University of Pennsylvania

Title: Natural Language Research
Principal Investigators: Aravind Joshi, Mitch Marcus (Co-PI), Mark Steedman (Co-PI), Bonnie Webber (Co-PI)

Project Summary: The main objective is the development of natural language interactive systems. Based on our earlier work on cooperative systems for interaction through language, further significant contributions were made in the following major areas: (1) development of linguistic formalisms (syntax, semantics, and pragmatics) and the corresponding algorithms; (2) extensions of deterministic parsing; (3) some aspects of discourse coherence; (4) some aspects of tense and anaphora; (5) a theory of response structure and response generation; and (6) initiation of research in information carried by intonation and its role in spoken language systems.

The basic research and system development is expected to lead to: (1) characterization of information carried by (a) syntax, semantics, and discourse structure, (b) their relationship to information carried by intonation, (c) development of methods for applying their information in generation and understanding; (2) development of architecture for integration of utterance planning, lexical, syntactic, and intonational choice; (3) development of incremental strategies for applying syntactic, semantic and pragmatic knowledge in understanding and generating language, especially spoken language.

Summary of accomplishments:
Accomplishments before FY 88:

- Preliminary design of a modular architecture for system integration for cooperative responses;
- Integration of RTM (Rus-Text-Mumble) with DECTALK, providing output for RTM;
- Completed pilot system for converting machine efficient proofs as input to natural language generation systems.

Accomplishments for FY88:

- Completed the design and development of a flexible architecture for a multi-functional cooperative response system;
- Development of linguistic formalisms for integrating syntax and semantics. Preliminary implementation of the algorithms. Established convergence of several formalisms;
- Characterization of tense and anaphora in discourse;
- Extension of deterministic parsing;
- Development of an architecture for interaction between the strategic and tactical components of a generation system.

Objectives for FY89:

- Characterization of the information carried by syntax, semantics, discourse, and intonation and its use in generation and understanding;
- Further development and implementation of deterministic parsing systems;
- Pilot implementation of the multi-functional cooperative response systems;
- Design and integration of a sizable grammar in the TAG parsing algorithms. Development of new strategies based on 'lexicalized' grammars;
- Development of incremental strategies for language processing (understanding and generation), especially in spoken language.