



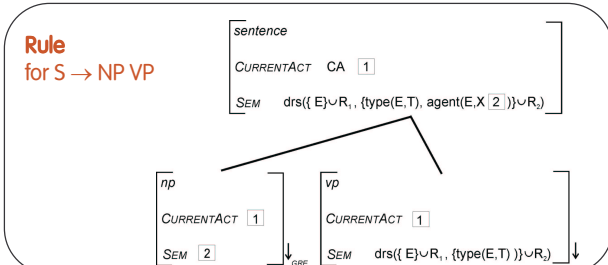
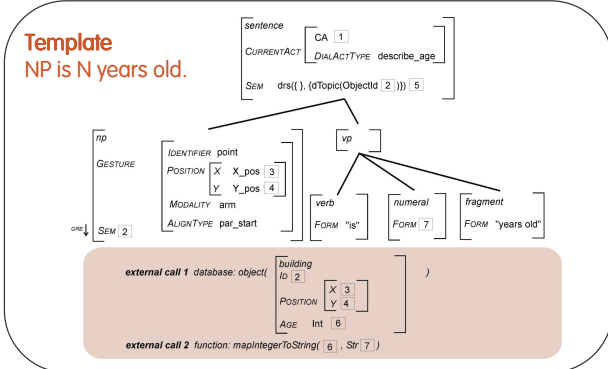
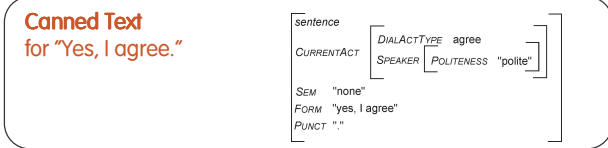
# The Multimodal Natural Language Generator – MNLG

Paul Piwek – ITRI

## What is the MNLG?

The MNLG is a platform for multimodal generation. The input is a series of communicative acts. For each act, the input specification includes the dialogue act type, semantics (in terms of DRT), speaker (including information about personality), addressee and common ground at the outset of the act. The output of the MNLG is the input representation enriched with specifications for the realization of the communicative acts by means of language and gestures. Both input and output are represented in the XML-compliant RRL language ([3]). Generation consists of tree construction: the goal is to construct a tree with only terminal leaves that has the input specification as its mother node. Selection of trees for expansion from a tree repository relies on unification and supports seamless integration of canned text, templates (cf. [5]) and generation rules (see [2] for details) and personalized referring expressions generation (see [4] for details).

## MNLG Tree Repository Examples



## Applications



The MNLG is one of the components of the NECA eShowroom application. This application generates dialogues between a virtual car seller and buyer. Users can control the personality, role and interests of the interlocutors. NECA was an EU project involving ITRI (Brighton), DFKI, ÖFAI, Sysis AG and IPUS (Saarland).

## EPOCH iGuide

The MNLG was used in the EPOCH EU Network of Excellence for a demonstrator in the cultural heritage domain. It is the generation component for a virtual tour guide for the world heritage site of Wolfenbüttel. The demonstrator subproject involved ITRI (Brighton) and the Universities of Braunschweig and East Anglia.



## Implementation

The MNLG has been implemented in Sicstus Prolog, using Profit [1] for representing feature structures. The MNLG is fully documented and makes extensive use of the Sicstus module system to facilitate maintenance and reusability.

## Performance

For the eShowroom application, with a total of 138 generation trees, the generation time for individual dialogue acts ranges from 1/100 to 1/10 of a second, depending on the generation strategy (select-first-match versus generate-first-N-and-select).

## Acknowledgements

Thanks are due to my colleagues in the NECA project, especially, Kees van Deemter who coordinated the NECA project at ITRI. Integration of and resources development for the MNLG for the EPOCH iGuide application was carried out by Ondrej Pacovsky, Michel Génèreux and Karina Rodriguez.

## References

- [1] G. Erbach (1995). PROFIT 1.54 User's guide. University of the Saarland, December 3, 1995.
- [2] P. Piwek (2003). A flexible pragmatics-driven language generator for animated agents. *Proc. of EACL (Research Notes)*, Budapest, Hungary, 2003.
- [3] P. Piwek, B. Krenn, M. Schröder, M. Grice, S. Baumann and H. Pfler (2002). RRL: A Rich Representation Language for the Description of Agent Behaviour in NECA. *Proc. of AAMAS workshop Embodied Conversational Agents - let's specify and evaluate them!*, Bologna.
- [4] P. Piwek, J. Masthoff and M. Bergensträhle (to appear). Reference and Gestures in Dialogue Generation: Three Studies with Embodied Conversational Agents. In: *Proc. of AISB Virtual Social Agents Symposium*, University of Hertfordshire, Hatfield, UK, 2005.
- [5] K. van Deemter, E. Krahrmer, and M. Theune (1999). Plan-based vs. template-based NLG: A false opposition? *Procs of K999 Workshop 'May I Speak Freely?'*, DFKI, Saarbrücken, 1999.