering an Open and Optimal roadmap towards consumer oriented ambient intelligence*, 2001-2004), I participated to the design of a multimodal system involving a touch screen for a travel information task. It was an opportunity to test the efficiency of reference domains and the related software architecture (Landragin *et al.*, 2004). More recently, in ITEA EMODE project (*Enabling Adaptive Multimodal Interfaces*, 2005-2007), I used to coordinate some research work on the use of model-driven

engineering methods for the design of multimodal dialog systems. This idea was a reaction to the increasing complexity of the phenomena we want to address and the increasing complexity of software production. This is an ongoing project and results are not yet available.

2 Future of Spoken Dialog Research

In my view, implementations of semantic and pragmatic modules in spoken dialog systems are still far away from relevant linguistic theories. Then, the field of spoken dialog research for the next years may include the formalization of these theories (Relevance Theory, for instance). A lot of work is still to be done in syntax-semantics, syntax-pragmatics, and semantics-pragmatics interfaces, and in their computational adjustments. Concerning multimodal systems, the current increasing diversity of devices and phenomena (emotions, for instance) we want to process has a negative consequence on evaluation. I think that one objective for the generation of young researchers can be the design of a unified methodology for the evaluation of multimodal dialog systems.

3 Suggestions for discussion

Here are three possible topics for discussion:

- Modeling deep understanding abilities: How can we combine formal semantic theories (e.g., DRT, SDRT) to cognitive aspects such as attention or salience, which put forward some particular semantic items?
- Modeling reasoning abilities: How can a system manage indirect and composite speech acts together with implicitness, intentions, beliefs and current task? In this way, how can we formalize the notion of relevance from Relevance Theory (if considered necessary)?
- Evaluating multimodal systems: Global competition, universal metrics for comparing disparate modalities and systems.

References

Frédéric Landragin. 2007. Taking Situational Factors into Account when Resolving Anaphora: an Approach based on Events and Salience. *6th Discourse Anaphora and Anaphor Resolution Colloquium (DAARC'07)*, Lagos, Portugal.

Frédéric Landragin. 2006. Visual Perception, Language and Gesture: A Model for their Understanding in Multimodal Dialogue Systems. *Signal Processing*, 86(12), Elsevier, The Netherlands.

Frédéric Landragin. 2005. Indirect Speech Acts and Collaborativeness in Human-Machine Dialogue Systems. *First International Symposium on the Exploration and Modelling of Meaning (SEM-05)*, Biarritz, France.

Frédéric Landragin, Alexandre Denis, Annalisa Ricci, and Laurent Romary. 2004. Multimodal Meaning Representation for Generic Dialogue Systems Architectures. *Fourth International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon, Portugal.

Frédéric Landragin, and Laurent Romary. 2004. Dialogue History Modelling for Multimodal Human-Computer Interaction. *Eighth Workshop on the Semantics and Pragmatics of Dialogue (SemDial-Catalog'04)*, Barcelona, Spain.

Frédéric Landragin, Antonella De Angeli, Frédéric Wolff, Patrice Lopez, and Laurent Romary. 2002a. Relevance and Perceptual Constraints in Multimodal Referring Actions. In: Kees van Deemter, and Rodger Kibble (Eds.), *Information Sharing*, CSLI Publications, Stanford, CA.

Frédéric Landragin, Nadia Bellalem, and Laurent Romary. 2002b. Referring to Objects with Spoken and Haptic Modalities. *Fourth IEEE International Conference on Multimodal Interfaces (ICMI'02, Pittsburgh, PA)*, IEEE CS Press, Los Alamitos, CA.

Frédéric Landragin, Nadia Bellalem, and Laurent Romary. 2001. Visual Salience and Perceptual Grouping in Multimodal Interactivity. *First International Workshop on Information Presentation and Natural Multimodal Dialogue*, Verona, Italy.

Biographical Sketch

Frédéric Landragin has completed his PhD in 2003, at the LORIA laboratory and the University of Nancy, France, in the area of computer science and computational linguistics. After some months as a post-doctoral engineer, he joined the THALES Research and Technology laboratory at Palaiseau, France. In 2006, he got a permanent researcher position at the CNRS, the French National Center for Scientific Research, and he is currently working at the LaTTICe laboratory at Paris. His research program for the next years is structured around the notion of salience in dialog.