March 1992

CLiFF Notes: Research In Natural Language Processing at the University of Pennsylvania

Faculty & Graduate Students
University of Pennsylvania

University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-92-22.

Edited by Barbara Di Eugenio and Libby Levison.

CLiFF Notes: Research In Natural Language Processing at the University of Pennsylvania

Abstract

The Computational Linguistics Feedback Forum (CLIFF) is a group of students and faculty who gather once a week to discuss the members' current research. As the word "feedback" suggests, the group's purpose is the sharing of ideas. The group also promotes interdisciplinary contacts between researchers who share an interest in Cognitive Science.

There is no single theme describing the research in Natural Language Processing at Penn. There is work done in CCG, Tree adjoining grammars, intonation, statistical methods, plan inference, instruction understanding, incremental interpretation, language acquisition, syntactic parsing, causal reasoning, free word order languages, ... and many other areas. With this in mind, rather than trying to summarize the varied work currently underway here at Penn, we suggest reading the following abstracts to see how the students and faculty themselves describe their work. Their abstracts illustrate the diversity of interests among the researchers, explain the areas of common interest, and describe some very interesting work in Cognitive Science.

This report is a collection of abstracts from both faculty and graduate students in Computer Science, Psychology and Linguistics. We pride ourselves on the close working relations between these groups, as we believe that the communication among the different departments and the ongoing inter-departmental research not only improves the quality of our work, but makes much of that work possible.

Comments

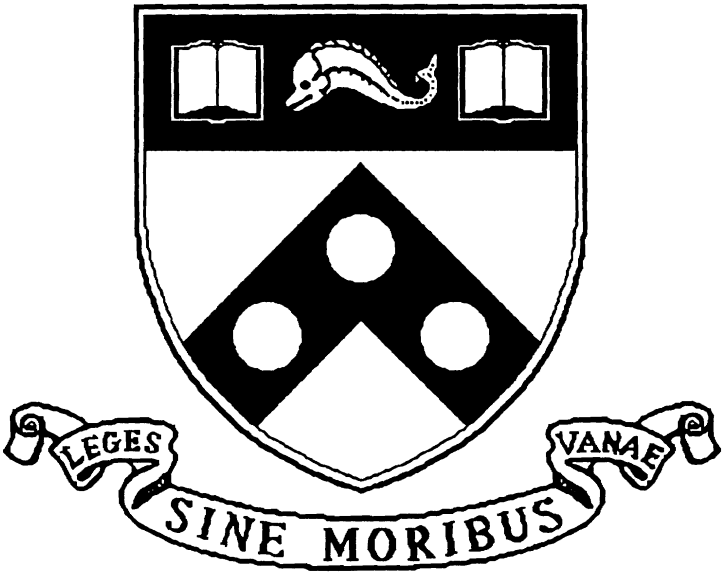
University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-92-22.

Edited by Barbara Di Eugenio and Libby Levison.

CLiFF Notes
Research In Natural Language Processing At The University of Pennsylvania

MS-CIS-92-22
LINC LAB 218

Faculty
and
Graduate Students



University of Pennsylvania
School of Engineering and Applied Science
Computer and Information Science Department
Philadelphia, PA 19104-6389

March 1992

CLiFF Notes

Research in Natural Language Processing at the University of Pennsylvania

Annual Report: Spring 1992, No. 2

Technical Report MS-CIS-92-22
LINC Lab 217

Department of Computer and Information Science
University of Pennsylvania
Philadelphia, PA 19104-6389

editors: Barbara Di Eugenio & Libby Levison

Contents

1 Introduction

The Computational Linguistics Feedback Forum (CLiFF) is a group of students and faculty who gather once a week to discuss the members' current research. As the word "feedback" suggests, the group's purpose is the sharing of ideas. The group also promotes interdiscipinary contacts between researchers who share an interest in Cognitive Science.

There is no single theme describing the research in Natural Language Processing at Penn. There is work done in CCG, Tree adjoining grammars, intonation, statistical methods, plan inference, instruction understanding, incremental interpretation, language acquisition, syntactic parsing, causal reasoning, free word order languages, ... and many other areas. With this in mind, rather than trying to summarize the varied work currently underway here at Penn, we suggest reading the following abstracts to see how the students and faculty themselves describe their work. Their abstracts illustrate the diversity of interests among the researchers, explain the areas of common interest, and describe some very interesting work in Cognitive Science.

This report is a collection of abstracts from both faculty and graduate students in Computer Science, Psychology and Linguistics. We pride ourselves on the close working relations between these groups, as we believe that the communication among the different departments and the ongoing inter-departmental research not only improves the quality of our work, but makes much of that work possible.

1.1 Addresses

Contributors to this report can be contacted by writing to:

contributor's name
department
University of Pennsylvania
Philadelphia, PA 19104

In addition, a list of publications from the Computer Science department may be obtained by contacting:

Technical Report Librarian
Department of Computer and Information Science
University of Pennsylvania
200 S. 33rd St.
Philadelphia, PA 19104-6389
(215)-898-3538
holland@central.cis.upenn.edu

1.2 Funding Agencies

Support for members of the CLiFF group comes from several sources including: AFOSR, Benjamin Franklin Partnership Program (ATC, Southeastern PA), DARPA, Deere and Company, FMC Corporation, General Electric (Schenectady), Lockheed Engineering and Management Services (NASA Johnson Space Center), Martin-Marietta Denver Aerospace, MOCO Inc., NASA Goddard (through University of Iowa UICR), the National Library of Medicine, the National Science Foundation, NSF CISE Grant CDA88-22719, Siemens Research, and ARO Grant DAAL03-89-C-0031 including participation by the U.S. Army Human Engineering Laboratory, Natick Laboratory, TACOM, and NASA Ames Research Center.

Acknowledgements

Thanks to all the contributors.

Comments about this report can be sent to:

Barbara Di Eugenio or Libby Levison
Department of Computer and Information Science
University of Pennsylvania
200 S. 33rd St.
Philadelphia, PA 19104-6389
dieugeni@linc.cis.upenn.edu or libby@linc.cis.upenn.edu

Charles Augustine

Department of Computer and Information Science
augustin@gradient.cis.upenn.edu

Conversion of neural network parser for German

Keywords: Neural Networks, Machine Translation, Parsing

PARSEC is a natural language parsing system implemented using neural networks. It was developed by Ajay Jain at Carnegie Mellon University [1]. It was originally developed as a standalone system for parsing English, and was later adapted to be included in the JANUS speech to speech translation system [2]. During the summer of 1991, I converted PARSEC to parse German as part of a project to convert the JANUS system to take German rather than English speech as its input.

The task domain was a set of 12 dialogs dealing with registering for a conference. After two months of work, German PARSEC could process the 53 sentences in the first three dialogs of the conference registration task with eight errors, half of which result from a change in the representation of numbers, which can be easily fixed.

The PARSEC parser represents words as a set of about 40 binary features. The features include simple part of speech tags and agreement information such as number, gender, and tense. The parser groups the words into phrases, and phrases into clauses, assigning role labels to the words and phrases.

At the beginning of this project the most interesting open question was probably whether or not this parser depended significantly on the structure of English, and would not be able to be trained to parse German. Of course, some modification was obviously needed in the coding of lexical features and role labels, however, the results of the summer work showed that the basic architecture did not have to change, at least in order to parse the relatively simple sentences in the first three of the dialogs. More work seems to be needed in refining the lexical features, as well as tole and phrase labeling, and I have been working on these changes.

References

[1] Ajay N. Jain and Alex H. Waibel. Robust Connectionist Parsing of Spoken Language. In *IEEE Proceedings of the 1990 International Conference on Acoustics, Speech, and Signal Processing.*

[2] Alex Waibel, Ajay N. Jain, Arthur E. McNair, Hiroaki Saito, Alexander G. Hauptmann, Joe Tebelskis. Janus: A Speech-to-speech Translation System Using Connectionist and Symbolic Processing Strategies. In *IEEE Proceesings of the 1991 International Conference on Acoustics, Speech, and Signal Processing.*

Norman I. Badler

Professor, Department of Computer and Information Science
badler@central.cis.upenn.edu

Graphical Behaviors and Animated Agents

Keywords: Graphics, Animation

One concern of the Computer Graphics Research Lab is in simulating human task behavior and understanding why the visualization of the appearance, capabilities and performance of humans is so challenging. Our research has produced a system, called *Jack*™ , for the definition, manipulation, animation and human factors analysis of simulated human figures. *Jack* permits the envisionment of human motion by interactive specification and simultaneous execution of multiple constraints, and is sensitive to such issues as body shape and size, linkage, and plausible motions. Enhanced control is provided by natural behaviors such as looking, reaching, balancing, lifting, stepping, walking, grasping, and so on. Although intended for highly interactive applications, *Jack* is a foundation for other research.

The very ubiquitousness of other people in our lives poses a tantalizing challenge to the computational modeler: people are at once the most common object around us, and yet the most structurally complex. Their everyday movements are amazingly fluid, yet demanding to reproduce, with actions driven not just mechanically by muscles and bones but also cognitively by beliefs and intentions. Our motor systems manage to learn how to make us move without leaving us the burden or pleasure of knowing how we did it. Likewise we learn how to describe the actions and behaviors of others without consciously struggling with the processes of perception, recognition, and language.

Present technology lets us approach human appearance and motion through computer graphics modeling and three-dimensional animation, but there is considerable distance to go before purely synthesized figures trick our senses. We seek to build computational models of human-like figures which manifest animacy and convincing behavior. Towards this end, we

- Create an interactive computer graphics human model.
- Endow it with reasonable biomechanical properties.
- Provide it with "human-like" behaviors.
- Use this simulated figure as an agent to effect changes in its world.
- Describe and guide its tasks through natural language instructions.

There are presently no perfect solutions to any of these problems; ultimately, however, we should be able to give our surrogate human directions that, in conjunction with suitable symbolic reasoning processes, make it appear to behave in a natural, appropriate, and intelligent fashion. Compromises will be essential, due to limits in computation, throughput of display hardware, and demands of real-time interaction, but our algorithms aim to balance the physical device constraints with carefully crafted models, general solutions, and thoughtful organization.

† *Jack* is a trademark of the University of Pennsylvania.

The *Jack* software is built on Silicon Graphics Iris 4D workstations because those systems have the 3-D graphics features that greatly aid the process of interacting with highly articulated figures such as the human body. Of course, graphics capabilities themselves do not make a usable system. Our research has therefore focused on software to make the manipulation of a simulated human figure easy for a rather specific user population: human factors design engineers or ergonomics analysts involved in visualizing and assessing human motor performance, fit, reach, view, and other physical tasks in a workplace environment. The software also happens to be quite usable by others, including graduate students and animators. The point, however, is that program design has tried to take into account a wide variety of physical problem-oriented tasks, rather than just offer a computer graphics and animation tool for the already computer-sophisticated or skilled animator.

As an alternative to interactive specification, a simulation system allows a convenient temporal and spatial parallel "programming language" for behaviors. The Graphics Lab is working with the Natural Language Group to explore the possibility of using natural-language instructions (such as those found in assembly or maintenance manuals) to drive the behavior of our animated human agents. (See other work of the AnimNL group: Baldwin, Di Eugenio, Geib, Levison, Moore, Webber and White.)

Even though *Jack* is under continual development, it has nonetheless1 already proved to be a substantial computational tool in analyzing human abilities in physical workplaces. It is being applied to actual problems involving space vehicle inhabitants, helicopter pilots, maintenance technicians, foot soldiers, and tractor drivers. This broad range of applications is precisely the target we intended to reach. The general capabilities embedded in *Jack* attempt to mirror certain aspects of human performance, rather than the specific requirements of the corresponding workplace.

We view the *Jack* system as the basis of a virtual animated agent that can carry out tasks and instructions in a simulated 3D environment. While we have not yet fooled anyone into believing that the *Jack* figure is "real," its behaviors are becoming more reasonable of its repertoire of actions more extensive. When interactive control becomes more labor intensive than natural language instructional control, we will have reached a significant milestone toward an intelligent agent.

Breck Baldwin

Department of Computer and Information Science
breck@linc.cis.upenn.edu

Determining What Individuals are in a Discourse Model

Keywords: Discourse Processing, Anaphora Resolution, Instruction Understanding

My research involves determining what individuals exist in the discourse model from the perspective of reference resolution. In particular I am interested in examining the status of individuals in the context of actions which are described by instructions.

Consider the following instructions:

1. *Mold a house from playdough. Squish (the house/the playdough/*them).*

After interpreting the first sentence, there are two new discourse entities, the house and the playdough. Therefore we would think that we can refer to them with a plural pronoun, such as they. However, *squish them* is infelicitous. I consider the infelicity of "Squish *them" to be a diagnostic which indicates that there is only one squishable entity in the discourse model after the squishing event. This is immediately problematic given that the antecedents for "the house" and "the playdough" appear to be available, and it is fairly clear that the NPs do not have the same antecedent. A possible argument for the distinctness of the antecedents is that "the house is beautiful' does not entail that "the playdough is beautiful".

Given that there appears to be distinct discourse markers for the playdough and the house, why isn't plural anaphora felicitous? A reasonable explanation is that in some sense the markers for "the house" and "the playdough" are the same individual, but distinct discourse markers. How to represent and reason about such a many-to-one-mapping is the task of my thesis.

At this point I am not very far along in having a theory which solves the problem, but I am exploring with Bonnie Webber possible explanations that involve extensions to Fred Landman's theory of groups.

References

[1] Fred Landman. "Groups, I" *Linguistics and Philosophy*, 12, 1989.

[2] Fred Landman. "Groups, II" *Linguistics and Philosophy*, 12, 1989.

[3] B. L. Webber *A Formal Approach to Discourse Anaphora.* Ph.D. thesis, Harvard University. 1978.

Eric Brill

Department of Computer and Information Science
brill@unagi.cis.upenn.edu

Automatic Acquisition of Syntactic Knowledge

Keywords: Language Acquisition, Automated Corpus Annotation

Learning a Language

Over the past couple of years, we have run a number of experiments that point toward the feasibility of extracting syntactic information from a large corpus of text using distributional analysis. Many of our techniques are variants on techniques originally proposed by Zellig Harris [4]. Progress has been made in designing procedures able to automatically learn morphology, word classes and phrase structure [1, 2]. These procedures begin with only minimal knowledge of the language they are learning.

Toward Automatic Corpus Annotation

Large corpora annotated with part of speech and phrase structure information are now being used by many researchers in computational linguistics. We hope to develop our phrase structure learner to allow for the fast and automatic bracketing of large corpora. We have written a rule-based part of speech tagger that automatically acquires its rules. Contextual tag disambiguation information is captured in a small number of simple rules (70-90) [3]. The stochastic taggers that are currently being used to tag text capture contextual information in thousands of contextual probabilities, while their performance is comparable to our rule-based tagger.

References

[1] Brill, Eric. (1991) Discovering the lexical features of a language. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics, Berkeley, CA.*

[2] Brill, E. and Marcus, M. (1992) Automatically Acquiring Phrase Structure Using Distributional Analysis. Proceedings of the DARPA Speech and Natural Language Workshop, Morgan Kaufmann.

[3] Brill, Eric. (1992) A Simple Rule-Based Part of Speech Tagger. Proceedings of the Third Conference on Applied Computational Linguistics (ACL), Trento, Italy.

[4] Harris, Zellig. (1951) *Structural Linguistics.* Chicago: University of Chicago Press.

Sharon Cote

Department of Linguistics
cote@linc.cis.upenn.edu

Null Arguments in English

Keywords: Null Anaphora, Implicit Arguments, Discourse Models

English does not allow the extensive use of null anaphora found in pro-drop or so-called "discourse-oriented" languages. As a result, little effort has been devoted to examining the nature of the nonovert (or "null") arguments that are a part of the language or to determining the effect that using a null argument has on English discourse. This research is an attempt to describe and explain the grammatical nature of null arguments in English, to compare these null arguments with those found in other languages, and to incorporate them into a model of local discourse structure.

There are a number of different types of null arguments in English and it is impossible for one explanation to encompass them all. For example, though we can safely state that in most cases transitive verbs in English occur with an overt NP (or its trace) in object position, there are nonetheless some types of apparently transitive verbs which do, under the right circumstances, occur without overt objects. These null objects have varied discourse properties, behaving to a greater or lesser degree like pronominals depending on the type of verb. (For a discussion of some discourse properties of two types of these null objects, see [1].) Syntactically, none of them behave like pronominals. In particular, although like pronouns, they must be free in their local domain; they are unable to bind anaphors as pronouns do. (See (1) below.) I argue that some null objects are actually implicit arguments affected in the lexicon. For others, in particular null sentential complements, I am currently considering the possibility that they are the result of discourse deletion.

(1) Harry called *∅ /her about pictures of herself.

Subjects may also be null in English although, with the exception of imperatives, this occurs only in informal speech. Null subjects do behave syntactically like pronouns, as shown in (2). This is consistent with two grammatical explanations. The first, which has been suggested before, is that null subjects are the result of phonological deletion of initial, destressed material. The second is that they are phonologically null pronouns. I argue that the evidence used so far for choosing the former explanation over the latter is inconclusive.

(2) ∅ /I bought myself a new car today!

In languages with pro-drop or freely available null pronominals (called 'zeroes' or 'zero pronominals') these null arguments often play a discourse role like that of destressed, overt pronouns in other languages. What role(s) do the null arguments in English play? I am examining the more specialized roles of English null arguments and incorporating them into CENTERING THEORY [2], a model of local discourse coherence and disambiguation. In the process, I am augmenting the existing theory to handle deictic and event reference, and considering the form and type of information that should be available to a centering model.

References

[1] Cote, Sharon (1992). *Discourse Functions of Two Types of Null Objects in English.* Presented at the 66th Annual Meeting of the Linguistic Society of America, Philadelphia, PA, 9-12 January.

[2] Grosz, B., A. K. Joshi and S. Weinstein (1987). *Toward a Computational Theory of Discourse Interpretation.* Unpublished ms.

[3] Walker, Marilyn, Masayo Iida, and Sharon Cote. Centering in Japanese Discourse. In *COLING-90: Proceedings of the 13th International Conference on Computational Linguistics*, Helsinki, 1990.

[4] Walker, Marilyn, Masayo Iida, and Sharon Cote. Japanese discourse and the process of centering. *In submission*, 1992.

Michel F. DeGraff

Department of Computer and Information Science
degraff@linc.cis.upenn.edu

Creole Grammars and Acquisition of Syntax

Keywords: Haitian Syntax, Principles and Parameters, Language Acquisition, Creole Languages

For my dissertation, I am studying various syntactic properties of the Haitian language within the Principles-&-Parameters framework most prominently brought forward by Noam Chomsky. Haitian emerged around the XVII century from the contact between French and African languages (the majority of them from West-Africa). In addition to its intrinsic descriptive importance, a detailed syntactic study of Haitian will advance our understanding of the still controversial nature of the creolization process.

Thus far, I have scrutinized three intriguing aspects of the syntax of Haitian ('intriguing' in the sense that they seem radically different from corresponding aspects in both the superstrate and substrates): 1) its unexpected *pro-drop* status, i.e, the systematic occurrence of certain clauses with no overt subjects; 2) the patterns through which the language expresses predication; 3) certain constraints on movement of predicates and the availability of a pro-form for a certain class of predicates. Using the analyses I provide for these aspects of Haitian, I investigate possible links between processes of syntax acquisition and the genesis of Creole grammars.

Regarding pro-drop, Haitian and, at least, a few other Creoles, admit phonetically null subjects in contexts where overt subjects may appear. I determine that Haitian exhibits two properties comparable to that of pro-drop languages like Spanish and Italian: 1) lexical expletive subjects are not obligatory with existential, weather, and '*seem*-type' verbs; 2) 'subject' pronouns are not in subject position, but are clitics spelling out the agreement features of the inflection phrase, and identifying an empty pronominal *pro* in subject position. My analysis of Haitian as a null-subject language explains its lack of [COMP–trace] effects. Also, an account for the obligatory co-reference between matrix arguments and null embedded arguments in so-called serial verb constructions (and other properties of these constructions) is provided, using the structures posited for the analysis of INFL.

In analyzing the patterns of predication in Haitian, I focus on the mechanisms that regulate the (non-)appearance of the morpheme *se* in such constructions. In matrix present-tense affirmative clauses, *se* occurs when the predicate position is filled by a proper name or a nominal occurring with a determiner, i.e., when the predicate is a DP. By contrast, AP, PP and bare NP predicates are string-adjacent to their subjects. In other words, no copula appears between subject and AP, PP, and NP predicates. I argue that *se* is not a copula, but a resumptive nominal which spells out the trace left by movement of the base-generated subject from a DP Small Clause to the S-structure subject position (the Spec(IP) position). This spell-out is necessary when the trace is not head-governed. Thus, *se* is used with DP predicates in certain contexts in order to escape a violation of the Empty Category Principle (ECP): in order to save the structure, the non-head-governed trace is spelled-out as a resumptive nominal, *se*, which, being overt, is not subject to the ECP. With DP predicates, *se* is absent only when head-government is assured by a head outside of the DP Small Clause. My analysis of the syntax of predication in Haitian supports the

distinction between NP and DP and leads to interesting cross-linguistic claims about the structure of Small Clauses.

One other property of Haitian revolves around the morpheme *ye*. *Ye* typically occurs with predicate questioning and predicate clefting. *Ye* never appears when the predicate remains in situ. Moreover, *ye* only shows up with non-verbal ([−V]) predicates, i.e., prepositional, nominal and interrogative predicates. I argue that *ye* is a resumptive pro-form for non-verbal predicates. At S-Structure, *ye* spells out the trace left by $\overline{\text{A}}$-movement of a [−V] predicate. I motivate such spell-out from a general ban on predicate traces in Haitian and the observation that $\overline{\text{A}}$-movement of arguments in English, for example, is much freer than that of predicates. Finally, I strengthen the proposal by arguing that my theta-theoretical assumptions explain independent phenomena in the grammar of Haitian.

As explained above, my dissertation research focuses on the syntax of Haitian. But then, why the title "Creole Grammars and Acquisition of Syntax"? Through this title, I hope to characterize my present research as one step in a longer-term investigation of the relationship between the acquisition of syntax and the genesis of Creole languages. Indeed, among the few unquestionable pre-requisites to creolization figure the attempts to acquire a language in a weakly-triggering environment, that is, an environment in which the primary linguistic data from the target language is, in many instances, too variable and/or impoverished for reliable parameter-(re)setting.

In my present work, I discuss how the above analyses for certain aspects of Haitian syntax support a particular working hypothesis about Creolization; namely, the hypothesis that *some* properties of Creole languages have their genesis in universal strategies of language learning. In fact, crucial traits of my analyses were arrived at by paying attention to characteristics which have been claimed to hold of the child's initial grammar, and by testing whether (some approximation of) these characteristics hold of Haitian. As a corollary, I argue that these properties of Haitian cannot be straightforwardly explained by evoking similar properties in the substrates or the superstrate.

Take pro-drop for example, Haitian would correspond to the child's grammar along the corresponding parameter, assuming the results of some major acquisition studies. Moreover, because neither French nor the West-African languages involved in the genesis of Haitian seem to admit null subjects in the contexts in which Haitian does, the pro-drop setting of Haitian might be due exclusively to the creolization process. (Contrarily to colloquial French, Haitian admits of null subjects even in embedded contexts.)

In the same vein, my analysis of the syntax of predication in Haitian implies that this language does not have a copulative verb equivalent to *être* 'be' in French (its superstrate) or *nyé* 'be' in Ewe (one of its Kwa substrates). I have shown that Haitian *se*, although historically related to French *c'est*, manifests syntactic properties crucially different from those of its Romance ancestor. Even more striking is the fact that most other Kwa languages seem to have copulas which are categorical. Interestingly children, in their first stages of acquiring English, use mainly 'copula-less' sentences.

As mentioned above, one of my long-term goals is to extrapolate this sort of acquisition-based approach to other Creole languages, given that the study of one Creole language does not suffice to prove anything about Creole genesis. Also, in future research, I would like to use data from Creolization to both formalize a theory of syntactic markedness and elaborate a model of acquisition.

A theory of markedness based on creolization would predict that the grammar of a Creole language should be less marked than that of the parent languages. Testing this prediction involves a comparison of every Creole language with its parents.

From a Language Acquisition point of view, this theory of markedness lends itself to an attractive model of syntax acquisition. This model would arise from the view that certain subparts of the grammar of a Creole language constitute reasonable first hypotheses of language structure, to be incrementally debugged upon linguistic evidence from the target language. One fascinating question is this: What is it that underlyingly unifies these subparts of Creole grammars which most directly reflect innate grammatical properties and distinguishes them from other parts of the grammar which show influences from the substrates and/or superstrates? I believe that answering this question will point to solutions to the "logical problem of language acquisition" — and a version thereof made more complex by the socio-linguistic conditions surrounding Creolization.

References

[1] Roger W. Andersen. *Pidginization and Creolization as Language Acquisition.* Newbury House, Rowley, Mass. 1983.

[2] Derek Bickerton. The Language Bioprogram Hypothesis. *Behavioral and Brain Sciences*, 7, 1984.

[3] Noam Chomksy. *Lectures on Government and Binding.* Foris Publications, Dordrecht, 1981.

[4] Noam Chomsky. *Knowledge of Language.* Praeger, New York, 1986.

[5] Noam Chomsky. *Barriers.* MIT Press, Cambridge. 1986.

[6] Michel F. DeGraff Creole languages & parameter setting: a case study using Haitian and the pro-drop parameter. Paper presented at the *Leiden Workshop on Creole Languages and Language Acquisition.* To appear in Wekker, ed., *Creole Languages & Language Acquisition*, Mouton de Gruyter, Berlin.

[7] Michel F. DeGraff. The Syntax of Predication in Haitian. *22nd Annual Meeting of the North-Eastern Linguistics Society.* University of Delaware. October 1991.

[8] Michel F. DeGraff. Predication in Haitian: *Se* vs. French *C'est. 22nd Annual Linguistic Symposium on Romance Languages.* University of Texas at El Paso and Universidad Autónoma de Ciudad Juárez in Mexico. February 1992.

[9] Michel F. DeGraff. On the structure of Small Clauses: Evidence from Haitian. *GLOW Workshop on African and Creole Languages.* Universidade de Lisboa, Portugal. April 1992.

[10] Michel F. DeGraff Is Haitian Creole a pro-drop language? In Byrne & Holm, eds., *The Atlantic Meets the Pacific: Selected Papers from the Society for Pidgin & Creole Linguistics.*, John Benjamins, Amsterdam. In press.

[11] Nina M. Hyams. *Language Acquisition and the Theory of Parameters.* D. Reidel Publishing Company, Dordrecht. 1986.

[12] Thomas Roeper & Edwin Williams (eds.). *Parameter Setting.* D. Reidel Publishing Company, Dordrecht. 1987.

Barbara Di Eugenio

Department of Computer and Information Science
dieugeni@linc.cis.upenn.edu

Goals and Actions in Natural Language Instructions

Keywords: Instruction Understanding, Discourse Processing, Plan Inference

Human agents are extremely flexible in dealing with Natural Language instructions: they are able to both adapt the plan they are developing to the input instructions, as well as to adapt the input instructions to the plan they are developing. Borrowing the term from [1], I call this two-way adaptation process *accommodation.*

That accommodation is indeed taking place is shown by minimal pairs such as

1. *Cut the square in half.*
2. *Cut the square in half to create two triangles.*

(1) admits an infinite number of different executions, because there are infinite ways to cut a square in half; however, in (2), the purpose clause *to create two triangles* reduces the number of possible executions to one, namely, cutting the square along (one of) the diagonal(s).

Naturally occurring data provide evidence for the following claims:

1. The actions an agent has to perform when s/he is given instructions have to be *computed* from the descriptions given in the instructions themselves, as opposed to simply *extracted* from such descriptions. In all previous work that deals with instructions – e.g. [3], [9] – the action to be executed is equated with the logical form extracted by the parser, which for both (1) and (2) would be something like *cut(agent, square, in-half).* However, in (2) such a logical form is underspecified with respect to the action that has to be executed, *cut the square in half* **along the diagonal**. This augmented description must be computed.

2. The goals that an agent is trying to achieve guide this computation: it is *create two triangles* that allows us to infer that the intended action is *cut the square in half* **along the diagonal**. Many goals are explicitly stated in the instructions themselves, as (2) shows.

3. Action descriptions found in instructions can not be expected to exactly match the knowledge that an agent may have about actions and their characteristics. Presumably an agent will have knowledge about *cutting a square in half along the diagonal* as a method for *creating two triangles.* This puts a requirement on the inference mechanisms: they must be flexible enough to deal with action descriptions which are underspecified with respect to the stored knowledge.

Further evidence for these claims comes from my analysis of naturally occurring instructions: in particular, I have examined two syntactic constructions, *purpose clauses*, such as *to create two triangles* in (2), and *negative imperatives.* In [4], I show what pragmatic functions these constructions fulfill, and what kind of inferences an agent must perform to understand them.

I am developing a model of instruction understanding consistent with the claims put forward above. Such model consists of:

1. A speaker/hearer model of imperatives that extends the one presented in [3].

2. A representation of actions and of goals that is able to support accommodation inferences. I will propose that a hybrid KR system [2] should be used, whose primitives are those proposed in [7] for the semantic representation of verbs and actions.

3. Inference mechanisms that contribute to building the structure of the intentions – the plan graph – that the agent develops while interpreting instructions.

My work is related to the "Animation from Natural Language Instructions" project [10, 1] – see contributions by Webber, Badler, Baldwin, Geib, Jung, Levison, Moore, White.

References

[1] Norman Badler, Bonnie Webber, Jugal Kalita, and Jeffrey Esakov. Animation From Instructions. In N. Badler, B. Barsky, and D. Zeltzer, editors, *Making Them Move: Mechanics, Control and Animation of Articulated Figures*, pages 51–93. Morgan-Kaufmann, 1990.

[2] Ronald Brachman, Deborah McGuinness, Peter Patel Schneider, Lori Alperin Resnick, and Alexander Borgida. Living with CLASSIC: when and how to use a KL-ONE-like language. In J. Sowa, editor, *Principles of Semantic Networks*, Morgan Kaufmann Publishers, Inc., 1990.

[3] Philip Cohen and Hector Levesque. Rational Interaction as the Basis for Communication. In J. Morgan, P. Cohen, and M. Pollack, editors, *Intentions in Communication*, MIT Press, 1990.

[4] Barbara Di Eugenio. *Goals and Actions in Natural Language Instructions.* Technical Report MS-CIS-92-07, University of Pennsylvania, 1991.

[5] Barbara Di Eugenio and Bonnie Webber. Plan recognition in Natural Language instructions, in *First International Conference on AI Planning Systems*, 1992.

[6] Barbara Di Eugenio. Understanding Natural Language instructions: the case of purpose clauses, in *Proceedings of the ACL*, 1992.

[7] Ray Jackendoff. *Semantic Structures. Current Studies in Linguistics Series*, The MIT Press, 1990.

[8] David Lewis. Scorekeeping in a language game. *Journal of Philosophical Language*, 8:339–359, 1979.

[9] Steven Vere and Timothy Bickmore. A basic agent. *Computational Intelligence*, 6:41–60, 1990.

[10] Bonnie Webber, Norman Badler, Barbara Di Eugenio, Libby Levison, and Michael White. Instructing Animated Agents. In *Proc. US-Japan Workshop on Integrated Systems in Multi-Media Environments*, Las Cruces, NM, 1991.

Robert Frank

Department of Computer and Information Science
frank@linc.cis.upenn.edu

Syntactic Locality and Tree Adjoining Grammar: Grammatical, Acquisition and Processing Perspectives

Keywords: Syntax, Language Acquisition, Parsing, Government-Binding Theory

My research focuses on the question of the nature of syntactic locality, i.e. what restrictions there are on the structural relations which may be grammatically relevant. Within generative linguistic theory, and in particular within Government-Binding theory, this question has usually been addressed by explicitly defining locality conditions under which the relations of government and movement may obtain. Instead, I have investigated this question in the context of the Tree Adjoining Grammar (TAG). This constrained grammatical formalism provides the theory of grammar with a module determining how phrase structure representations are composed, and, in so doing, restricts the class of possible grammatical principles to those expressible over the local domains of TAG's elementary structures, which are determined by the function-argument structure of a single lexical head. Therefore, grammatical locality is not derived from a special stipulation, but instead falls out from general principles underlying the sorts of representations which the grammar has available.

In my dissertation, I explore the implications of this conception of syntactic locality on the statement of a grammatical theory as well as on the construction of psychological models of parsing and acquisition. With respect to the concerns of the theory of grammar, I investigate the principles that govern the well-formedness of basic clausal structure. In a theory of grammar with TAG as an underlying theory of phrase structure, this translates into the question of the nature of the elementary structures that the formalism allows. By viewing these structures as minimal function-argument structures, I show that the requirement that all grammatical dependencies be treated within the local domain of the formalism's elementary structures leads to analyses of a range of phenomena including small clauses, copular constructions, asymmetries between nominals and gerunds, parasitic gaps [1], and amount quantifier scope.

Turning to questions of syntactic performance, I demonstrate that the "TAG hypothesis", i.e. that the theory of grammar is expressed as restrictions on a set of local domains, provides a solution to the apparently incompatible demands of grammatical transparency on the one hand, and computational efficiency on the other. We might impose this as desiderata of a parsing model for a principles and parameters theory. Previous principle-based parsers have elegantly implemented the grammatical theory in the parsing-as-deduction paradigm, but have eschewed claims of psychological validity as a result of the computational properties of logical deduction. I show that by taking into account the nature of grammatical principles, in particular that they express constraints over elements in local domains, a parsing model can be constructed which transparently implements the linguistic theory, yet operates incrementally and in linear time [2, 3].

Finally, within the domain of acquisition, I show that the assumption that children are unable to perform one of the combinatorial operations of the TAG formalism, *adjoining*, explains the relative difficulty experienced in the acquisition of a seemingly disparate set of

syntactic phenomena, including Wh-movement, control constructions, raising, subject-aux inversion and realization of tense [4]. Performing the adjoining operation is more computationally complex than the other operation of the formalism, *substitution.* Thus, this expression of the grammar allows us to explain precisely why certain constructions occur later than others. If this work is on the right path, it represents a new kind of explanation for the time course of acquisition through properties of formal and computational complexity.

In addition to these topics, I am also currently investigating, within the Gold paradigm, questions of formal learnability in the context of principles and parameters theories with and without locality restrictions of various kinds. Also, I am looking at the phenomenon of word order variation in Germanic and its implications for the theory of movement and the typology of syntactic positions [5].

References

[1] Robert Frank. Parasitic Gaps and Locality Conditions. In *Proceedings of the 27th Regional Meeting of the Chicago Linguistic Society.* University of Chicago. 1991.

[2] Robert Frank. *Computation and Linguistic Theory: A Government-Binding Theory Parser Using Tree Adjoining Grammar.* Technical Report MS-CIS-90-29 (LINC Lab 171), University of Pennsylvania. 1990.

[3] Robert Frank. Licensing and Tree Adjoining Grammar in Government-Binding Parsing. In *Proceedings of 28th Annual Meeting of the Association for Computational Linguistics.* University of Pittsburgh. 1990.

[4] Robert Frank. Formal Grammar and the Acquisition of Complex Sentences. In *Proceedings of the 8th Eastern States Conference on Linguistics.* G. Westphal, J. Dai and B. Ao (editors). Ohio State University. 1991.

[5] Robert Frank, Young-Suk Lee and Owen Rambow. Scrambling as Non-Operator Movement and the Special Status of Subjects. In *Proceedings of the 3rd Leiden Conference for Junior Linguists.* University of Leiden. 1992.

Christopher W. Geib

Department of Computer and Information Science
geib@linc.cis.upenn.edu

ItPlanS (an intentional planning system): Integrating Intention and Planning

Keywords: Planning, Intentions, Elimination of Preconditions

The manner in which an agent choses to accomplish a given task is often determined not only by the actual action or goal, but also by the agent's other intentions. For example, breaking the door off its hinges is a way of satisfying the goal of having a door open. However, this is not an acceptable solution in most cases.

Intentions play two roles in our planning deliberations. First, they are the ends of our deliberations. That is to say, those things to be achieved in the future. Second, they also constrain the manner in which other intentions are satisfied. Bratman[1] gives a detailed description of these two different facets of intentions in planning.

The goal of this work is to provide a principled treatment of the role of intentions in planning. To this end, the role of preconditions in action representations must be reevaluated. Preconditions often implicitly encode intentions the agent must hold for the action to achieve its expected effects. For example, in STRIPS[2] the action PICKUP(a) had as a precondition that the block "a" must be clear. This is not logically required for moving blocks; it is not even physically required. This condition must be true only if the agent has intentions not to move other blocks, if the objects on "a" make the action physically impossible for the agent in that world state, or if the objects on "a" would cause a result that is undesired in the situation. Notice, all of these possibilities are dependent on the situation in which the action is performed.

In these cases, preconditions are encoding situation-dependent, intentional information in a universal manner. In order to use intentions in making planning decisions, these intentions must be represented explicitly. In the end, this explicit encoding of intentions and the use of situated reasoning will allow for the complete elimination of preconditions.

The Intentional Planning System (ItPlanS) is a first step toward a planner that gives intentions and situated reasoning the role in planning they deserve. ItPlanS is a hierarchical intention planner. The system plans by recursively expanding intentions, developing lower level intentions, until one is found that can be immediately satisfied by a single action. This action is performed and the decomposition process is repeated. The only structures maintained across each iteration is a list of the top level intentions. This means the decompositions for a given intention can vary from iteration to iteration allowing ItPlanS to adapt its plan to a changing world environment. Since ItPlanS decides on the next action based on the world state, the system's knowledge, and its intentions, it might be characterized as a reactive system rather than a planning system. Regardless of how ItPlanS is classified, the implications of this work are valid to all planning paradigms.

References

[1] Bratman M. *Intentions, Plans, and Practical Reason*, Harvard University Press, Cambridge, Mass. 1987.

[2] Fikes R.E. and Nilsson N.J. Strips: a new approach to the application of theorem proving to problem solving, *Artificial Intelligence*, 2:189, 1971.

Abigail S. Gertner

Department of Computer and Information Science
agertner@linc.cis.upenn.edu

Critiquing the management of trauma patients

Keywords: Critiquing, AI in medicine, Plan inference, Generation

I have recently begun working with a group at the University of Pennsylvania who are developing a system (TraumAID) to provide decision support during the initial definitive management of severe injuries [5]. I am concerned with improving the form of interaction between the system and the physician user. In particular, I am working on developing a *critiquing mode*, in which the system evaluates the course of action proposed by the physician rather than presenting the physician with a plan that it has determined to be optimal. Critiquing is a technique that has been incorporated into the interfaces of several medical expert systems [4, 3, 1, 2]. It has the advantage that it acknowledges the role of the physician as the primary *decision maker* while constantly monitoring his proposed plan to ensure that it does not deviate significantly from the system's preferred plan. Whenever such differences are detected, the system produces an appropriate comment, alerting the physician is aware that a potential problem exists. This approach is particularly effective given the subjectivity and variability of preferred treatments that often play a role in medical decision-making.

Previous critiquing models have been based on rule-based reasoning systems that produce a recommended treatment plan given the relevant patient data. Thus they are primarily applicable to domains that are sufficiently predictable to support such an inflexible determination of plans. In trauma care, on the other hand, the advising system must be able to adapt to a constantly changing situation. Each new piece of evidence triggers reevaluation of conclusions and potential changes to the proposed course of action. In addition, trauma management requires the execution of urgent therapeutic procedures, while at the same time pursuing additional treatment goals through diagnostic activity. Both diagnostic and therapeutic procedures must be scheduled as efficiently as possible with both urgency and logistical considerations in mind. The flexibility of TraumAid 2.0, the current version of the system, derives from the integration of a *rule-based reasoner* to determine a set of diagnostic and therapeutic goals to be addressed, with a *planner* that reasons from the set of pending goals and the current situation to produce an appropriate course of action.

The critiquing module for TraumAID must consist of several components: a plan-inference component to analyze the physician's proposed sequence of actions, a plan-comparison component to compare the physician's proposed plan with the plan developed by TraumAID, a content generator to determine what the critique should comment on and to generate explanations, and a final output component. The user should be able to choose how picky he wants the system to be (comment on everything vs. give only essential warnings) and how much detail the critiques should contain.

In addition to contributing to a more acceptable interface between system and physician, I am also interested in exploring the educational potential of the critiquing model. The fact that the user must first propose a plan to the system, together with the system's ability to evaluate that proposal and to explain its reasoning, suggest that this approach could effectively be applied to a tutorial setting.

References

[1] Curtis P. Langlotz and Edward H. Shortliffe. Adapting a consultation system to critique user plans. *International Journal of Man-Machine Studies*, 19:479–496, 1983.

[2] Perry L. Miller. *A Critiquing Approach to Expert Computer Advice: ATTENDING.* London: Pittman Press, 1984.

[3] Ivan Rankin. The deep generation of text in expert critiquing systems. Licentitate Thesis no.184, Linkoping University, 1989.

[4] Johan van der Lei. *Critiquing Based on Computer-Stored Medical Records.* PhD thesis, Erasmus University, 1991.

[5] Bonnie L. Webber, Ron Rymon, and John R. Clarke. Flexible support for trauma management through goal-directed reasoning and planning. To appear in *Artificial Intelligence in Medicine.*

Jane A. Gillette

Department of Psychology
gillette@cattell.psych.upenn.edu

The Role of Syntax in Verb Acquisition

Keywords: Language Acquisition, Syntax

This research is being done in collaboration with Lila Gleitman. The focus of our research is to investigate the connection between syntax and semantics and the implications of these connections for language acquisition. Specifically, the topic is the Mapping Problem: how children come to connect certain phonetic objects (e.g. /iyt/) with specific concepts (here, the concept 'eat'). Also related is the issue of how children acquire the argument structures and syntactic structures associated with each such item. For example, eating involves an action of an animate being on an "edible", and appears in both transitive (*John eats an apple*) and intransitive (*John eats*) syntactic structures.

The mapping problem is usually assessed to be accomplished through the child's attention to what is occurring in the world when the word is uttered. However, little is known about the aspects of an observed scene that children will encode into the meaning of a new verb uttered in the presence of these scenes. My goal is to look at those aspects of the learning situation, in particular linguistic information, which are useful and or necessary for the child to successfully solve this word learning problem. The recent literature suggests that verbs line up rather poorly with their extra-linguistic contexts of use. Therefore it has been suggested that linguistic (syntactic) evidence is important for extracting the meanings of verbs.(Landau & Gleitman, 1985; Gleitman,1990)

I am interested, particularly, in the process by which young children come to learn the meanings of abstract verbs such as think, and know. Such verbs clearly pose a problem for the view that situational concomitants are sufficient to explain how the meanings are extracted from observing a scene. Perhaps not coincidentally, these "non-observable" verbs or verbs with a mental component license a unique range of syntactic structures. For example, one can say *John thinks that the building is tall*; but not *John eats that....* Adults are very sensitive to the presence or absence of this kind of syntactic information such that given a scene to observe, and a novel verb (e.g. gorp) to learn, they rarely conjecture abstract verbs like think without syntactic information in the scene. However, adults always guess such verbs correctly with appropriate syntactic information (Lederer et al, 1991). It has also been shown that syntactic information influences children's conjectures about the meaning of a new verb when a scene has more than one interpretation such as give and take (Fisher et al, forthcoming). It is not known whether or not children also use this syntactic information to acquire abstract verbs, and this is the hypothesis that we are investigating.

References

[1] Fisher,C., Hall,G., Rakowitz,S. & Gleitman,L. When it is better to receive than to give: syntactic and conceptual constraints on vocabulary growth. In it Journal article), (forthcoming).

[2] Gleitman,R. The structural sources of verb meanings. In *Language Acquisition*,1, 3-55, 1990.

[3] Landau,B. & Gleitman,R. *Language and experience: evidence from the blind child* Harvard University Press, Cambridge, MA, 1985.

[4] Lederer,A., Gleitman,R., Gleitman,H., & Ariel,S. *Evidentiary sources of verb learning: scenes, co-occurrence, syntactic, and selectional information in mother's speech to babies.* Paper presented at the Child Language Conference, Boston, MA, 1991.

Daniel Hardt

Department of Computer and Information Sciences
hardt@linc.cis.upenn.edu

Verb Phrase Ellipsis and Discourse

Keywords: Ellipsis, Discourse, Semantics

It is generally acknowledged that the process of understanding a discourse results in the construction of a discourse model which contain semantic objects evoked in the discourse. Many accounts of pronominal anaphora make use of a discourse model, treating the semantic objects in the discourse model as the possible antecedents for pronominal expressions. I propose to account for Verb Phrase ellipsis in an analogous way, by extending the discourse model to include semantic objects associated with Verb Phrases, as well as those associated with Noun Phrases.

This approach accords with the intuition that Verb Phrase ellipsis is based on identity of predication; an intuition reflected in the Logical Form identity theories of Sag [4] and Williams [6], as well as many subsequent syntactic approaches. In contrast to these approaches, the current approach imposes a semantic identity condition; the meaning of the elided VP is identified with a VP meaning stored in the discourse model. I treat semantic objects as "file change potentials", or relations on contexts.

There is a wide range of evidence that the identity condition governing VP ellipsis must be semantic rather than syntactic. There are many cases of VP ellipsis in which there is no syntactic VP that can be identified as the antecedent: for example, the antecedent may be contextually supplied, it may be associated with a Noun Phrase, or it may be a combination of two or more contextually available predicates. When there is a syntactic VP as antecedent, there are many cases which violate a syntactic identity condition between antecedent and ellipsis: for example, there may be a wh-trace in the antecedent, although a wh-trace would not be syntactically permitted in the ellipsis site. I argue that all of these cases can be handled in the semantic approach being advocated here.

A computer implementation of the proposed approach is under way; it consists of some simple extensions to the "Incremental Interpretation" system of [3]. The system will be tested on several hundred examples of VP ellipsis collected from the Brown Corpus.

References

[1] D. Hardt. A Discourse Model Account of VP Ellipsis. *ESCOL 1991.* Baltimore, MD.

[2] D. Hardt. Ellipsis and Discourse. Dissertation Proposal, U. Pennsylvania, 1991.

[3] F. Pereira & M. Pollack. Incremental Interpretation. *Artificial Intelligence.* 1991.

[4] I. A. Sag. *Deletion and Logical Form.* Ph.D. thesis, U. Massachusetts-Amherst. 1976.

[5] B. L. Webber *A Formal Approach to Discourse Anaphora.* Ph.D. thesis, Harvard University. 1978.

[6] E. Williams. *Discourse and Logical Form. Linguistic Inquiry,* 8(1):101-139. 1977.

Michael Hegarty

Postdoc, Institute for Research in Cognitive Science
mhegarty@linc.cis.upenn.edu

The Semantic Function of Factive Complementizers

Keywords: Complementizers, Factive Verbs

A structural basis can be given for the presuppositions induced by the use of certain expressions. In a definite description, analyzed along the lines of Higginbotham [4], the determiner theta–binds a variable introduced by the head noun. Assuming that the domain bound by the definite determiner is the discourse frame, a presupposition of existence is generated. The present work seeks to generalize this result to instances of theta–binding of an event variable.

In factive ascriptions, such as *John noticed that Bill broke the vase*, the speaker presupposes the truth of the complement clause. No such presupposition is involved in the use of propositional ascriptions, such as *John believes that Bill broke the vase.* This difference can be given a structural basis in terms of the event structure of factive and propositional complement clauses, appealing to the event variable of Higginbotham [4] [5]. In a factive complement, the complementizer has the semantic function of discharging the event variable of the complement clause, thus binding the event variable within the discourse frame. As a result, the factive complement is interpreted as presupposed by the speaker. In a propositional complement, the complementizer does not have a semantic function; the event variable therefore propagates up to the CP node of the complement clause, where it is existentially quantified out in semantic composition with the propositional verb. The resulting semantics for propositional ascriptions exhibits formal parallels to Higginbotham's [3] analysis of perception sentences, except that in the propositional ascription the domain of events and individuals must be relativized to the mental model of the subject of the ascription.

This explains various aspects of the distribution of factive and propositional tensed complements, including why the object expletive *it* is available with factive tensed complements, and not with propositional tensed complements.

Complementizers and Antecedent Government

Keywords: Complementizers, Adjunct Extraction, Antecedent Government

In the literature, it is argued that *that*–type complementizers have semantic content in tensed factive complement clauses (Hegarty [2]) and in complement clauses of inherently negative verbs (Laka [7]). Adjunct *wh-* movement is blocked out of both complement types. This suggests that antecedent government is blocked over semantically contentful complementizers. This idea is approached by assuming that semantically contentful complementizers are obligatorily present throughout a syntactic derivation, whereas complementizers which play no semantic role delete at some point in the derivation, an application of Chomsky's [1] Principle of Full Interpretation. On this basis, domains of antecedent government can be defined in a way similar to Kayne's [6] notion of a g-projection set, but with a built–in sensitivity to the presence of semantically contentful complementizers. Applying the results

to adjunct extraction from infinitival complement clauses, it seems that the extension of antecedent government domains is also conditioned by the presence of (in)dependent tense operators within a domain. The nature of this constraint needs to be carefully worked out in terms of an account of the structure of infinitival complement clauses, a focal point of current work on this project. The system developed to this point accounts for patterns of adjunct extraction out of tensed indicative and subjunctive complements in Western Romance languages (French, Spanish, Catalan, Italian), but with subsidiary assumptions about the obligatoriness and semantic function of complementizers in these languages. The status of some of these subsidiary assumptions is not clear, and is another focal point of current work on this project.

References

[1] Noam Chomsky. Some Notes on the Economy of Derivation and Representation. In I. Laka and A. Mahajan, editors, *MIT Working Papers in Linguistics 10*, 1989.

[2] Michael Hegarty. *Adjunct Extraction and Chain Configurations.* PhD thesis, MIT, 1992.

[3] James Higginbotham. The Logic of Perceptual Reports: An Extensional Alternative to Situation Semantics. *The Journal of Philosophy*, 80:100–127, 1983.

[4] James Higginbotham. On Semantics. *Linguistic Inquiry*, 16:547–594, 1985.

[5] James Higginbotham. Elucidations of Meaning. *Linguistics and Philosophy*, 12:465–517, 1989.

[6] Richard Kayne. Connectedness. In R. Kayne, *Connectedness and Binary Branching*, Foris, 1983.

[7] Itziar Laka. *Negation in Syntax: On the Nature of Functional Categories and Projections.* PhD thesis, MIT, 1990.

James Henderson

Department of Computer and Information Science
henders@linc.cis.upenn.edu

An Incremental Connectionist Parser Using Partial Descriptions of Phrase Structure

Keywords: Syntactic Parsing, Connectionism, Linguistic Representation

In the investigation of natural language, the question arises as to how it is possible that people can understand language with the speed and reliability with which they do. Given the vast amount of grammatical information and the proliferation of possible analyses at any given point in a parse, the answer must involve a large amount of parallelism. This work looks at one proposed model of massively parallel computation and tries to determine whether that model is computationally adequate for parsing natural language syntax.

The model of massively parallel computation investigated here is the connectionist computational architecture proposed in [4]. This architecture is especially suited for parsing natural language because, unlike many other connectionist models, it directly manifests the symbolic interpretation of the information it stores and manipulates. Virtually all characterizations of natural language syntax have relied heavily on symbolic representations. Like other connectionist models, this architecture is massively parallel and does its computations with units which have roughly the same computational properties as neurons. In this architecture, a computing module can represent entities, store predications over those entities, and use pattern–action rules to dynamically manipulate this stored information. While this representation provides a rather general purpose computing framework, it has two limitations which pose significant problems for parsing natural language. First, it cannot represent arbitrary disjunction. This means that an incremental parser must operate essentially deterministically (as in [2]). Second, the memory capacity of any computing module is bounded. The number of entities which can be stored is bounded by a small constant, the number of predications is bounded, and the number of pattern–action rules is bounded. This means the parser must be able to forget information which it no longer needs to finish parsing the sentence.

In order to compensate for the above two limitations of the computational architecture while still maintaining a linguistically adequate representation of grammatical information, the parser developed in this work makes extensive use of partial descriptions of phrase structure trees. In [3] this approach was proposed as a way to parse deterministically. The partiality of the descriptions makes it possible to simply not specify information which has not yet been determined, rather than using disjunction to list possible alternatives. The specific instantiation of this approach to parsing which the parser uses is a grammatical formalism called Structure Unification Grammar (SUG, [1]). SUG provides a language for specifying partial descriptions of tree fragments, combines these descriptions using node equations, and provides a criteria for the successful completion of a derivation. Since the combination operation and the completion requirement both apply locally to nodes, a completed node which will not be involved in any further equations can be forgotten without interfering with the subsequent completion of a valid parse. This property allows the parser to forget nodes, and thus to parse arbitrarily long sentences with only a finite amount of

memory, as required by the second limitation of the computational architecture. In addition to providing the mechanisms necessary to compensate for the limitations of the architecture, SUG provides a powerful and perspicuous framework for specifying grammatical information. See [1] for an extensive discussion of how insights and analyses from a variety of other linguistic investigations (including D-Theory, TAG, CCG, and LFG) can be captured in SUG.

The parser developed in this work uses SUG grammars and parses following SUG derivations. The parser stores in its memory the current intermediate description of the derivation, and incrementally combines grammar entries for the words of the sentence with this intermediate description. The nonterminal nodes of the intermediate descriptions are the entities in the parser's memory, phrase structure information about these nodes is represented in the predications in the parser's memory, and the pattern–action rules are used to encode general properties of SUG structures and to encode how each grammar entry can be combined with an intermediate description. This model of parsing is interesting independently of the fact that it is connectionist because of its implications for efficient natural language parsing. Since the grammar is represented in pattern–action rules which compute in parallel, the parser's speed is independent of the size of its grammar. Its output is incremental, monotonic, and does not include disjunction. Its disambiguation mechanism provides a simple parallel interface for the influence of higher level language modules. It parses in quasi-real time (constant time per word). Interestingly, assuming neurologically plausible timing characteristics for the computing units, a biological implementation of the parser could parse simple sentences with a worst case throughput in the vicinity of six words per second, which is real time for spoken language. More complex sentences (which require time for disambiguation) would not be much slower.

Currently the parsing model has been developed and for the most part implemented using the Rochester Connectionist Simulator. An empirical study to verify that this model is computationally adequate for parsing natural language is in progress. Also, the use of probabilistic information in making disambiguation decisions is being added. This parsing model is particularly well suited for the use of such information because of the ability of the connectionist architecture to do evidential reasoning, and because SUG's large domain of locality allows probabilistic information to be stated within grammar entries. Future work includes applying this parser to specific natural language processing tasks and investigating the significance of this parser as a model of human linguistic performance.

References

[1] James Henderson. Structure unification grammar: A unifying framework for investigating natural language. Technical Report MS-CIS-90-94, University of Pennsylvania, Philadelphia, PA, 1990.

[2] Mitchell Marcus. *A Theory of Syntactic Recognition for Natural Language.* MIT Press, Cambridge, MA, 1980.

[3] Mitchell Marcus, Donald Hindle, and Margaret Fleck. D-theory: Talking about talking about trees. In *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*, Cambridge, MA, 1983.

[4] Lokendra Shastri and Venkat Ajjanagadde. From simple associations to systematic reasoning: A connectionist representation of rules, variables and dynamic bindings. Technical Report MS-CIS-90-05, University of Pennsylvania, Philadelphia, PA, 1990.

Beth Ann Hockey

Department of Linguistics
beth@linc.cis.upenn.edu

Prosody and the Interpretation of *okay*

Keywords: Prosody, Cue Phrase, Discourse Structure

Okay can contribute semantic content or discourse information to an utterance. Hirschberg and Litman(87) and Litman and Hirschberg(90) claim that sentential and discourse uses of similar, ambiguous 'cue phrases' can be distinguished by prosodic features.The question I address here is whether finer distinctions than discourse vs. sentential use can be made on the basis of prosody. To this end, studying *okay* has several advantages. First, *okay* has multiple discourse uses. Consequently, in addition to disambiguating between discourse and sentential use, hearers need to distinguish among the various discourse uses. Second, *okay* occurs with high frequency in conversation, so tokens are relatively easy to collect. Third, speakers have strong intuitions that at least some varieties of *okay* are easily distinguishable perceptually.

I investigate the relation between prosody and interpretation for the lexical item *okay* using two approaches: 1) forming natural groupings of F0 contours and of contexts, and identifying correlations between the two types of groupings; and 2) performing a perceptual experiment in which subjects interpret prosodically varied tokens presented in isolation. Data for the study is from taped dialog of a task requiring two participants separated by a barrier to cooperatively reconstruct a paperclip design (barrier task).

Grouping of F0 contours is done using characteristics such as relative F0 height of the first and second syllables and general shapes of the two syllables (e.g. rising, falling, degree of rise or fall). Because the study is restricted to *okay*, it is relatively easy to divide tokens into natural intonational classes by sorting pitch contours visually and auditorily, without relying on any previously-assumed system of description. This classificatory independence is an advantage since the fit between existing descriptive systems and natural data is often dubious.

Initial grouping of contexts into interpretive categories is based on: 1) observed conversational behavior, such as turn taking and 2) discourse segmentation/change or continuation of theme of discussion. Following Litman and Hirschberg (90), I assumed that *now* indicates the beginning of a new discourse segment. Therefore, I considered the presence of *now* immediately following *okay* to be diagnostic of discourse segment completion. Similarly *and* immediately following *okay* was considered indicative of discourse segment continuation.

Grouping the barrier task data reveals three especially clear F0 contour types. One F0 contour type (ct1) is flat. The two syllables have very close F0 values and each syllable remains at its value for most of the syllable duration.

The second contour type (ct2) has a first syllable higher than the second with an abrupt transition. Both syllables have constant F0 value so are basically flat.

In the third contour type (ct3), the first syllable is flat or slightly falling. The second syllable is rising. The second syllable begins higher than the end of the first and ends considerably higher than any point in the first syllable.

Other groupings of contours in the barrier task data contain fewer tokens than ct1, ct2 and ct3 and will require examination of additional data for reliable characterization. For

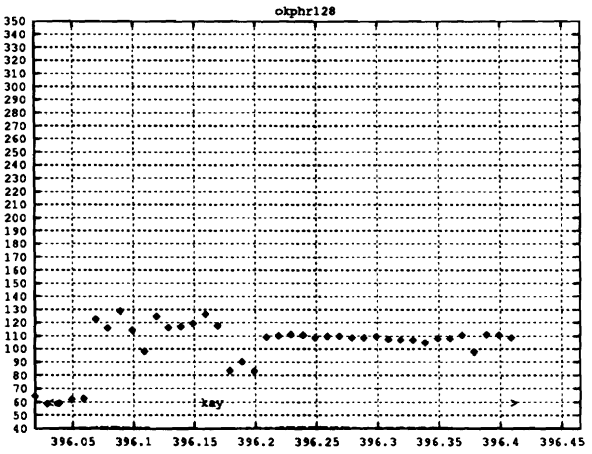


Figure 1: ct1

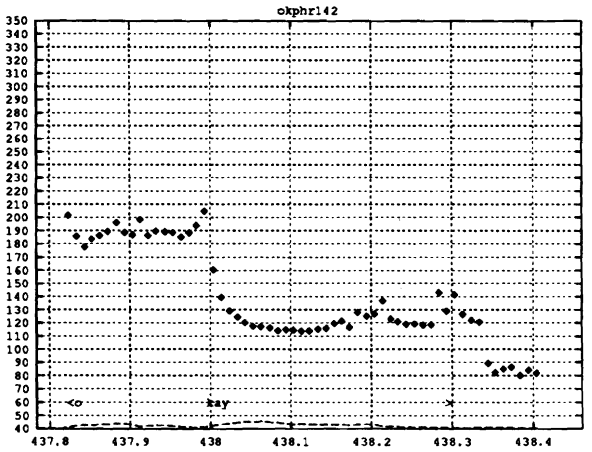


Figure 2: ct2

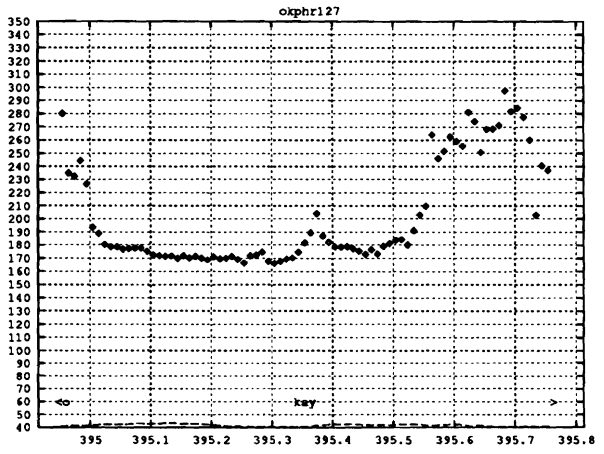


Figure 3: ct3

tokens of ct1, ct2 or ct3, contour type is predictive of context type. 86% of the tokens with ct1 were from a context in which the speaker was continuing an instruction that had been started in a previous utterance. 88% of the tokens with ct2 were in a context where the *okay* marked the end of a task or subtask. 89% of the tokens with ct3 were from a context in which the speaker was passing up a turn and letting the other person continue.

The perception experiment consists of presenting subjects with instances of *okay* in isolation to which they respond with interpretations of the instances. The experiment provides a way of testing the perceptual reality of proposed categories, and of generating interpretation categories that might not have been revealed through other methods.In the pilot study, subjects could replay tokens at will and the form of their responses was fairly unconstrained. Tokens for the pilot were selected from a separate data base with sound quality as the sole selection criterion. The newest version of the experiment allows subjects a limited time to respond, and the number and timing of repetitions is controlled. Stimuli for this experiment are from the barrier task data and represent a range of variation based on the results of the 'grouping' phase of the research.

Although clearly not random, the results of these experiments are fairly messy. However, the improvement in interpretability between the pilot and the newer experiment indicates that further refinements will yield reasonable results. Consequently, a major part of my current work is developing and testing improvements to the perceptual experiment.

References

[1] Barbara J. Grosz and Candace L. Sidner. Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12(3):175–204, 1986.

[2] Julia Hirschberg and Diane Litman. Now let's talk about *Now*: identifying cue phrases intonationally. In *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics*, 1987.

[3] Beth A Hockey. The function of okay as a cue phrase in discourse. December 1989. Department of Linguistics, University of Pennsylvania.

[4] Diane Litman and Julia Hirschberg. Disanbiguating cue phrases in text and speech. In *Proceedings of Coling90*, 1990.

[5] Cynthia Ann McLemore. *The Pragmatic Interpretation of English Intonation: Sorority Speech*. PhD thesis, University of Texas at Austin, 1991.

[6] J. Pierrehumbert and J. Hirschberg. The meaning of intonational contours in the interpretation of discourse. In *Intentions in Communication*, MIT Press, 1990.

[7] Maria Swora and Mary Beckman. Talk given LSA annual meeting 1991, January 1991.

Beryl Hoffman

Department of Computer and Information Science
hoffman@linc.cis.upenn.edu

Word Order Variation in Turkish

Keywords: Turkish, Scrambling, Pragmatics, CCGs

The strict word order in English allows us to identify the grammatical relations in a sentence. However in Turkish, and in many other "free word order" languages, a rich system of case markings identifies the predicate-argument structure of a sentence, whereas word order serves a pragmatic function. In Turkish the most common word order in simple transitive sentences is SOV (Subject Object Verb). However, all of the permutations of the sentence seen below are grammatical in the proper discourse situations.

(1) a. Esra gazeteyi okuyor.
Esra newspaper-acc read-present.
Esra is reading the newspaper.

b. Gazeteyi Esra okuyor.

c. Esra okuyor gazeteyi.

d. Gazeteyi okuyor Esra.

e. Okuyor gazeteyi Esra.

f. Okuyor Esra gazeteyi.

The pragmatic functions of these word order variations roughly consist of moving into the sentence-initial topic position, moving into the immediately pre-verbal focus position, or moving backgrounded information into post-verbal positions [1]. An active area of my research is to determine the specific pragmatic functions of the word order variations in Turkish. I am also interested in how prosody interacts with these word orders and the pragmatic information structure.

I have also been working on a CCG model for free word order languages. In Turkish, elements with overt case-marking generally can scramble freely, even out of embedded clauses. This suggest a CCG approach where case-marked elements are functions which can combine with one another and with verbs in any order. The primary advantage of using CCGs is the ease with which any two adjacent elements in a sentence can be combined in an incremental manner. We only have to fine-tune the combinatory operations in CCG to handle free word order. Karttunen [2] has proposed a Categorial Grammar formalism to handle free word order in Finnish. However, CCGs allow the operations of composition and type raising which have been useful in handling a variety of linguistic phenomena including long distance dependencies and nonconstituent coordination [3]; these operations will also play an essential role in an analysis of scrambling.

In complex Turkish sentences with clausal arguments, elements of the embedded clauses can be extracted and scrambled to positions in the main clause, i.e. long distance scrambling. Long distance scrambling appears to be no different than local scrambling as a syntactic

and pragmatic operation. Generally, long distance scrambling is used to move an element into the sentence-initial topic position or to background it by moving it behind the matrix verb.

(2) a. Fatma [Esra'nın okula gittiğini] biliyor.
Fatma [Esra-GEN school-LOC go-GER-3SG-ACC] know-PROG.
Fatma knows that Esra goes to school.

b. Okula$_i$ Fatma [Esra'nın e$_i$ gittiğini] biliyor.
school-LOC$_i$ Fatma [Esra-GEN e$_i$ go-GER-3SG-ACC] know-PROG.

c. Fatma [Esra'nın e$_i$ gittiğini] biliyor okula$_i$.
Fatma [Esra-GEN e$_i$ go-GER-3SG-ACC] know-PROG school-LOC$_i$.

I would like to develop a uniform analysis that handles both local and long distance scrambling. Further development of the CCG combinatory rules is necessary to handle long distance scrambling and to fully take advantage of our choice to maintain the verbs and case-marked elements as functors.

Prosody and pragmatic information must be incorporated into any account of free word order languages. Although there are not many syntactic restrictions on word order in Turkish, there are semantic and pragmatic restrictions on word order that we must take into account. Further research is necessary to decide how best to use prosody and pragmatic information within a CCG model to interpret Turkish.

References

[1] Eser Emine Erguvanli. *The Function of Word Order in Turkish Grammar.* University of California Press. 1984.

[2] Lauri Karttunen. "Radical Lexicalism". Paper presented at the Conference on Alternative Conceptions of Phrase Structure. New York. 1986.

[3] Mark Steedman. Dependency and Coordination in the Grammar of Dutch and English, *Language*, 61, 523-568. 1985.

Aravind K. Joshi

Professor, Department of Computer and Information Science
joshi@linc.cis.upenn.edu

Fixed and Flexible Phrase Structure: Coordination in Tree Adjoining Grammars in the Manner of Combinatory Categorial Grammars

Keywords: Grammar Formalisms, Coordination, TAGs

TAGs are like phrase structure grammars in the sense that the elementary structures are trees (although not derived from a phrase structure grammar) and the derived structures are trees. They are unlike phrase structure grammars in the sense that the combining operations are tree operations (substitution and adjoining) and not string operations (concatenation and substitution).

Categorial grammars, especially the Combinatory Categorial Grammars (CCG) [1, 2] under certain conditions (which hold for grammars written in CCGs) are equivalent to TAGs[3, 4]. CCGs are unlike phrase structure grammars and also TAGs, in the sense that a CCG does not assign a unique tree structure to a sentence (even in the unambiguous case); it is quite flexible in the assignment of the structure. Almost any contiguous sequence of words can be put together as a constituent in a CCG. This property is exploited by CCGs to give a very elegant account of a wide range of coordination phenomena.

Lexicalized TAGs (with substitution and adjunction) are similar to CCGs in the sense that for each lexical item the elementary tree(s) which is (are) 'anchored' on that item can be regarded as the (structured) category (categories) associated with that item. Then for any sequence of lexical items (contiguous or non-contiguous) we can assign a (structured) category. We will attempt to show how a CCG-like account for coordination can be constructed in the framework of lexicalized TAGs. We are also examining gapping and other related phenomena in this context. To the extent it is successful, it shows that an account of coordination can be constructed along the lines of CCG without having to construct constituents corresponding to sequences of lexical items that will ordinarily not be grouped as constituents. More specifically, constituency is defined in the elementary structures and this constituency is preserved and no additional constituent types have to be created. In a CCG being a function is the same thing as being a constituent and vice versa. In our approach we try to show how these two aspects can be kept apart while still realizing the kind of flexibility in the constituent structure that a CCG allows. We also examine some of the processing implications of this approach.

As a followup of this work, we are also exploring the possibility of developing a calculus in which we assign 'partial proofs' as categories to lexical items, allowing for a TAG-like account within the categorial framework.

Machine Translation

Keywords: TAG, Machine Translation

Lexicalized Tree Adjoining Grammar (LTAG) is an attractive formalism for linguistic description mainly because of its extended domain of locality and its factoring of recursion from the domain of local dependencies. LTAG's extended domain of locality enables one to localize syntactic dependencies (such as filler-gap), as well as semantic dependencies (such as predicate-arguments). The aim of this paper is to show that these properties combined with the lexicalized property of LTAG are especially attractive for machine translation.

The transfer between two languages, such as French and English, can be done by putting directly into correspondence large elementary units without going through some interlingual representation and without major changes to the source and target grammars. The underlying formalism for the transfer is "synchronous Tree Adjoining Grammars". Transfer rules are stated as correspondences between nodes of trees of large domain of locality which are associated with words. We can thus define lexical transfer rules that avoid the defects of a mere word-to-word approach but still benefit from the simplicity and elegance of a lexical approach.

We rely on the French and English LTAG grammars that have been designed over the past two years jointly at University of Pennsylvania and University of Paris 7-Jussieu.

References

[1] Mark Steedman. Combinatory grammars and parasitic gaps. In *Natural Language and Linguistic Theory*, 5, 403–439, 1987.

[2] Mark Steedman. Gapping and constituent coordination. In *Linguistics and Philosophy*, 13, 207-264, 1990.

[3] David Weir and Aravind Joshi. Combinatory categorial grammars: Generative power and relationship to linear context-free rewriting systems. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics (ACL)*, 1988, Buffalo, NY.

[4] David Weir. *Characterizing mildly context-sensitive grammar formalisms.* Ph.D. Dissertation, University of Pennsylvania, Philadelphia, PA.

Moon Ryul Jung

Department of Computer and Information Science
moon@graphics.cis.upenn.edu

Human Posture Planning

Keywords: Posture Planning, Situated Agents, Collision Avoidance

The problem I am studying involves devising a computational process by which the human body coordinates and controls its body parts to achieve given tasks; this process is called *posture planning*. The work is reported in [3, 1, 2]. The posture planning process is designed using motion-related rules and constraints commonsensically observed. That is, posture planning is done not at the physiological or biomechanical levels, but directly at the behavioral level. While motor behavior studies typically try to understand local motions of the lower body (e.g., for walking and jumping) or the upper body (e.g., reaching), we consider the whole body; in particular, coordination between the lower body and the upper body.

As part of **AnimNL** project, the ultimate input to the posture planner are natural language instructions. The objective of this study, however, is to provide the most bottom level interface between language related processes and agent animation. Even when particular ways of achieving given instructions are determined (see Di Eugenio, Levison, White) based on the meaning of verbs and the structural and functional features of objects referred to in instructions, and a partial sequence of 'primitive actions' are determined by a traditional AI symbolic planner (see Geib), those primitive actions are *goals* from the perspective of motion planning. Specifically, I call these goals *task-space goals*. They do not specify which body parts to move nor how to move them. Rather, I am concerned only with the geometric aspects of gross body motions. Moreover, I am particularly interested in *approaching* the goal region and *performing* the given task.

The process of posture planning for *task-space goals* is briefly summarized as follows. Task-space goals typically do not specify body parts to be used to achieve given goals. For instance, the task-space goal defined by the instruction *pick up the block* does not indicate what body part to use. Nor does the instruction indicate whether the upper body (torso) or the lower body is to be used. In other words, task-space goals do not specify which body resources to use. The agent must decide where and how much to move the body parts. Moreover, in order to determine a motion of a body part, the base joint relative to which a body part is moved must also be determined. An example of determined or quantified task-space goals is:

> Move the *palm center* of the right hand to *Cartesian space position* (x,y,z) using the *hip joint* as the base joint of the motion.

To obtain quantified task-space goals, the posture planner suggests goal placements (positions/orientations) of important body parts (or directions and distances of their movements), with respect to the task-space (Cartesian space). This requires capturing dependencies between body parts motions. Additionally, it requires predicting ramifications of intermediate decisions and verifying that they do not interfere with the global goal. Motion rules used for the task-space reasoning are described in terms of qualitative spatial relations between

the moving body parts and the objects around them, e.g., *in-front-of, forward, left-of, right-of, above, under, away-from, horizontally-away-from, parallel, perpendicular, within-view-of, within-comfortably-reachable-region.* The qualitative geometric relations are in turn defined by quantitative geometric relations.

Each quantified task-space of a body part has a base joint. When a goal is being considered, the sequence of body segments between the base joint and the body part is called an **active joint chain**. A given goal of a body part is ultimately achieved by rotations of the joints in the active joint chain. The body has multiple goals and hence multiple active joint chains. These are combined to form the global joint chain. Motions of joints in the global joint chain are supposed to coordinate with one another so that all the spatial goals of multiple body parts may be achieved. In the case of multiple spatial goals of body parts, the planner must deal with situations in which one goal demands that a joint rotate in one direction and another goal demands that same joint rotate in the opposite direction. This is an example of *goal interference* in AI planning, although at a lower level. To solve this problem, joint angles of a limb are considered *resources* for task-space goals for the end-point of that limb. Then the problem of joint goal interference is resolved by allocating joint resources to multiple spatial goals such that each spatial goal has at least one joint contributing to it. If this resolving strategy fails, it means that the multiple goals in question cannot be achieved *at the same time.* In such a case, the current planning decision is backtracked and alternative ways are tried.

References

[1] Moon Ryul Jung and Norman Badler. *Generating Human Motions via Symbolic Reasoning*, Technical Report MS-CIS-92-06, University of Pennsylvania.

[2] Moon Ryul Jung and Norman Badler. *Posture Planning for Agent Animation.* To Appear in *The First International Conference on AI Planning Systems*, College Park, Maryland, June, 1992.

[3] Moon Ryul Jung, Jugal Kalita, Norman Badler, and Wallace Ching. Simulating Human Tasks Using Simple Natural Language Instructions. In *Winter Simulation Conference*, Phoenix, Arizona, December 1991.

Shyam Kapur

Postdoc, Institute for Research in Cognitive Science
skapur@linc.cis.upenn.edu

Formal Modeling of Natural Language Acquisition

Keywords: Formal Learning Theory, Natural Language Acquisition

My Ph.D. thesis [1] is a formal investigation of certain aspects of Language Learning. Variations of some existing models along the directions suggested by studies in Natural Language Acquisition have been studied. We have obtained characterizations for families of languages learnable under some commonly assumed constraints [3]. We have also incorporated a stochastic element in our model (along the lines of *pac*-learning) and obtained learning algorithms that exploit 'indirect negative evidence' generated using this assumption. In addition, we have looked at the applications of formal principles, such as the *Subset Principle*, and pinpointed the inaccuracies in the past applications and suggested alternatives [4].

In a further development of the construction in my thesis, we have recently shown that learning can succeed when only this kind of indirect negative evidence is used. At each step of the learning algorithm a prediction made on the basis of the current grammar is put to test. The outcome completely determines the behavior of the learner. In this construction, the extensional relationship between various possible languages ceases to be a parameter. The learning algorithm is uniform, simple and robust.

We expect to develop the learning model using input from different directions. One part of research is involved in determining how well this learning model is consistent with existing data on child language acquisition. In order to verify some predictions such a model makes, we hope to conduct new psychological studies. We also intend to run the learning algorithms on large corpora of text from different languages. It is expected that as a byproduct of this procedure new understanding of the principles and parameters of the Universal Grammar may be obtained [2].

References

[1] Shyam Kapur. *Computational Learning of Languages.* PhD thesis, Cornell University, September 1991. Technical Report 91-1234.

[2] Shyam Kapur. Some applications of formal learning theory results to Natural Language acquisition. To be presented at a symposium at Cornell University on 'Syntactic Theory and First Language Acquisition: Cross Linguistic Perspectives', April 1992.

[3] Shyam Kapur and Gianfranco Bilardi. Language learning without overgeneralization. In *Proceedings of the 9th Symposium on Theoretical Aspects of Computer Science (Lecture Notes in Computer Science Series).* Springer-Verlag, 1992.

[4] Shyam Kapur, Barbara Lust, Wayne Harbert, and Gita Martohardjono. Universal grammar and learnability theory: the case of binding domains and the subset principle. In *Knowledge and Language: Issues in Representation and Acquisition.* Kluwer Academic Publishers, 1992. In preparation.

Michael H. Kelly

Assistant Professor, Department of Psychology
kelly@cattell.psych.upenn.edu

Phonology and Grammatical Category

Keywords: Language Processing, Acquisition, Phonology, Lexical Structure

During language acquisition, children must learn the grammatical categories of their language, such as noun and verb, and how to assign words to the appropriate class (e.g., "car" is a noun, "go" is a verb). Adults also must assign words to grammatical classes quickly and accurately. In understanding how these tasks are accomplished, researchers try to determine what sources of information are available and used to make grammatical category assignments. Most research on this question focuses on semantic and syntactic information for grammatical class. My own research examines a relatively neglected source of such information, namely phonology. Perhaps certain phonological features are correlated with different grammatical classes, with the concomitant possibility that children and adults have learned and use these correlations.

My investigations of this question have led me to conclude that phonology has been seriously underestimated as an informational source for grammatical class. In particular, large correlations between phonology and grammatical class exist, they can involve thousands of words in the lexicon of a language, children and adults have implicit knowledge of these correlations, and specific hypotheses about the causes of these correlations can be proposed and evaluated. For example, disyllabic nouns and verbs differ in stress in English, with nouns being more likely to have first syllable stress (e.g., compare the pronunciations of "record" in "I bought a record at the store" versus "I will record the concert"). In a variety of experiments, I have shown that adult English speakers have knowledge of this correlation. For example, disyllabic pseudowords such as "bontoon" are more likely to be pronounced with first syllable stress if they act as nouns in sentences than if they act as verbs. In addition to examining speaker knowledge of this stress difference, we have explored a possible basis for its evolution in English. In particular, we have argued that the noun-verb stress difference is due to two factors: (a) a general preference for rhythmic alternation in language, and (b) the tendency for verbs to be more likely than nouns to appear in rhythmic contexts that bias them toward second syllable stress. Using this general hypothesis, we have discovered large rhythmic contexts differences between nouns and verbs, and have used these contexts to predict where the noun-verb difference should be strongest in the English lexicon. Furthermore, a variety of experiments have indicated that stress patterns on words can be altered in predicted ways by the rhythmic contexts in question. These experiments demonstrate that the causes of certain phonological differences between nouns and verbs can be elucidated experimentally. More generally, various characteristics of language change can be subjected to standard psycholinguistic experimental methods. In the future, we plan to relate these various phenomena to models of lexical representation and access in human memory.

References

[1] Michael H. Kelly. Rhythmic alternation and lexical stress differences in English. *Cognition*, 30, 107-137, 1988.

[2] Michael H. Kelly and J. Kathryn Bock. Stress in time. *Journal of Experimental Psychology: Human Perception and Performance*, 14, 389-403, 1988.

[3] Michael H. Kelly. Rhythm and language change in English. *Journal of Memory and Language*, 28, 690-710, 1989.

[4] Kelly, M. H. Using sound to solve syntactic problems: The role of phonology in grammatical category assignments. *Psychological Review*, in press.

Anthony S. Kroch

Professor, Department of Linguistics
kroch@linc.cis.upenn.edu

Formalizing the Theory of Grammar Using TAG

Keywords: Grammar Formalisms, Theory of Grammar, TAG

Research over the last several years has demonstrated the utility of the TAG formalism in empirical research on natural language syntax. Use of the formalism allows the linguist to capture syntactic generalizations that would be represented as constraints on movement in a transformational grammar rather than as constraints on the well-formedness of elementary structures. Due to the limited generative capacity of TAG, the theory of grammar that results from the change in representation is more highly constrained than standard transformational theory. For example, the principle of subjacency, independent of other principles in a transformational grammar, falls out as a corollary of the TAG formalism. The possibility of successfully translating well-motivated transformational analyses into TAG has been demonstrated for a number of constructions, among which the most important has been wh- movement, including such complex features of the construction as the parasitic gap phenomenon and the phenomenon of long movement. In addition, there has been work on NP movement, on extraposition, and on the complex West Germanic verb-raising construction. Recent research on scrambling has also yielded promising results. Currently, our research on the linguistic application of the TAG formalism centers on specifying in detail the proper representation of elementary TAG structures and evaluating the empirical utility of various extensions to the formalism, with the aim of improving the coverage and conceptual elegance of TAG analyses of core grammatical phenomena.

Patterns of Grammar in Language Use and Change

Keywords: Statistics and Language, Language Change

Work on the history of the English auxiliary system has revealed a surprising statistical pattern in the frequency of use of modern versus Middle English forms. When sentences from the late Middle English corpus are grouped by sentence type into negative interrogatives, affirmative interrogatives, negative declaratives, and affirmative declaratives, the frequency of use of the periphrastic auxiliary *do* differs substantially by type. This difference follows the ordering given; and under assumptions long standard in studies of language change, the ordering of frequencies would be taken to reflect a temporal ordering of contexts. Specifically, the use of *do* would have been supposed to enter the language context by context following the frequency ordering, and the rate of spread would have been differentiated by context in the same way. Statistical analysis, however, reveals that the rate of spread of the *do* form is the same in all contexts. Furthermore, this rate is the same as that of the spread of preverbal positioning of prosodically weak sentential adverbs, which, under a well-motivated and standard syntactic analysis, is a reflex of the same grammatical change as the one that motivates the use of periphrastic *do*. The parallelism across contexts suggests the following "constant rate" hypothesis for language change: When alternations in different surface

contexts reflect competition at a single locus in an underlying grammatical system, the rates of change in the frequencies of the alternating forms over time will be the same for all contexts. In other words, change takes place at the level of the grammar, not at the level of the surface contexts where its effects are observed. The first support beyond the original Middle English case found for the constant rate hypothesis was in certain previously described historical changes in Portuguese and French. More recently, the hypothesis has received further support in diachronic studies of Old English and Yiddish phrase structure that were specifically designed to test it.

References

[1] Frank, Robert. 1991. "Parasitic Gaps and Locality Conditions." In *Proceedings of the 27th Regional Meeting of the Chicago Linguistic Society.* University of Chicago.

[2] Kroch, Anthony S. and Joshi, Aravind. 1985. "The Linguistic Relevance of Tree Adjoining Grammar." University of Pennsylvania Department of Computer and Information Science Technical Report no. MS-CIS-85-16.

[3] Kroch, Anthony S. 1987. "Unbounded Dependencies and Subjacency in a Tree Adjoining Grammar." In A. Manaster-Ramer, (ed.), *Mathematics of Language.* Philadelphia: John Benjamins.

[4] Kroch, Anthony S. and Joshi, Aravind. 1987. "Extraposition in a Tree Adjoining Grammar." In Huck, et al. (eds.), *Proceedings of the Conference on Discontinuous Constituents. Syntax and Semantics*, vol. 20. New York: Academic Press.

[5] Kroch, Anthony S. 1989. "Asymmetries in Long Distance Extraction in a Tree Adjoining Grammar." In M. Baltin and A.S. Kroch, (eds.), *Alternative Conceptions of Phrase Structure.* Chicago: University of Chicago Press, pp. 66-98.

[6] Kroch, Anthony S. 1989. "Reflexes of Grammar in Patterns of Language Change." *Language Variation and Change.* vol 1, no. 3, pp. 199 – 244.

[7] Kroch, Anthony S. 1989. "Language Learning and Language Change." Comments on Lightfoot, D. "The Child's Trigger Experience: Degree 0 Learnability." *The Behavioral and Brain Sciences*, vol. 12, no. 2, pp. 348 – 9.

[8] Kroch, Anthony S. 1989. "Function and Grammar in the History of English: Periphrastic 'do'." In Fasold and Schiffrin, (eds.), *Language Variation and Change: Current Issues in Linguistic Theory.* vol 52. Philadelphia: John Benjamins, pp. 133 – 172.

[9] Kroch, Anthony S. and Santorini, Beatrice. 1991. "The Derived Constituent Structure of the West Germanic Verb Raising Construction." In Freidin, (ed.), *Principles and Parameters in Comparative Grammar.* Cambridge: MIT Press, pp. 269 – 338.

[10] Lee, Young-Suk and Rambow, Owen. 1992. "Scrambling in Korean: a TAG Analysis". Paper presented at the annual meeting of the Linguistic Society of America, Philadelphia.

[11] Pintzuk, Susan. 1991. *Phrase Structures in Competition: Variation and Change in Old English Word Order.* University of Pennsylvania Dissertation.

[12] Pintzuk, Susan. 1992. "Phrase Structure Change in the History of Old English." to appear in Language Variation and Change.

[13] Santorini, Beatrice. 1992. "Phrase Structure Change in the History of Yiddish." to appear in Language Variation and Change.

Chang-Bong Lee

Department of Linguistics
cblee@unagi.cis.upenn.edu

A Tag Analysis of 'Tough' Constructions in English

Keywords: TAG Analysis

The paper attempts to provide a unified account of the so-called 'Tough' constructions in English, using the framework of Tree Adjoining Grammar (TAG). 'Tough' constructions have been paid great attention in linguistic theory since they show an interesting linguistic property of filler-gap dependency. In this paper, I will first present the previous analysis of 'Tough' construction done in the framework of GPSG and point out its limitations. It will be shown that such a context-free formalism as GPSG is not sophisticated enough to capture the linguistic fact of 'Tough' construction. This problem and some other motivations lead us to use a new formalism, the TAG analysis developed by Joshi(1975) and others, which is mildly context-sensitive in its nature. In particular, my analysis will be based on a recent extension of TAG, Lexicalized Unification-Based Tree Adjoining Grammars (UTAG) developed in Vijay-Shanker and Joshi(1988). I will argue that to properly deal with 'Tough' constructions we need two properties of TAG formalism. One is that in TAG the constraints are specified at the domain of locality and they are expanded in an unbounded way by applying the adjoining operation. This property of TAG is contrasted with that of the context-free grammar (CFG) formalism; for instance, in GPSG the constraints are implemented in each rule by some kind of feature- specification, which is passed up and down around the tree. It will be argued that the TAG formalism has a crucial advantage over CFG formalism in giving a simple unified account of 'Tough' constructions. The second property is the extra mildly context-sensitive mechanism of UTAG; namely, the constraint on the adjoining operation with respect to feature-specification on the node to be adjoined. I will propose a feature called GAP to be used to capture the linguistic information as to whether or not the given node has a gap above or below in the rest of the tree. It will be argued that interacting with this feature specification, the theory-internal constraint on the adjoining operation in TAG contributes to correctly deriving only well-formed sentences and rules out the ill-formed ones with respect to 'Tough' constructions.

References

[1] Abeille, A, Bishop, K, Cote, S, and Schabes, Y. *A Lexicalized Tree Adjoining Grammar For English.* Technical Report MS-CIS-90-24, University of Pennsylvania 1990.

[2] Chae, H.R. Gap Licensing in Tough and Similar Constructions. *Unpublished ms. Dept. of Linguistics, Ohio State University.*

[3] Gazdar, Klein, Pullum and Sag. *Generalized Phrase Structure Grammar.* Blackwell, 1985.

[4] Joshi, A, Levy, L.S and Takahashi, M. Tree Adjunct Grammars. *Journal of Computer Systems Science*, 1975.

[5] Kroch, A and Joshi, A. *Linguistic Relevance of Tree Adjoining Grammars.* Technical Report MS-CIS-85-18, University of Pennsylvania, 1985.

[6] Kroch, A. Unbounded Dependencies and Subjacency in a Tree Adjoining Grammar. *Mathematics of Language*, 1988.

[7] Quirk, Gech and Svartvik. *A Comprehensive Grammar of the English Language.* Longman, 1985.

[8] Vijay-Shanker, K and Joshi, A. Feature Structure Based Tree Adjoining Grammars. In *Proceedings 12th International Conference on Computational Linguistics.* Budapest, 1990.

Young-Suk Lee

Department of Linguistics
ysl@linc.cis.upenn.edu

Scrambling and Movement Types

Keywords: Operator, Non-operator, A-movement, A′-movement

In the history of generative syntax, there have been several major issues concerning scrambling phenomena in languages such as German, Hindi, Japanese, Korean, Turkish, etc. In these languages, ordering of the arguments of a verb is relatively free compared to English, indicating that a clause in these languages is less structured than in English. The first issue, which became known as the *configurationality parameter*, in the early 1980's, is whether scrambling is due to base-generation (cf. [2]), or to movement. The second issue, on which more recent studies on scrambling have focused, is what type of movement scrambling belongs to, e.g. A vs. A′-movement (under the assumption that it is movement cf. [6], [4], [5]). A third issue is whether or not local and long distance scrambling are the same syntactic phenomenon.

In my thesis, I primarily address the issue of what type of movement scrambling is, by examining the nature of the position occupied by a scrambled element with regard to scope and binding facts. This provides the answer for the question of uniformity of local and long distance scrambling as well. I also consider constraints on scrambling, and conclude that the absence/presence of the category AGR leads to parametric variations among scrambling languages with respect to the extraction domain, and the absence/presence of subject/object asymmetry in scrambling.

I argue that scrambling is movement of a third type, namely, movement to a non-operator A′-position, along the line of [3]. The operator/non-operator distinction is based upon a semantic operator-variable relation at LF, and the A/A′-distinction, on θ-/non-θ-position, as defined in [1]. Crucial evidence for scrambling being non-operator movement comes from interpretations of scrambled wh-phrases, negative polarity items, and quantifiers. A scrambled scope element may undergo reconstruction to its base-position, and the position occupied by a scrambled element itself behaves as a variable. Examination of the data concerning weak crossover and strong crossover suggests that scrambling is A-movement, while considerations of parasitic gap facts and the existing definitions of A-movement/A-chain indicate that scrambling is A′-movement. Given this, I conclude that scrambling is movement to a non-operator A′-position, in which non-operator positions subsume the classical A-positions.

As for constraints on scrambling, in Korean there is no subject/object asymmetry, and a scrambled clause does not form an island for further extraction. I derive these properties by hypothesizing that the verb governs all of its arguments including the scrambled ones at S-structure, which is in turn due to the absence of the category AGR in the Korean language. This hypothesis is supported by the distribution of nominative case in untensed clauses and event nominal clauses, and gives a natural explanation for the differences between German and Korean concerning domain of and constraints on scrambling.

References

[1] Noam Chomsky. *Lectures on Government and Binding*, Foris: Dordrecht, 1981.

[2] Ken Hale. Preliminary Remarks on Configurationality. In *Proceedings of NELS*, 1982.

[3] Robert Frank, Young-Suk Lee and Owen Rambow. Scrambling as non-operator movement and the special status of subjects. In *Proceedings of the Third Leiden Conference for Junior Linguists*, 1992.

[4] Anoop Mahajan. *The A/A-Bar Distinction and Movement Theory*. PhD thesis, MIT, 1990.

[5] Mamoru Saito. Long distance scrambling in Japanese. In *Journal of East Asian Linguistics*, Kluwer Academic Publishers, 1992.

[6] Gert Webelhuth. *Syntactic Saturation Phenomena and the Modern Germanic Languages*. PhD thesis, University of Massachussetts at Amherst, 1989.

Libby Levison

Department of Computer and Information Science
libby@linc.cis.upenn.edu

Lexical Semantics of Instructions for Animation

Keywords: Natural Language Processing, Lexical Semantics, Graphics, User Interfaces

I am interested in improving human-computer interaction by augmenting machine understanding of human instructions and incorporating graphics into human-computer interfaces. My current research is in understanding instructions for the purpose of generating animations. As a member of the **Anim**ation and **N**atural **L**anguage project (AnimNL) (see Badler, Baldwin, Di Eugenio, Geib, Jung, Moore, Webber and White), I work between the Language, Information and Computation (LINC) Lab and the Graphics Lab, building animation definitions of those instructions which result in physical actions.

If the purpose of an interaction is cooperation on a task, the computer must understand the user's instructions and act appropriately. This is relevant at the level of the human-computer interface as well as in the domain of the AnimNL project, in which we want to instruct a graphics program to generate certain animations. In building a system which will interpret the user's instruction, I am specifically interested in verb-object relations; I have identified wide variations in intended action which occur when a verb appears with different objects. For example, the verb *open* is associated with two distinct physical actions in the instructions *open the door* and *open the soda can.* If we believe that each verb has a unique meaning then we must account for these variations in interpretation at the sentence or the instruction level. I would argue that each verb has a partial, core meaning; this meaning is completed in an utterance with information carried by the verb's object, as well as by understanding the *intention* of the given instruction. For example,the definition of *open* might be something like: *provide access to.* One fact that I know about the door to my apartment, either from living in my house or through visual perception, is the door's degrees of freedom. (I might also know that this is a heavy door and that it is hung to swing shut if not propped open.) If part of the definition of *open* includes moving its object, then interpreting how to *open my door* entails checking how my door moves – what its degrees of freedom are – either translation or rotation. If the door is marked as allowing translations then I probably have a sliding-glass door; not only do rotations indicate a hinged door, but positive rotation implies pulling, while negative requires pushing. These observations suggest that building animation definitions based solely on the verb or exclusively on the object won't work. I advocate instead a hybrid system in which the core meaning of the verb makes use of (geometrical) information associated with the object.

Investigating these definitions requires an application that allows the user to give the system task instructions, as well as providing the user an easy way to check the computer's interpretation of those instructions – in other words, to verify that the correct action is performed. The animation of repair instructions satisfies this requirement: to generate an animation the computer must understand the instructions, and the resulting animation provides an easy way for an engineer to (visually) check the correctness of both the original instructions and the interpretation – the resulting simulation. (In addition, this application has real-world utility: rather than read an instruction manual, a technician or trainee can watch an animation of a simulated agent performing a repair or maintenance task.)

My research uses the *Jack*™ modeling system (a software package developed in the Computer Graphics Research Lab), which provides 3D-modeling capabilities as well as extensive human factors and anthropometric analysis tools. At the same time that I am examining linguistic issues in the instructional texts, I am investigating methodologies which will enable an engineer to produce simulations of task-level actions despite possibly limited knowledge of low-level animation techniques. I am using a minimal set of *animation directives*, (*Jack* animation instructions like **`move left foot`** or **`bend torso`**) to define higher-level actions such as **grasp**, **attach** or **open**. I call these composites **task-actions**. I hope to provide a richer set of task-actions definitions as well as a utility for defining new task-actions. These action descriptions, from the viewpoint of animation, will allow an engineer with minimal knowledge of graphics to generate animations. The interpretation process will save the engineer from defining multiple animation procedures such as *open-door*, *open-book* and *open-jewelry-box*. I am trying for an economy of action definitions, relieving the engineer of the burden of specifying detail which the system might well be able to deduce.

In summary, then, I believe that I can classify both the verbs and their objects in instructional texts according to their lexical semantics: the verbs based on the underlying physical action, the objects dependent on geometrical information. I am building a high-level utility, within the *Jack* framework, which will determine, in a given instantiation, exactly how to apply the verb to its object by reasoning about such things as the geometry of the object. I will use *Jack* animation directives – primitives which describe high-level motor control – to build compositional definitions of the physical actions underlying the instructional verbs. These *task-actions* will describe the tasks to be performed at a high-level and not on a movement-by-movement basis.

I hope to relate this work in instruction understanding and animation to research in user interfaces. Imagine a user who requests: *print my file*. Simply sending the file to the printer is adequate for a brief C program, but is not the most felicitous act if the file is a LaTeX file; similarly, the system should prevent the user from printing a `.obj` file. Instead the system can check the type of the file, and either format it correctly before printing or suggest an alternative action. This functionality will allow the user to describe the high-level intended action, while the system determines the low-level details based on the central meaning of *print* and "geometrical" knowledge of the file in question.

References

[1] Norman Badler, Bonnie Webber, Jugal Kalita, and Jeffrey Esakov. Animation From Instructions. In N. Badler, B. Barsky, and D. Zeltzer, editors, *Making Them Move: Mechanics, Control and Animation of Articulated Figures*, pages 51–93. Morgan-Kaufmann, 1990.

[2] Barbara Di Eugenio, and Michael White. On the Interpretation of Natural Language Instructions. Submitted to *1992 Int. Conf. on Computational Linguistics* (COLING-92), Nantes, France, July 1992.

[3] Barbara Di Eugenio and Bonnie Webber. Plan recognition in Natural Language instructions, in *First International Conference on AI Planning Systems*, 1992.

[4] Barbara Di Eugenio. Understanding Natural Language instructions: the case of purpose clauses, in *Proceedings of the ACL*, 1992.

† *Jack* is a trademark of the University of Pennsylvania.

[5] Jeffrey Esakov and Norman Badler. An Architecture for High Level Task Animation Control. In P.A.Fishwick and R.S.Modjeski, editors, *Knowledge-Based Systems: Methodolody and Applications*. Springer Verlag, 1989.

[6] Libby Levison. Action Composition for the Animation of Natural Language Instructions. Technical Report MS-CIS-91-28, University Of Pennsylvania, August, 1991.

[7] B. Webber, N. Badler, B. Di Eugenio, L. Levison, and M. White. Instructing Animated Agents. In *Proceedings of the First US-Japan Workshop on Integrated Systems in Multi-Media Environments*, Las Cruces, NM, December 1991.

Mark Liberman

Professor, Department of Linguistics
myl@unagi.cis.upenn.edu

Language Sound Structure

Keywords: Phonetics, Phonology

Phonetic variation sometimes seems to be an annoyance, or even an embarrassment to the elegant structures of phonological description. On the contrary, the infinite variability of phonetic interpretation rescues phonological categories and relations from the problem of their essential finiteness. Phonetic variation provides an inexhaustible body of evidence whose statistical structure reveals the nature of the underlying processes, and can thus can help settle questions about phonology that might otherwise depend on evaluating the elegance of alternative accounts of the structure of a finite set of word forms.

At least, this ought to be the case. There are some problems: often, the physical measurements that we would like to have are difficult to get; very large amounts of data are usually required, due to the complexity of the underlying processes; finally, since phonetics deals with what happens when people actually talk, its interpretation requires consideration of many things besides the structure of the linguistic message narrowly conceived.

One way forward is to concentrate on cases where easily-derived acoustic measurements are fairly close to linguistically-motivated dimensions (e.g. vowel formants, F0); to use experimental designs that maximize variation in dimensions that help choose among alternative models, while minimizing sources of unmodeled variation; and to use computer technology to make the collection and interpretation of large data sets as efficient as possible. I've been exploring this area of research since the mid 70's, largely through models of prosody (pitch and timing). Since coming to Penn in 1990, I've tried to design Linguistics classes so that students will be able to work on real questions of this type by the end of the second semester. This spring, we're concentrating on problems that arise in the phonology and phonetics of Igbo, especially vowel harmony and tone.

Speech and Natural Language Technology

Keywords: Speech Recognition, Speech Synthesis, Text Understanding

While working at AT&T Bell Laboratories (1975-1990), I spent much of my time developing and implementing speech and NLP technology, mainly in the area of speech synthesis, but also to some extent in speech recognition and text analysis. Like most of my colleagues in these fields, I learned that the most efficient way to build the best-performing systems was to rely on models derived from large bodies of speech and text.

One problem that became apparent was the difficulty of acquiring adequate corpora for research and development. Such acquisition (including necessary "clean up" and annotation efforts) is unglamorous, time consuming and expensive. Nevertheless, the performance of inductive algorithms is directly dependent on the amount of data they are based on. In the mid-80's there was a great deal of duplication of effort; no one had as much data as they wanted; smaller groups, especially in universities, often had a hard time getting started at

all; comparison of competing algorithms was difficult because they were usually trained and tested on different bodies of mutually-unavailable material. The experience of the DARPA-sponsored speech recognition effort provided a positive example of how valuable shared data could be in fostering a research community as well as producing concrete results.

For all these reasons, I began working several years ago on efforts to produce and distribute large-scale resources for research in speech and natural language technology. I helped to found and run the ACL Data Collection Initiative, which is now centered here at Penn, funded by grants from GE and NSF; I serve on the boards of the Center for Lexical Research, which Yorick Wilks directs at New Mexico State, and the Penn Treebank, directed by Mitch Marcus; we are providing four gigabytes of English text for a DARPA-organized project on document retrieval, routing, and understanding; and Penn has been designated as the host institution for the DARPA-initiated Linguistic Data Consortium. Although a great deal of work remains to be done, we have come a long way in providing the infrastructure for research and development in this area of work. Penn is playing a leading role both in developing the resources for such research, and in exploring the research problems themselves.

Models of Linguistic Inference

Keywords: Language Learning, Linguistic Theory

For entirely practical reasons, the last decade has seen an upsurge of engineering interest in models of speech and language that learn crucial parameters by statistical induction from large bodies of speech or text. Such models are favored simply because they are cheaper to produce and maintain, and work better.

Having participated in this "sea change" through engineering work in speech synthesis, speech recognition, and text analysis, I've been interested in exploring the lessons it offers for linguistic theory. The most obvious one concerns the famous question of "negative evidence," which obviously has a very different status in abstract models of language that induce (or even bound) a probability measure over the infinite set of sentences that they admit. Under appropriate assumptions, access to positive evidence in such cases can provide the same information as access to negative evidence.

A second important issue is the distinction between the number of parameters in a model and the inherent complexity of inducing them from (possibly noisy) evidence. There can obviously be cases where a very large number of parameters are computationally easy to estimate, given adequate data; and also cases where optimal estimation of a relatively small number of parameters is computationally intractable.

In general, it seems to me that linguistic argumentation about language learning over the past few decades has been based on an unwisely narrowed conception of the inductive process and its outcome. Broadening the horizons a bit is likely to lead to quite different conclusions, or at least different boundary conditions on theorizing.

References

[1] Cecil H. Coker, Kenneth W. Church and Mark Y. Liberman. Morphology and Rhyming: Two Powerful Alternatives to Letter-to-Sound Rules for Speech Synthesis. In *Proceedings of the ESCA Workshop on Speech Synthesis*, 1990.

[2] Joan Feigenbaum, Rebecca Wright and Mark Liberman. Crypographic Protection of Databases and Software. In J. Feigenbaum and M. Merritt, editors, *Distributed Computing and Cryptography*, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, AMS and ACM, v. 2, pages 161-172, 1991.

[3] S. E. Levinson, A. Ljolje and Mark Liberman. Speaker Independent Phonetic Transcription of Fluent Speech for Large Vocabulary Speech Recognition. in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, 1989.

[4] Mark Liberman. Text on Tap: the ACL Data Collection Initiative. in *Proceedings of DARPA Workshop on Speech and Natural Language Processing*, 1989.

[5] Mark Liberman and Kenneth W. Church. Text Analysis and Word Pronunciation in Text-to-Speech Synthesis. In S. Furui and M. Sondhi, editors, *Advances in Speech Technology*, Marcel Dekker, 1991.

[6] Mark Liberman The Trend Towards Statistical Models in Natural Language Processing. In E. Klein and F. Veltman, editors, *Natural Language and Speech*, Springer-Verlag, 1991.

[7] Mark Liberman and Richard Sproat. The Stress and Structure of Modified Noun Phrases in English. In I. Sag and A. Szabolsci, Eds., *Lexical Matters*, Stanford, 1992.

[8] Mark Liberman and Cynthia McLemore. The Structure and Intonation of Business Telephone Greetings. submitted to Penn Review of Linguistics, 1992.

[9] Evelyn Tzoukermann and Mark Liberman. A Finite-state Morphological Processor for Spanish. In *Proceedings of COLING90*, 1990.

Mitch Marcus

Professor, Department of Computer and Information Science
mitch@linc.cis.upenn.edu

Deducing Linguistic Structure from Large Corpora

Keywords: Parsing, Stochastic Natural Language Processing, Automatic Language Acquisition, Annotated Databases

Automatic Acquisition of Linguistic Structure

Within the past several years, a widening circle of researchers have begun to investigate a new set of techniques for the use of trainable systems in natural language processing. The early successes of these new techniques, coupled with other advances, have allowed the emergence of a new generation of systems that both extract information from and summarize pre-existing text from real-world domains.

A group of us at Penn have initiated a research program to see how far the paradigm of trainable systems can take us towards the fully automatic syntactic analysis of unconstrained text and towards the automatic acquisition of grammatical structure from both annotated and unannotated text corpora. This research investigates both statistical and symbolic learning methods using both supervised and unsupervised approaches. We are operating under the assumption that this work should proceed by attempting to combine two different traditions often viewed as mutually exclusive: the research program of generative grammar, as set forth originally by Noam Chomsky, and the research paradigm of distributional analysis, as developed by the American structural linguists resulting in the mathematical and computational work of Zellig Harris. For an overview of this point of view, see [8].

Our research into distributional analysis has already yielded results which are both surprising and encouraging. We have investigated how accurately the phrase structure of sentences taken from unconstrained free text can be determined without an explicitly encoded grammar at all, using only automatically compiled distributional statistics of a corpus of text tagged for part of speech (POS). Two years ago, we reported results [7] on a new sentence analyzer which subdivides text tagged for POS into smaller and smaller (unlabelled) grammatical constituents. On a reserved test set, this parser misplaced about 2 to 3 brackets per sentence for sentences of length less than 15 words, and about 5 to 6 brackets on sentences from 30 to 60 words in length.

This year, we have developed a new distributional technique [3] for discovering a grammar from a corpus of material tagged only with POS data. This technique provides complete binary-branched parses for POS tag sequences; for short sentences (6-15 words) the resulting parses are consistent with the input tag sequence 71% of the time for simple sentences excluding sentences with coordinate structures or quotations. An additional 11% of sentences are incorrect by one bracket. Work is under way to extend this technique to handle a wider range of the more difficult grammatical phenomena.

To allow this technique to be applied to completely unannotated text, Brill is concurrently experimenting with techniques to automatically derive a tag set for a corpus of text, again using only distributional facts [2]; results to date are very encouraging. Brill has also developed a part of speech tagger [4] which uses only a single best-guess for part of speech

for each word in its lexicon, coupled with a set of non-probabilistic correction rules which the tagger learns itself. To date, this algorithm has a 95% correct tagging rate on a set of 80 tags with a very limited dictionary, tested on a reserved test corpus, after acquiring about 80 rules at the end of about 24 hours of computation on a Sun 4/490. This result hints that perhaps the success of recent work in the automatic acquisition of linguistic structure is due in large measure to the methodology of trainable systems, and less so to the use of explicitly stochastic techniques.

Stochastic Parsing

In an experiment last year, we investigated how distributional facts can be used to choose between the multiple grammatically acceptable analyses of a single sentence. The resulting parser, Pearl, [6] differs from previous attempts at stochastic parsers in that it uses a richer form of conditional probabilities based on context to predict likelihood. Tested on a naturally-occurring corpus of sentences requesting directions to vary locations within a city (the MIT Voyager corpus), the parser correctly determined the correct parse (i.e. gave the best parse first) on 37 of 40 sentences. We are now beginning a collaboration with the Continuous Speech Recognition Group at IBM's Thomas Watson Laboratory to develop a new generation of stochastic parsers, based on decision tree technology utilizing a rich set of linguistic predicates, and trained on output from both the Penn Treebank (see below) and the Lancaster Treebank. (A first version of such a parser [1] developed at IBM last summer, with Magerman's participation, can be viewed as an extension of Pearl.)

The Penn Treebank Project

For the last several years, a group of us have been working on the construction of the Penn Treebank, a data base of written and transcribed spoken American English annotated with detailed grammatical structure. This data base, although now only in preliminary form, has been distributed to a wide variety of groups in the US and elsewhere, providing training material for a wide variety of approaches to automatic language acquisition, a reference standard for the rigorous evaluation of some components of natural language understanding systems, and a research tool for the investigation of the grammar and prosodic structure of naturally spoken English.

The preliminary corpus now consists of over 4 million words of running text annotated with POS tags, with a subset of 1.4 million words now assigned skeletal grammatical structure. Annotation is first done using software provided by AT&T Bell Labs, and then hand-corrected. The annotation team currently consists of Rob Foye, Lisa Frank, Leslie Dossey, Robert MacIntyre, Victoria Tredinnick, and Alissa Hinckley and is supervised by Mary Ann Marcinkiewicz. Among other materials, the corpus now includes 1 million words of Dow-Jones News Service articles (from the ACL-DCI) annotated for both POS and skeletal grammatical structure, and an additional 1.5 million words annotated only for POS, as well as a retagged version of the Brown corpus, with half in skeletal syntactic analysis as of 3/92. Included on the first ACL-DCI CD-ROM sampler are 3.5 million words of Penn Treebank material annotated for POS, and a very small sample of skeletally parsed material. In addition to the CD-ROM, over 30 copies of parts of the preliminary corpus have been distributed directly by us to date to groups in the U.S., Europe and Asia. We hope to annotate another 2.5-3 million words before the conclusion of the first phase of the Treebank project at the end of this calendar year.

We look forward to distributing our preliminary corpus as widely as possible, and will

freely distribute most of it for use on an "as-is" basis. To obtain a copy of the corpus, send email to maryann@unagi.cis.upenn.edu.

References

[1] E. Black, F. Jelinek, J. Lafferty, D. Magerman, R. Mercer, and S. Roukos, Towards History-based Grammars: Using Richer Models for Probabilistic Parsing, *M. Marcus, editor, Proceedings, DARPA Speech and Natural Language Workshop*, February, 1992, Morgan-Kaufmann, to appear.

[2] E. Brill, Discovering the Lexical Features of a Language, *Proceedings, 29th Annual Meeting of the Association for Computational Linguistics (ACL)*, Berkeley, June 1991.

[3] E. Brill, Using Distributional Analysis to Automatically Acquire Phrase Structure, *M. Marcus, editor, Proceedings, DARPA Speech and Natural Language Workshop*, February, 1992, Morgan-Kaufmann, to appear.

[4] E. Brill, A Simple Rule-based Part of Speech Tagger, Proceedings of the Third Conference on Applied Computational Linguistics, Trento, Italy, to appear. Also, *M. Marcus, editor, Proceedings, DARPA Speech and Natural Language Workshop*, February, 1992, Morgan-Kaufmann, to appear.

[5] E. Brill, D. Magerman, M. Marcus and B. Santorini, Deducing linguistic structure from the statistics of large corpora, *Proceedings, DARPA Speech and Natural Language Workshop*, June, 1990, Morgan-Kaufmann.

[6] D. Magerman and M. Marcus, PEARL — Probabilistic Chart Parser, *Proceedings, Fifth Annual Meeting of the European Association for Computational Linguistics (EACL)*, Berlin, April 1991.

[7] D. Magerman and M. Marcus, Parsing a natural language using mutual information statistics, *Proceedings, Eighth National Conference on Artificial Intelligence (AAAI 90)*, 1990, AAAI Press/MIT Press.

[8] M. Marcus. Directions in Natural Language Understanding. In *1990 W. Gear, editor, Computation & Cognition, Proceedings of the First NEC Research Symposium*, SIAM, 1991.

Cynthia A. McLemore

Postdoc, Institute for Research in Cognitive Science
cam@unagi.cis.upenn.edu

Intonation and Discourse Interpretation

Keywords: Intonation, Prosody, Phonology, Pragmatics

My research areas include intonation, phonetics and phonology, sociolinguistics, and pragmatics. Current work focuses on the role of intonation in discourse interpretation; in addition, I have been investigating the status of the prosodic hierarchy in grammar, in collaborative work with Mark Liberman (Penn) and Anthony Woodbury (Texas).

References:

1991a. *The Interpretation of English Intonation: Sorority Speech.* Austin: University of Texas Ph.D. dissertation.

1991b. "On the Nature of Prosodic Phrasing" (with M. Liberman and A. Woodbury). Paper presented at Declarative Perspectives on the Syntax-Prosody Interface, Research Institute for Language and Speech, Rijksuniversiteit Utrecht, The Netherlands.

1991c. "Boundary Tone Values and Shared Knowledge." Paper presented at Grammatical Foundations of Prosody and Discourse, University of California at Santa Cruz.

1991d. "The Interpretation of L*H in English." In, C. McLemore, ed., *Texas Linguistics Forum 32: Discourse.* Austin: Department of Linguistics and the Center for Cognitive Science.

1991e. Editor, *Texas Linguistics Forum 32: Discourse.* Austin: Department of Linguistics and the Center for Cognitive Science.

1992a. "The Structure and Intonation of Business Telephone Greetings" (with M. Liberman). Submitted to *The Penn Review of Linguistics.*

1992b. "Situated Symbols: Patterns of Intonational Use in an All- Female Community." Paper presented at the American Association of Applied Linguistics, Seattle, WA.

In prep. "Why Rise?" (with C. Zeller and M. Liberman). University of Pennsylvania.

Carol A. Miller

Department of Psychology
cmiller@cattell.psych.upenn.edu

Symmetric Predicates

Keywords: Symmetry; Similarity; Argument Structure

Symmetry. Consider a relation r between two entities such that if A r B, then B r A. Such a relation is *symmetric*. Tversky (1977; Tverksy & Gati, 1978) have claimed that the concepts SIMILAR and DIFFERENT are not symmetric, contrary to intuition. One basis for this claim is the finding that subjects, when presented with sentences such as:

(1) North Korea is similar to China.

(2) China is similar to North Korea.

judge (1) to be more true than (2); further, (1) seems to most people to be more felicitous than (2). We argue that the basis for the intuition that similar, along with many other predicates, is symmetrical, can be reconstructed, and that the difference in interpretation between (1) and (2) is a function of linguistic rather than conceptual structure.

Apparent asymmetry of intuitively symmetric relations. Following Tversky's reasoning, if similar is not symmetric, then neither are a lot of other predicates, like equal or meet. Consider the following pairs of sentences:

(3) a The humblest citizen is equal to the president.

b The president is equal to the humblest citizen.

(4) a My sister met Meryl Streep.

b Meryl Streep met my sister.

Subjects exhibit consistent preferences for one of the sentences in each of these pairs (almost invariably the a alternative) and give clear and consistent reasons for their preferences, just as Tversky found. Yet if any predicates in the English language encode symmetric relations, surely equal and meet are among them. It is axiomatic that if A is equal to B then B is equal to A, and if my sister met Meryl Streep, Meryl could hardly avoid meeting my sister.

Intuitions about symmetry. We conducted a series of studies to map out the relationship between meaning and structure for symmetric predicates. A group of subjects was asked to rate 40 predicates, without sentential context, for symmetry. The predicates were chosen by the experimenters in the hopes of covering a broad range of degrees of symmetry, from very symmetric through borderline to not at all symmetric. Subjects' ratings did indeed cover the entire scale, and subjects agreed reliably on their ratings. On the basis of these ratings, the predicates were divided into 2 halves, and symmetric vs. asymmetric was used as a variable in the other experiments.

Consistent differences between symmetric and asymmetric predicates. The syntactic behavior of symmetric and asymmetric predicates was found to be different. According to subjects, asymmetric predicates may occur freely in intransitive constructions, but symmetric predicates may only do so when the subject position contains a plural or conjoined NP.

(5) a The swimmer and the lifeguard drowned.

b The lifeguard drowned.

(6) a My sister and Meryl Streep met.

b *My sister met.

The construction in (6a) is not truly "intransitive". At some level, the structure underlying (5a) is that in (7) (see Gleitman, 1965).

(7) The swimmer drowned and the lifeguard drowned.

Underlying (6a), however, is (8a), not (8b).

(8) a My sister met Meryl Streep and Meryl Streep met my sister.

b My sister met and Meryl Streep met.

The reciprocality underlying (6a) as opposed to (5a) is demonstrated by an experiment in which subjects compare sentences of the following types.

(9) a The swimmer and the lifeguard drowned each other.

b The swimmer and the lifeguard drowned.

(10) a My sister and Meryl Streep met each other.

b My sister and Meryl Streep met.

Sentence pairs such as (9)–those with asymmetric predicates–are judged to be different in meaning, while, sentence pairs like (10)–those with symmetric predicates–are judged to mean the same thing.

Sentence structure changes interpretation. Symmetric and asymmetric predicates, then, clearly fall into different classes on the basis of their syntactic behavior. Yet in a construction such as in (3) and (4), which we call the directional construction, reversing the order of the nouns in the sentence with a symmetric predicate yields a difference in interpretation, a fact which is also true of asymmetric predicates. The change is not as great for symmetric predicates, however. We conclude that the asymmetry of interpretation of sentences observed by Tversky is certainly real, but it can be accounted for by the syntactic structure of sentences such as those in (3) and (4), which unlike (6a), require that the nouns in the sentence, one serving as subject and the other as object, be construed as playing different roles with respect to the predicate. We believe that the symmetry of the concept SIMILAR (and other intuitively symmetric concepts) remains stable, but the interpretation of the predicate which encodes a concept can be influenced by the structure of the sentence in which it appears.

References

[1] Gleitman, L. R. Coordinating conjunctions in English. *Language*, 41(2), pp. 260-293, 1965.

[2] Tversky, A. Features of similarity. *Psychological Review*, 84, (4) pp. 327-350, 1977.

[3] Tversky, A, and Gati, I. Studies of similarity. In E. Rosch and B. Lloyd (Eds.), *Cognition and categorization* (pp. 79-98). Hillsdale, New Jersey: Erlbaum, 1978.

Michael B. Moore

Department of Computer and Information Science
mmoore@linc.cis.upenn.edu

Combinations of intensional logics

Keywords: Semantics, Context, Perception, Knowledge, Belief, Action, Instructions

In both natural human languages and formal computer languages, the notion of context plays an important role in determining the interpretation of expressions. There is a variety of different types of context that may be "consulted" in this regard including environments (in programming language theory), discourse context, temporal context, and various epistemic contexts for the agents involved. Intensional logic (a broad category of logic including modal logic, provability logic, tense logic, temporal logic, and dynamic logic) provides a mathematical framework in which to characterize these contexts. Since there are many types of context which can influence the interpretation of an expression, it is desirable to have ways of combining different notions of context into a single intensional logic. Johan van Bentham [9] provides a good survey of intensional logic and its combinations.

Previous research in knowledge representation and reasoning has studied a small number of different combinations of specific intensional logics, primarily the combination of an epistemic logic with a temporal logic into logics of knowledge and action. Robert C. Moore [6] pioneered theories of knowledge and action. Recent alternatives include work by Morgenstern [7] and Lesperance [5]. Rosenschein's situated automata [8] provide one way to directly implement a system specified by such logics. Davis [2] has contributed theories relating perception and knowledge.

My research aims at a general theory of combining different intensional logics which will be applied to develop a logic of perception, knowledge, belief, and action. It was originally motivated by problems of reference in natural language instructions which require accommodation. In the prototypical scenario of instructed action an instructor presents instructions to an agent, the agent interprets the instructions, and the agent acts based on their interpretation of the instructions. Each of these three activities (presentation, interpretation, and action) can be decomposed into sub-activities. Various sub-activities may be interleaved in different ways. I am currently exploring scenarios where the instructions can not be interpreted until after some action has been performed – in particular, scenarios where the objects the agent is instructed to manipulate are not known to the agent when the instructions are presented. The agent must then find those objects before completing the interpretation and carrying out the instructed action. This has led me to study the interaction between perception and action on the one hand and the knowledge and beliefs of the instructor and the agent on the other.

I am providing some of the logical foundations for knowledge representation and inference in the AnimNL project [1]. Intensional logic supports many different inferences throughout the process of translating from natural language expressions to human figure animations. Reasoning about future perceptual contexts supports inference about situations where the agent can expect to find referents for referring expressions and, consequently, about what actions the agent is able to accomplish. As the actions are being executed, the logical theory of the interaction between perceptions and propositional attitudes guides the design

of procedures for maintaining consistency of the world model and knowledge base of the agent. The intensional context also supports interpretation of indexical expressions.

An early product of this research has been the development of a theorem prover for combined intensional logics. I have found it easier to characterize these logics by providing an axiomatic description of the accessibility relation(s) than to provide Hilbert-style axioms for the logics directly. Other approaches to theorem proving in modal and intensional logics are discussed by Frisch and Scherl [4]. The interactive theorem prover has been implemented in PROLOG based on a Gentzen sequent calculus. The meta-interpreter for the theorem prover is implemented following the tactical style described by Felty and Miller [3]. Two sets of similar sequent rules are provided, one for proofs in the intensional logic, and one for proofs about the accessibility relation(s). The connection between these two sets of rules is through the $\Box$-left rule which has the following form:

$$\frac{\mathcal{R} \vdash_R (wRt) \qquad \Gamma \cup \{t{:}A, w{:}\Box A\} \vdash \Delta}{\Gamma \cup \{w{:}\Box A\} \vdash \Delta}$$

Previous formalisms of knowledge and action have been based on the assumption that the knowledge effects can be finitely characterized for each action. By incorporating a theory of perception into a combined intensional logic, I can represent indirect knowledge effects which have not previously been addressed. The effects of moving from one room to another include a change in the perceptual context of the agent. Since there are an infinite variety of different perceptions which may confront the agent in the next room and lead to an infinity of different knowledge states, it is unreasonable to expect to finitely represent the knowledge gained by the "move" action without using some rule-like finite encoding of infinite information. The approach I take is to represent these knowledge effects indirectly since movement changes location, location constrains the perceptual context, and perceptions ground knowledge. Changes in the perceptual context are encoded as constraints on the accessibility relations for the actions which cause the change.

References

[1] Norman Badler, Bonnie Webber, Jeff Esakov, and Jugal Kalita. Animation from instructions. In Badler, Barsky, and Zeltzer, editors, *Making them Move*. MIT Press, 1990.

[2] Ernest Davis. Inferring ignorance from the locality of visual perception. In *Proceedings AAAI*, pages 786–791, St. Paul, MN, 1988.

[3] Amy Felty and Dale Miller. Specifying theorem provers in a higher-order logic programming language. Technical Report MS-CIS-88-12, Department of Computer and Information Science, University of Pennsylvania, 1988.

[4] Alan M. Frisch and Richard B. Scherl. A general framework for modal deduction. In Allen, Fikes, and Sandewall, editors, *Principles of Knowledge Representation and Reasoning: Proceedings of the second international conference*, pages 196–207, 1991.

[5] Yves Lesperance. *A Formal Theory of Indexical Knowledge and Action.* PhD thesis, Computer Systems Research Institute, University of Toronto, 1991. also Technichal Report CSRI-248.

[6] Robert C. Moore. A formal theory of knowledge and action. In R.C. Moore and J. Hobbs, editors, *Formal Theories of the Commonsense World.* Ablex Publishing, Norwood NJ, 1984.

[7] Leora Morgenstern. Knowledge preconditions for actions and plans. In *Proceedings IJCAI*, pages 867–874, Milano, Italy, 1987.

[8] Stanley J. Rosenschein. Synthesizing information-tracking automata from environment descriptions. In *KR89*, pages 386–393, 1989.

[9] Johan F. A. K. van Bentham. *A Manual of Intensional Logic.* Number 1 in Lecture Notes. CSLI, Stanford, second edition, 1988.

Christine Nakatani

Department of Computer and Information Science
nakatani@linc.cis.upenn.edu

Japanese phrase structure and tree-adjoining grammar

Keywords: Japanese Phrase Structure, Tree-Adjoining Grammar, Agglutinative Morphology

The configurational parameters of Japanese phrase structure remain the focus of much debate for both transformational and non-transformational grammarians. Far from settled are issues such as whether there exists a VP node in Japanese, the properties of case-marking, and the precise structures for numerous complex predicates. My work provides one non-transformational analysis of Japanese phrase structure that focuses squarely on the productive and ubiquitous agglutinative processes of Japanese morphology. This approach has required the integration of theoretical findings arising from the modern Western syntax tradition with observations recorded in the native system of Japanese grammar, and has crucially made use of the mildly context-sensitive expressive power of tree-adjoining grammar (TAG).

The backbone of the Japanese tree-adjoining grammar (JTAG) that I have developed is the system of open-ended affixation for deriving complex predicates. Reanalysis of morpho-syntactic inflectional dependencies described in the native grammar lead to the discovery of an "unbounded" dependency between the predicate root and its inflectional subcategorization (i.e. the conclusive, or *syusi* form). The "unbounded" nature of this dependency motivates the choice of tree-adjoining grammar for the formal description of the Japanese inflectional system, which is viewed as the foundation upon which any adequate analysis of Japanese phrase structure must be built. Formalization of this "unbounded" dependency in lexicalized, unification-based TAG yields a grammatical system with several distinctive properties. In JTAG, we dispense with explicit infinitival clauses and do not require the VP node to represent important hierarchical relations. Also, case-marking within clauses is inherent. Most importantly, agglutination is realized as tree-adjunction of affixes into partially derived sentential trees, enabling open-ended affixation during syntactic derivation.

Presently, I am expanding the syntactic coverage of the grammar and studying its consequences for semantic interpretation (e.g. compositional semantics for complex predicates, interpretation of "empty" categories). The goal of further work is to understand the interaction between formal phrase structural constraints (as expressed in JTAG) and various properties and features of the resultant grammatical system that may be said to derive from the settings of configurational parameters in an agglutinative language such as Japanese.

Michael Niv

Department of Computer and Information Science
niv@linc.cis.upenn.edu

A Computational Model of Syntactic Processing: Disambiguation from Interpretation

Keywords: Syntactic Processing, Garden Path Sentences

Natural language contains ambiguities on many levels, of which my dissertation addresses two: part of speech selection and syntactic structure assignment. A central question is how are people able to cope so effortlessly with the considerable computational task of language understanding? One intriguing hint about this process of syntactic disambiguation is the existence of grammatical sentences such as Bever's example:

The horse raced past the barn fell.

in which ambiguity early in the sentence 'tricks' the reader/hearer into committing to the ultimately incorrect analysis. The overall research strategy then is to collect evidence of situations where the process succeeds and fails, and to construct theories in computational terms. While most extant theories of sentence processing have recognized the role of meaning, most consider purely structural aspects as well. My thesis is that it is solely meaning which determines which grammatical alternative is chosen. It follows that the processor is a very simple device, consisting of a blind all-paths syntactic-rule-applier and a meaning-based controller which performs the disambiguation. Here I consider three aspects of my project: a reexamination of a structural disambiguation strategy, formulation of a parser, and my proposed ambiguity resolution scheme.

One of the most successfully exploited structural disambiguation strategies is Right Association [4]. It states that modifiers prefer to attach as low as possible in the phrase structure. While other structural disambiguation strategies have recently been argued to be artifactual,[1] arising only in a limited set of circumstances, I am aware of no such claims about Right Association. Through an investigation of the Penn Treebank corpus of syntactically annotated newswire text, I demonstrate that this principle is often violated, especially when the modifier in question is "syntactically heavy". It follows that the data adduced in support of this principle can be explained by the same competence mechanism which is responsible for other heaviness related phenomena such as dative shift and heavy-NP shift. The need for the structural disambiguation criterion is eliminated.

Examination of sentences such as the example above indicates that the meaning of a word is integrated into the meanings of the various syntactically defined possibilities immediately after the word is encountered. This condition of timely semantic analysis, along with the desideratum of simplicity in the parser, places certain requirements upon the form of the competence grammar. I adopt and extend Steedman's work on Combinatory Categorial Grammar (CCG, [6]) as a formalism which satisfies these constraints. In particular I propose a parser for CCG which is able to effectively cope with the additional nondeterminism which CCG entails. I investigate certain choices in the formulation of the parsing operation called 'revealing' [5] and the consequences of those choices for the whole system. This

[1] Minimal Attachment [3], [1]; Late Closure Altmann, (forthcoming.)

work is applicable to bottom-up parsers of any grammatical formalism which share CCG's associativity of derivation.

For the central project of the dissertation – a demonstration of how meaning could be used to resolve all syntactic ambiguity – I construct a model of an interpreter which considers certain aspects of meaning: filler-gap relations, reference resolution, thematic relations, and a form of heaviness. The role of reasoning is minimized. Using this model, I explore many specific questions, among them:

- How long is ambiguity maintained before it is resolved?
- How many competing analyses can be maintained at a time?
- What role does similarity play in disambiguation?

I test the resulting model on human performance data available from psycholinguistic research and from other naturally occurring and artificially constructed examples.

References

[1] Gerry Altmann and Mark Steedman. Interactiom with Context during Human Sentence Processing. *Cognition* 30. 1988.

[2] Tom Bever. The Cognitive Basis for Linguistic Structures. In J. R. Hayes ed. *Cognition and the Development of Language.* 1970.

[3] Stephen Crain and Mark Steedman. On not being led up the garden path. In Dowty *et al.* eds. *Natural Language Parsing: Psychological Computational and Theoretical Perspectives.* Cambrige. 1985.

[4] John Kimball. Seven Principles of Surface Structure Parsing. *Cognition 2(1).* 1973.

[5] Remo Pareschi and Mark Steedman. A Lazy Way to Chart Parse with Combinatory Grammars. *Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics.* 1987

[6] Mark Steedman. Grammar, Interpretation and Processing form the Lexicon. in *Lexical Representation and Process.* Marslen-Wilson ed. Cambridge, Massachusetts. MIT Press. 1989

Charles L. Ortiz

Department of Computer and Information Science
clortiz@linc.cis.upenn.edu

The Role of Causality in Explanation

Keywords: Causal Reasoning, Conditionals, Explanation

We regularly draw on causal terms when trying to explain the behavior of systems or the effects of an action, or when we wish to support a prediction. For example, we might say that "by removing the chain we enable the removal of the wheel," or "the lights went out because of a short circuit caused by loose wiring." Despite the conspicuous presence of terms such as "causes", "enables", "prevents" "impede", etc, there have been very few attempts in AI at formalizing our commonsense notions of causality. However it seems reasonable to suggest that, if we routinely call upon these terms in the course of explaining behavior, our representations should, at the very least, be able to draw the sorts of distinctions we intend when using causal terms.

In AI, Shoham has made the only attempt I know of at a definition of a few of these terms [4]. In [2], however, I show that there are some problems with his attempts; in particular with his observation of the duality of enablement and prevention and with the status of negative events that can stand in causal relation.

The work described here is one part of my thesis research in causal reasoning. Two areas I am interested in involve the role that might be played by causal theories in language as used for explanation: (1) the ontological status of negative events in causal explanations (e.g., "Not opening the valve will prevent the tank from emptying" in which reference is made to two "events": a present non-event which influences a fictitious or hypothetical future event), and (2) the evaluation of counterfactuals.

With regards to (1), consider a tank filled with water which has a valve that allows water to escape. On Shoham's account, $\neg opening(Valve)$ prevents the tank from emptying because *opening* enables the emptying. However, it appears that a more sensible explanation of the prevention is in terms of the structure of the mechanism or in terms of a previous action that closed the valve; that is, a re-description of the negative act in positive terms. (This presents a difficulty for Shoham's formulation since it requires backward projection and it is well-known [3] that Shoham's approach produces counter-intuitive results.)

Unfortunately, there are still technical problems that arise when we admit negative events. Suppose we ask for a prediction, "What happens if we don't open the valve?" In the situation calculus this corresponds to proving the formula $holds(f, result(\neg opening(Valve), S))$. If we consider only a simple, idealized situation in which concurrent events are disallowed, then f will stand for a huge disjunction of all the states that could result from all of the actions in our repertoire (with the exception of *opening*); this is equivalent to no knowledge about the effect. In a sense this problem is related to the problem of equating causality with material implication: that is, we can certainly prove that $holds(full(tank), result(sneeze, s))$, but we don't want to say that sneezing caused the tank to stay full. Another problem is that there is no way to block the conclusion that there are possible states resulting from not opening the valve in which the tank is empty. For example, drilling a hole into the tank is true just in case we don't open the valve. One might argue that a way around this difficulty is to claim $\neg open$ really corresponds to "do nothing", in which case we get the

right result. But this doesn't seem quite right: if not smoking prevents you from getting cancer, it certainly is not the case that the "not smoking" is to be equated with a huge wait event that spans, say, ten years' time. Here, "not smoking" carries with it the implication that whatever you did did not conflict with the goals of "not smoking"; i.e, you don't give up smoking in order to become an alcoholic. Linguistically, this seems to be related to: "if you don't win the match then..." which carries with it the presupposition that you will play the match and loose and *not* that you will go to the movies instead.

In [5], Vermazen talks about negative acts such as resistings, simple refrainings, displacement refrainings, and disobedient refrainings. He suggests that negative acts can sometimes be re-described in terms of some positive act and offers a characterization of what constitutes a negative act; this characterization is in terms of an agent's prior pro-attitude to not performing the indicated act. Unfortunately, he has little to say about the role of negative events in causal relations nor their explanatory role.

Causal theories can also play a role in the evaluation of hypothetical and counterfactual statements. In philosophy, the study of counterfactuals has centered on a possible worlds analysis without demonstrating how one could arrive at a similarity measure on possible worlds necessary for such an analysis [1]. Adopting a causal approach might provide a more computationally oriented road to the evaluation of conditionals. It would also raise a number of interesting issues: the causal theory applied at the time point referenced by the counterfactual must abstract away unnecessary details. Similarly, the reconstruction of the situation representing the time point in the past will have to be some sort of "vivid context" of that "type" of situation; otherwise there is a danger of infinite regress: "If I hadn't played the match then I would have acted differently prior to that decision but then something would have made me act differently but then I would have..." Similar issues in localized reasoning also arise in the evaluation of hypotheticals such as "What would happen if you were allowed to turn left on red": only the relevant aspects of the situation must be considered as well as the proper causal theory so that reasonable predictions can be made.

References

[1] Lewis, David, *Counterfactuals*, Harvard University Press, Cambridge, Massachusetts, 1973.

[2] Ortiz, Charles L., *The Role of Causality in AI*, Dissertation Pre-proposal, Department of Computer and Information Science, University of Pennsylvania, February 1992.

[3] Stein, Lynn Andrea and Morgenstern, Leora, *Motivated Action Theory: A Formal Theory of Causal Reasoning*, MIT AI Laboratory, A.I. Memo 1338, December 1991.

[4] Shoham, Y., *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence*, MIT Press, Cambridge, Massachusetts, 1988.

[5] Vermazen, Bruce, Negative Acts. In *Essays on Davidson*, edited by Bruce Vermazen and Merrill Hintikka, Clarendon Press, Oxford, 1985.

Jong Cheol Park

Department of Computer and Information Science
park@linc.cis.upenn.edu

A Unification-Based Semantic Interpretation for Coordinate Constructs

Keywords: Unification, Semantic Interpretation, Coordination

Combinatory Categorial Grammar (CCG) has been offered as a theory of coordination in natural language [8], and it has usually been implemented in languages based on first order unification. Moore [4] however has pointed out that coordination presents problems for first-order unification-based semantic interpretation. I have shown in [6] that it is possible to resolve this problem by compiling into the lexicon the lambda reduction steps that are directly associated with coordination. The idea is to revise the semantics of coordination from the standard Montagovian approach [1] as in (1) to the one in (2), where these are for noun phrase coordination.

(1) λQ.λR.λP.(Q(P) & R(P))
(2) λQ.λR.λP.(Q(λY.exists(X,X=Y&P(X))) & R(λZ.exists(X,X=Z&P(X))))

In the work cited I have described how to apply this suggestion to CCG for the constructs shown below; we believe that it could equally well be applied to any lexicon based grammar formalism. Together with the notion of partial execution [7], this approach completely eliminates all the lambda expressions.

(3) Harry walks and every farmer walks.
(4) A farmer walks and talks.
(5) A farmer and every senator talk.
(6) Harry finds and a woman cooks some mushroom.
(7) Mary gives every dog a bone and some policeman a flower.

Since Jowsey's results [2] and [3] suggest that, in other respects, natural language semantics can be characterized in a first-order logic, the approach we have chosen make it possible to give uniform treatment within a first-order unification paradigm.

As an alternative on the other hand, we could choose to enforce a uniform treatment within second-order unification, using, for example, ideas in [5]. This approach is predicted to yield a more elegant solution not only for the coordinate constructs, but also for other issues. We are currently investigating ways to expand ideas in [9] for this purpose.

References

[1] David R. Dowty, Robert E. Wall and Stanley Peters. *Introduction to Montague Semantics.* D. Reidel Publishing Company, 1981.

[2] Einar Jowsey. *Montague Grammar and First Order Logic.* Edinburgh Working Papers in Cognitive Science: Categorial Grammar, Unification Grammar and Parsing, 1:143–194, 1987.

[3] Einar Jowsey. *Constraining Montague Grammar for Computational Applications.* PhD thesis, University of Edinburgh, 1990.

[4] Robert Moore. Unification-Based Semantic Interpretation. In *Proceedings of the 27th Annual Meeting of the ACL*, 1989.

[5] Gopalan Nadathur and Dale Miller. An Overview of λ-Prolog. In *Proceedings of the Fifth International Logic Programming Conference*, 1988.

[6] Jong C. Park. A Unification-Based Semantic Interpretation for Coordinate Constructs. To appear in *The Proceedings of the 30th Annual Meeting of the ACL*, 1992.

[7] Fernando C.N. Pereira and Stuart M. Shieber. *Prolog and Natural-Language Analysis.* CSLI Lecture Notes Number 10, 1987.

[8] Mark J. Steedman. Gapping as Constituent Coordination. *Linguistics and Philosophy*, 13:207–263, 1990.

[9] Gennaro Chierchia and Raymond Turner. Semantics and Property Theory. *Linguistics and Philosophy*, 111:261–302, 1988.

Scott A. Prevost

Department of Computer and Information Science
prevost@linc.cis.upenn.edu

Intonation in Spoken Language Generation

Keywords: Intonation, Natural Language Generation, CCGs

Many researchers have considered intonational structure to be distinct from surface syntactic structure, thereby complicating the processes of speech recognition and synthesis by requiring interaction between autonomous levels of structure. A theory that relates intonational structure with traditional syntactic structure has the advantage of simplifying the paths from speech recognition to interpretation and from semantic interpretation to speech generation. Steedman postulates that Combinatory Categorial Grammar (CCG) admits such a close coupling of the intonational and syntactic structures of English. Distinct intonational contours within an utterance contrast the theme (what the utterance is about) with the rheme (what the speaker is saying about the theme). Moreover, intonational cues distinguish the foci of the theme and rheme from that which is presupposed.

The goal of this project is to produce a discourse-driven utterance generator, employing Steedman's CCG/Intonation theory as well as speech synthesis techniques. The resulting system will consist of a natural language query system in a limited domain (such as inventory management) which derives appropriate query responses with proper intonational contours. As this project does not address the issues of speech recognition, input to the system is given textually with intonational melodies represented symbolically using Pierrehumbert's notation for pitch accents and boundary tones.

For a given input query, a CCG parser determines the prosodic constituents and semantic content of the question, identifying the open proposition, the theme, the rheme, the theme focus and the rheme focus. The open proposition has the form of an expression in the lambda calculus, whose variables are then instantiated by a simple Prolog database query.

By instantiating the variable in the open proposition, a Prolog utterance generator produces a natural language response to the question, employing the notions of theme and rheme to generate the appropriate intonational melody. Specifically, the theme of the question becomes the rheme of the response. Current work is aimed at appropriately distributing the theme and rheme melodies across the corresponding constituents in the generated response, with attention to the proper placement of the theme and rheme foci.

Currently the generator's output is represented as lists of words and pitch markings in Prolog. In future stages of the research I anticipate developing an interface between the response generator and speech synthesis systems, thereby producing spoken responses with appropriate intonational contours.

References

[1] Mark Steedman. Structure and Intonation. *Language*, 67:260–296, 1991.

[2] Mark Steedman. *Surface Structure, Intonation and Meaning in Spoken Language.* Technical Report MS-CIS-91-12, University of Pennsylvania, 1991.

Ellen F. Prince

Professor, Department of Linguistics
ellen@central.cis.upenn.edu

Discourse/pragmatics

Keywords: Discourse Functions of Syntax, Reference, Language contact, Yiddish

I am interested in that part of linguistic competence that underlies the use of particular linguistic forms in particular contexts, where the choice is not entailed by sentence-grammar or truth-conditional meaning. In particular, I am interested in the choice of referential expressions and syntactic constructions. I am also interested in the effects of language contact on this domain. The bulk of my research has focused on English and Yiddish.

References

[1] Ellen F. Prince. Fancy syntax and 'shared knowledge'. *Journal of Pragmatics* 1985.

[2] Ellen F. Prince. On the syntactic marking of presupposed open propositions. In Farley, Farley,, and McCullough, eds. *Papers from the Parasession on Pragmatics and Grammatical Theory, 22nd Regional Meeting, Chicago Linguistic Society.* 1986.

[3] Ellen F. Prince. Discourse analysis: a part of the study of linguistic competence. In Newmeyer, F., ed. *Linguistics: the Cambridge Survey.* Vol. 2. Linguistic theory: extensions and implications. NY/Cambridge: Cambridge Univ. Press. Pp. 164-82. 1988.

[4] Ellen F. Prince. The discourse functions of Yiddish Expletive *es* + Subject-Postposing. *Papers in Pragmatics* 2.176-94, 1988.

[5] Ellen F. Prince. On pragmatic change: the borrowing of discourse functions. In Kasher, A., ed. *Cognitive aspects of language use.* Amsterdam: Elsevier Pub. B.V. 1989.

[6] Wllen F. Prince. Syntax and discourse: a look at resumptive pronouns. In Reinman, S. et al., eds. *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics Society. Parasession on the Legacy of Grice.* 1990.

[7] Ellen F. Prince and Gregory Ward. On the topicalization of indefinite NPs. *Journal of Pragmatics,* 15.8.167-78, 1990.

[8] Ellen F. Prince. On syntax in discourse, in language contact situations. In Kramsch, C. and McConnell-Ginet, S., eds. *Text and context: cross-disciplinary perspectives on language study.* Boston: D.C. Heath. 1992

[9] Ellen F. Prince. On the discourse functions of syntactic form in Yiddish. In Goldberg, D., Herzog, M., Kirshenblatt-Gimblett, B., and Miron, D. eds. The Field of Yiddish. 5th collection. Studies in Yiddish language, folklore, and literature. Northwestern University Press/YIVO. Pp.59-86. In press.

[10] Ellen F. Prince. The ZPG letter: subjects, definiteness, and information-status. In Thompson, S. and Mann, W., eds. *Discourse description: diverse analyses of a fund raising text.* Philadelphia/Amsterdam: John Benjamins B.V. In press.

Owen Rambow

Department of Computer and Information Science
rambow@linc.cis.upenn.edu

Formal Aspects of Word Order Variation in Natural Language

Keywords: Natural Language Syntax, Formal Languages, Processing

I am studying the formal properties of word order variation in "free word order" languages. Most of my data comes from German. In German, there are two different mechanisms for changing the order of the arguments of a verb: topicalization and scrambling, illustrated below:

(1) [Das Buch]$_i$ hat der Lehrer [PRO meiner Mutter t$_i$ zu geben] versprochen.

(2) Ich glaube, daß [das Buch]$_i$ der Lehrer [PRO meiner Mutter t$_i$ zu geben] versprochen hat.

(E) (I think that) the teacher has promised to give the book to my mother.

In the topicalization case (1), the moved accusative NP *das Buch* – 'the book' – has moved into sentence-initial position, presumably into the SPEC position of CP. In the case of scrambling (2), it has not moved beyond the complementizer *daß* – 'that'. The two movements are different: in the case of topicalization, there is exactly one possible landing site, and exactly one constituent can move into it; in the case of scrambling, the scrambled NP can move to any position, and more than one NP may move simultaneously, so that all orders of the verbal arguments (and adjuncts) are possible. In joint work with Tilman Becker (Universität des Saarlandes, Germany) and Aravind Joshi [1], I have investigated the formal implication of this freedom. It is known that topicalization can be handled by Tree Adjoining Grammars [6]. However, it turns out that Linear Contex-Free Rewriting Systems [8] are not powerful enough to generate all relevant word orders [2]. Instead, one has to resort to more powerful TAG variants such as non-local MC-TAGs or FO-TAGs. With Giorgio Satta, I am currently exploring properties of formalisms that exhibit the requisite non-locality.

Interestingly, the asymmetry in the required power of the underlying formal systems correlates with an asymmetry in the linguistic facts. While much of the early discussion of scrambling in the GB literature attempted to show that scrambling and topicalization are in fact the same sort of syntactic process (namely, A'-movement), more recent evidence from binding facts has shown an asymmetry. In joint work with Bob Frank and Young-Suk Lee, I have been exploring the effect of scrambling in weak and strong cross-over configurations [3]; it appears that while scrambling generally patterns with A-movement, the subject has a special status with respect to binding facts. An asymmetry between topicalization and scrambling can also be found in coordination structures. These can be explained by referring to the different formalisms used to represent the two movement types.

Two further areas are relevant to the issue of topicalization and scrambling. First, pragmatic factors are crucial since both scrambling and topicalization are subject to discourse constraints. I am currently studying the matter within Centering theory [4]. Secondly, in embedded contexts, processing heavily affects the grammaticality judgments for certain scrambled word orders, so that a processing account is needed that can correctly predict

which word orders lead to degraded performance. I have been exploring this question in contexts in which scrambling interacts with extraposition (the so-called "Third Position").

Text Planning and Knowledge

Keywords: Text Generation, Text Planning

In a separate vein of interest, I have been looking into the types of knowledge needed for planning multi-paragraph texts during the text generation process. Recent approaches to text planning have stressed the importance of rhetoric (the first such approach being McKeown's TEXT [7]). However, it appears that the task of relating rhetorical goals to domain knowledge is difficult and needs a type of knowledge all of its own, *domain communication knowledge* [5]. In typical domain representations of existing knowledge-based systems, this knowledge may not be explicitly represented, since communication is not always a design objective. On the other hand, humans typically have problems distinguishing domain knowledge from domain communication knowledge, since domain knowledge is usually acquired through communication. The relationship between domain communication knowledge, domain knowledge and communication knowledge (such as rhetoric) remains to be investigated further, from both the theoretical and practical points of view.

References

[1] Tilman Becker, Aravind Joshi, and Owen Rambow. Long distance scrambling and tree adjoining grammars. In *Proceedings of the 5th Conference of the European Chapter*, ACL, 1991.

[2] Tilman Becker and Owen Rambow. Scrambling is beyond LCFRS. 1992. Unpublished Draft, University of Pennsylvania.

[3] Robert Frank, Young-Suk Lee, and Owen Rambow. Scrambling as non-operator movement and the special status of subjects. In *Proceedings of the Third Leiden Conference for Junior Linguists*, 1992.

[4] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Towards a computational theory of discourse interpretation. 1986. Draft.

[5] Richard Kittredge, Tanya Korelsky, and Owen Rambow. On the need for domain communication knowledge. *Computational Intelligence*, 7(4), 1991.

[6] A. Kroch. Subjacency in a tree adjoining grammar. In A. Manaster-Ramer, editor, *Mathematics of Language*, John Benjamins, Amsterdam, 1987.

[7] Kathleen McKeown. *Text Generation.* Cambridge University Press, Cambridge, 1985.

[8] D. J. Weir. *Context-Free Grammars to Tree Adjoining Grammars and Beyond.* Technical Report, Department of Computer and Information Science, University of Pennsylvania, Philadelphia, 1987.

Philip Resnik

Department of Computer and Information Science
resnik@linc.cis.upenn.edu

Combining Knowledge-based Resources with Statistical Discovery

Keywords: Statistical Methods in Natural Language, Lexical Acquisition

It has become common in statistical approaches to natural language to use measures of lexical association, such as the information-theoretic measure of mutual information, to extract useful relationships between words (e.g. [1, 2, 3]). Applications of lexical statistics include the discovery of typical collocations, the disambiguation of word senses, the determination of parsing preferences, and many others.

I have been developing a generalization of lexical association techniques that is intended to facilitate statistical discovery of facts involving word *classes* rather than individual words. Many of the applications of lexical statistics mentioned above would be improved by access to class-based information — for example, David Yarowsky and colleagues at Bell Laboratories have recently demonstrated the advantages of using class-based information for word-sense disambiguation. In addition, some areas of linguistic investigation would seem to *require* a move toward class-based rather than lexically-based statistics; one example is the acquisition of the selectional restrictions (and preferred arguments) of verbs.

Although the most straightforward approach to defining class-based association measures is to treat word classes simply as sets of words, direct use of such a definition is impractical because there are simply too many classes to consider. Rather than considering all possible classes, in my investigations I have been structuring the set of possible word classes by using a broad-coverage lexical/conceptual hierarchy [4]. Such a hybrid approach combines many of the advantages of "knowledge-based" methods with those of "knowledge-free" statistical methods: the statistical techniques are well suited to discovering linguistic regularities on the basis of large quantities of data, and capturing the details of typical usage; knowledge-based techniques provide aspects of linguistic organization that may not be easily recoverable statistically, and can be organized according to principles and generalizations that seem appropriate to the researcher regardless of what evidence there is in any particular corpus. Perhaps most important, neither the knowledge-base nor the corpus need be perfect, since the knowledge-base helps narrow the range of statistical search and the weight of large numbers helps compensate for errors or idiosyncrasies in a hand-built taxonomy.

Initial results are encouraging. Using the corpus resources of the Penn Treebank, I have applied the hybrid technique to the discovery of "prototypical" argument classes for verbs, and the following small selection is representative of the results:

Association Score	verb	object class
1.94	ask	⟨question,[question,...]⟩
0.16	call	⟨someone,[person,...]⟩
2.39	climb	⟨stair,[step,...]⟩
3.64	cook	⟨repast,[repast,...]⟩
0.27	draw	⟨cord,[cord,...]⟩
3.58	drink	⟨beverage,[beverage,...]⟩
1.76	eat	⟨nutrient,[food,...]⟩
0.30	lose	⟨sensory_faculty,[sense,...]⟩
1.28	play	⟨part,[character,...]⟩
2.48	pour	⟨liquid,[liquid,...]⟩
1.03	pull	⟨cover,[covering,...]⟩
1.23	push	⟨button,[button,...]⟩
1.18	read	⟨written_material,[writing,...]⟩
2.69	sing	⟨music,[music,...]⟩

Work in progress includes the application of this technique to models of lexical acquisition, and in particular the investigation of the different ways in which verbs can syntactically realize their arguments.

References

[1] Kenneth Church and Patrick Hanks. Word association norms, mutual information, and lexicography. In *Proceedings of the 23rd Annual Meeting of the ACL*, 1989.

[2] K. Church, W. Gale, P. Hanks, and D. Hindle. Parsing, word associations and typical predicate-argument relations. In *Proceedings of the International Workshop on Parsing Technologies*, August 1989.

[3] Donald Hindle. Noun classification from predicate-argument structures. In *Proceedings of the 28th Annual Meeting of the ACL*, 1990.

[4] George Miller, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. Five papers on WordNet. CSL Report 43, Cognitive Science Laboratory, Princeton University, July 1990.

[5] David Yarowsky. A Class-based Statistical Approach to Word-Sense Disambiguation. Presentation to CLiFF group, University of Pennsylvania. February 1992.

Robert Rubinoff

Department of Computer and Information Science
rubinoff@linc.cis.upenn.edu

Integrating Text Planning and Linguistic Choice by Annotating Linguistic Structures

Keywords: Natural Language Generation, Knowledge Representation

This work focuses on the relationship between text planning and linguistic choice in natural language generation. In particular, I am concerned with handling cases where linguistic and planning issues interact while still preserving the separation of the planning and linguistic processes. This is accomplished by having the linguistic component of the generator provide feedback to the planner in the form of annotations that describe the effects and consequences of particular linguistic decisions. These annotations abstract away from the details of linguistic structure, providing the planner with just the information it needs to evaluate the various linguistic options suggested by the linguistic component. This allows the two components to coordinate their decisions without the planner having to understand linguistic structures or the linguistic component having to understand plan structures or goals. I have implemented these ideas in the **IGEN** generator, which can produce texts in which linguistic decisions depend on the goals, preconditions, and structure of the text plan even though the linguistic component has no access to the plan.

One consequence of this approach to natural language generation is that the generator must explicitly consider issues of time pressure and resource limitations. The text planner and linguistic component incrementally refine and modify the text, constantly seeking to improve its quality and eliminate problems. Since this process can continue for an arbitrary amount of time, and indeed is not guaranteed to ever terminate, the generator must explicitly model both the quality of the current version of the text and the time pressure to produce some text. I am investigating handling this requirement by means of a separate component – which I call the "utterer" – that balances these factors. Some initial experiments that vary the time pressure or the available linguistic resources have shown how the **IGEN**'s output varies as the utterer responds to the changed constraints.

A secondary focus of my research is on how particular linguistic choices can be sensitive to and affect the speaker's (and hearer's) perspective on the things being talked about. The notion of perspective includes both the speaker's specific purposes in talking about something as well as the general body of beliefs and attitudes that the speaker has about the topics of discussion. For example, the choice between describing a set of beliefs as "myth" or as "religion" may depend both on whether the speaker intends to disparage or praise the beliefs and on the speaker's general attitudes. I have modeled this notion of perspective within a semantic network and inference rule framework by attaching to every network element and rule an associated perspective weight that indicates its prominence to reasoning. The perspective weights allow the representation to model the varying perspectives that the generator can take on the objects it talks about: varying the prominence of a concept modifies whether and how quickly the generator sees it, and varying the prominence of relations modifies which descriptions the generator considers appropriate to describe particular objects. In addition to the weights themselves, network elements and rules can have associated

perspective shifts that indicate how their use in reasoning can alter perspective weights for them or related elements; this allows a generator to consider how use of particular words or phrases can alter the hearer's perspective when making its choices.

Giorgio Satta

Postdoc, Institute for Research in Cognitive Science
gsatta@linc.cis.upenn.edu

A Computational Investigation of the Notion of Locality

Keywords: Rewriting Systems, Parsing

In current development of modern theories of grammar, the notion of locality plays a fundamental role, one of the goals being to find a restrictive device powerful enough to account for so-called long distance dependencies. Under a formal language theoretic perspective, it is interesting to investigate the status of locality, its generative limitations and its computational consequences, exploiting the formal setting of abstract rewriting systems that are general enough to represent the structural dependencies found in natural language. I am currently involved in this research.

The common interest within the computational linguistics field for rewriting systems that enlarge the generative power of context-free grammars (CFG) has led to the definition of a class of languages, called mildly context-sensitive (MCS), that remain far below the full power of context-sensitive languages (see [1] for discussion). The rewriting systems found in MCS enlarge, up to some extent, the domain of locality of context-free grammars by generalizing the concatenation operation in different ways. As a result, the generative power of these systems is increased, but we still observe efficient parsing properties, in fact polynomial (deterministic) time parsability. The rather surprising fact, that many of these systems have been shown to be weakly equivalent, has led researchers to generalize the elementary operations involved in only apparently different formalisms, in order to capture the underlying similarities.

Such an attempt is found in [6], where the class of linear context-free rewriting systems (LCFRS) is defined; for each system in the class, rewriting observes some bound on the domain of locality. Nonetheless, since concatenation is generalized in such a way that we can define local "crossing dependencies" among phrases, the parsing problem is no longer guaranteed to have efficient solutions, as shown in [4]. This result reveals an undesired dissimilarity between well-known formalisms such as TAG, HG, LIG and the subclass of LCFRS that is intended to generalize these formalisms.

The definition of an even more powerful class of rewriting systems, called nonlocal multi-component tree adjoining grammars, completely drops the locality restriction. Interestingly enough, we observe a further enlargement of the generative power, but at the cost of a more demanding computation: even if we fix the grammar in advance, there is evidence for the non-existence of polynomial parsing algorithms, as discussed in [3].

BiDirectional Context-Free Grammar Parsing for Natural Language Processing

Keywords: Context-free Grammar Recognition/Parsing, Covering Grammars

In recent years a number of natural language parsing algorithms have been proposed that adopt bidirectional strategies, i.e., do not analyze the input string in a strictly left-to-right

fashion. In fact, devices that are capable of processing input sentences in a bidirectional manner are very attractive: many arguments have been presented in the standard literature in favor of this claim. The idea of lexicalization, to begin with, has played an important role in favor of bidirectional parsing for obvious reasons. Furthermore, automatic speech recognition/understanding, hand-written input parsing and, more generally, all cases in which the input string may be corrupted, can gain some benefit from bidirectional strategies.

With the prospects outlined above, a systematization for bidirectional tabular parsing of context-free languages has been proposed in [5]. The work starts by reviewing well-known left-to-right tabular methods as based on "easy to process" grammars that cover the general form input grammars, following the approach proposed in [2]. Various classes of "bidirectional" covering grammars are then defined in order to study the general problem from a theoretical perspective: in the general case we have found that, as compared with unidirectional parsing, bidirectional parsing requires considerable additional computational effort. This important fact has already been observed from an intuitive point of view in the natural language parsing literature; the covering framework allows us to give a more precise account of the problem.

Our result has also been the starting point for the study of a standard technique which allows us to improve, in an average case perspective, the parsing performance of bidirectional tabular parsers. Such a technique has been exploited in the development of two parsing algorithms. The first algorithm is a head-driven bidirectional parser which works under the hypothesis that the input grammar comes with a specification of head elements found within each production. We have also studied a general algorithm for island-driven parsing: the algorithm permits analyses to start from any dynamically chosen positions within the input sentence, and combines, without restrictions, bottom-to-top and top-down processing. In both cases, the key problem of avoiding analysis redundancy has been solved in its full generality by exploiting our technique. As already mentioned, the proposed algorithms have been developed with the aim of yielding results to be used in natural language processing. The head-driven algorithm can successfully be employed in parsing of "mixed" languages where heads of different categories subcategorize for complements in different directions. The island-driven algorithm can be primarily used for speech or hand-written input parsing, or in those cases in which the usual left-to-right methods seem inadequate, as in text skimming.

References

[1] Aravind K. Joshi. The convergence of mildly context-sensitive grammatical formalisms. Presented at the Workshop on *Processing of Linguistic Structure*, Santa Cruz, 1987.

[2] R. Leermakers. How to cover a grammar. In *Proc. of the 27 th ACL*, pages 135–142, Vancouver, British Columbia, Canada, 1989.

[3] O. Rambow and G. Satta. Locality in rewriting systems. In preparation, 1992.

[4] G. Satta. Recognition of linear context-free rewriting systems. In 30th *Meeting of the Association for Computational Linguistics (ACL'92)*, 1992.

[5] G. Satta and O. Stock. Bidirectional context-free grammar parsing for natural language processing. Submitted, January 1992.

[6] K. Vijay-Shanker, David J. Weir, and Aravind K. Joshi. Characterizing structural descriptions produced by various grammatical formalisms. In 25th *Meeting of the Association for Computational Linguistics (ACL'87)*, 1987.

Jeffrey Mark Siskind

Postdoc, Institute for Research in Cognitive Science
qobi@unagi.cis.upenn.edu

Recovering Event Descriptions from Animated Movies

Keywords: Event Perception, Naive Physics, Lexical Semantics

When people observe the world they can generally determine whether certain events have happened. Furthermore, they can describe those events using language. For instance, after seeing John throw a ball to Mary, the observer can say that the event described by the utterance *John threw the ball to Mary* has happened, perhaps along with events described by other utterances. I am investigating mechanisms for explaining how such event perception may work.

My approach is motivated by recent experimental studies of adult visual perception and infant knowledge of object permanence (cf. Freyd et al. 1988, Baillargeon 1987 and Spelke 1988). In formulating this approach I advance three claims about event perception and the process of grounding language in visual perception. First, I claim that the notions of support, contact, and attachment play a central role in defining the meanings of simple spatial motion verbs in a way that delineates prototypical occurrences of events described by those verbs from non-occurrences. For example, *throwing* involves moving one's hand while grasping an object (attachment), resulting in the unsupported motion of that object. Prior approaches to lexical semantic representation (e.g. Miller 1972, Schank 1973, Jackendoff 1983, and Pinker 1989) focussed primarily on movement and lacked the ability to incorporate the crucial notions of support, contact, and attachment into the definitions of simple spatial motion verbs. Second, I claim that support, contact, and attachment relations between objects are recovered from images by a process of counterfactual simulation. For instance, an object A supports another object B if B does not fall when the short-term future of the image is predicted, but does fall if A is removed. Such counterfactual simulations are performed by a modular imagination capacity. Third, I claim that this imagination capacity, while superficially similar in intent to traditional kinematic simulation (cf. Cremer 1989 and Kramer 1990), is actually based on a drastically different foundation. This foundation takes the process of enforcing naive physical constraints such as substantiality, continuity, and attachment relations between objects to be primary. In doing so it sacrifices physical accuracy and coverage. This is in contrast to the traditional approach which achieves physical accuracy and coverage by numerical integration, relegating the maintenance of constraints to a process of secondary importance built around the numerical integration core.

The mechanisms which I have proposed as part of this work have been partially implemented in a computer program called ABIGAIL (Siskind 1992). ABIGAIL watches a computer-generated animated stick-figure movie and constructs descriptions of the objects and events that occur in that movie. The input to ABIGAIL consists solely of the positions, orientations, shapes, and sizes of the line segments and circles which constitute the image at each frame during the movie. From this input, ABIGAIL segments the image into objects, each object comprised of several line segments and circles, and delineates the events in which those objects participate. ABIGAIL's event perception processes rely on counterfactual simulation to recover changing support, contact, and attachment relations between objects in the movie. This is in contrast to prior approaches to this task (e.g. Badler 1975)

which were based solely on determining the spatial relations between objects in the image sequence, grounding verb meanings in static geometric predicates used to compute those spatial relations without counterfactual analysis.

Computational Models of Child Language Acquisition

Keywords: Language Bootstrapping

As part of the process of acquiring their native language, children must learn at least three things: the syntactic categories of words, their meanings, and the language-specific components of syntax. Such knowledge constitutes, at least in part, the language-specific linguistic knowledge which children must acquire to become fluent speakers of their native language. Initially, children lack any such language-specific knowledge. Yet they come to acquire that knowledge through the language acquisition process. My work attempts to answer the following question: *What procedure might children employ to simultaneously learn word-to-category mappings, word-to-meaning mappings, and the language-specific components of syntax, without any access to previously acquired language-specific knowledge?*

Prior work in this area often assumes an ordering in the acquisition of the different kinds of language-specific information. For instance, the semantic bootstrapping hypothesis (Grimshaw 1981 and Pinker 1984) claims that the child first learns word-to-meaning mapping without the aid of syntactic information. These word-to-meaning mappings can then be used to derive word-to-category mappings and the syntax of the language being learned. On the other hand, the syntactic bootstrapping hypothesis (Gleitman 1990, Fisher et al. unpublished) claims that children use syntactic information available from prosody to help determine word meanings. In my work I explore a third possibility: that constraint satisfaction techniques can be used to simultaneously acquire all language-specific information with no assumptions about which information is acquired first. Such techniques allow the learner to acquire partial knowledge from ambiguous situations and combine this partial knowledge across situations to infer a unique language model despite the ambiguity in the individual isolated situations.

I have implemented a series of computer programs which test this constraint-satisfaction-based learning strategy on linguistic theories of successively greater sophistication. In accord with current hypotheses about child language acquisition, these systems use only positive examples to drive their acquisition of a language model. MAIMRA (Siskind 1992), the first program developed, learns word-to-meaning and word-to-category mappings from a corpus pairing utterances with sets of expressions representing the potential meanings of those utterances hypothesized by the learner from the non-linguistic context. MAIMRA's syntactic theory is embodied in a fixed context-free grammar. DAVRA (Siskind 1992), the second program developed, extends MAIMRA by replacing the context-free grammar with a parameterized variant of $\overline{\text{X}}$ theory. Given the same corpus as MAIMRA, DAVRA learns the parameter settings for $\overline{\text{X}}$ theory in addition to a lexicon mapping words to their syntactic category and meaning. DAVRA has been successfully applied, without change, to tiny corpora in both English and Japanese, learning the requisite lexica and parameter settings despite differences in word order between the two languages. KENUNIA (Siskind 1992), the third program developed, incorporates a more comprehensive model of universal grammar supporting movement, adjunction, and empty categories, as well as more extensive parameterization of its $\overline{\text{X}}$ theory component. This model of universal grammar is based on recent linguistic theory and includes such notions as the DP hypothesis, VP-internal subjects, and

V-to-I movement. KENUNIA is able to learn the parameter settings of this model, as well as word-to-category mappings, in the presence of movement and empty categories.

References

[1] Renée Baillargeon. Object permanence in $3\frac{1}{2}$- and $4\frac{1}{2}$-month-old infants. *Developmental Psychology*, 23(5):655–664, 1987.

[2] James F. Cremer. *An Architecture for General Purpose Physical System Simulation—Integrating Geometry, Dynamics, and Control.* PhD thesis, Cornell University, April 1989. available as TR 89–987.

[3] Cynthia Fisher, Geoffry Hall, Susan Rakowitz, and Lila Gleitman. When it is better to receive than to give: syntactic and conceptual constraints on vocabulary growth. Unpublished manuscript received directly from author, 1991.

[4] Jennifer J. Freyd, Teresa M. Pantzer, and Jeannette L. Cheng. Representing statics as forces in equilibrium. *Journal of Experimental Psychology, General*, 117(4):395–407, December 1988.

[5] Lila Gleitman. The structural sources of verb meanings. *Language Acquisition*, 1(1):3–55, 1990.

[6] Jane Grimshaw. Form, function, and the language acquisition device. In C. L. Baker and J. J. McCarthy, editors, *The logical problem of language acquisition.* The M. I. T. Press, Cambridge, MA and London, England, 1981.

[7] Ray Jackendoff. *Semantics and Cognition.* The M. I. T. Press, Cambridge, MA and London, England, 1983.

[8] Glenn Andrew Kramer. Geometric reasoning in the kinematic analysis of mechanisms. Technical Report TR–91–02, Schlumberger Laboratory for Computer Science, October 1990.

[9] George A. Miller. English verbs of motion: A case study in semantics and lexical memory. In Arthur W. Melton and Edwin Martin, editors, *Coding Processes in Human Memory*, chapter 14, pages 335–372. V. H. Winston and Sons, Inc., Washington, DC, 1972.

[10] Steven Pinker. *Language Learnability and Language Development.* Harvard University Press, Cambridge MA, 1984.

[11] Steven Pinker. *Learnability and Cognition.* The M. I. T. Press, Cambridge, MA and London, England, 1989.

[12] Roger C. Schank. The fourteen primitive actions and their inferences. Memo AIM-183, Stanford Artificial Intelligence Laboratory, March 1973.

[13] Jeffrey Mark Siskind. *Naive Physics, Event Perception, Lexical Semantics, and Language Acquisition.* PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, January 1992.

[14] Elizabeth S. Spelke. The origins of physical knowledge. In L. Weiskrantz, editor, *Thought without Language*, chapter 7, pages 168–184. Clarendon Press, 1988.

Mark Steedman

Associate Professor, Department of Computer and Information Science
steedman@linc.cis.upenn.edu

Combinators and Grammars for Natural Language Understanding

Keywords: Computational Linguistics, Syntax and Semantics, Speech, Combinatory Logic, Cognitive Science

My research interests cover a range of issues in the areas of computational linguistics, artificial intelligence, computer science and cognitive science, including syntax and semantics of natural languages and programming languages, parsing and comprehension of natural language discourse by humans and by machine, natural language generation, and intonation in spoken discourse. I also work on formal models of musical comprehension.

Most of my research since completing my graduate work has been on two problems in computational linguistics. The first concerns a theory of natural language syntax and its relation to "incremental" syntactic and semantic processing of spoken and written language. The research demonstrates a direct relation between certain problematic natural language constructions and certain purely local, variable-free, combinatory operations on functions, such as functional composition. The constructions in question involve unbounded dependencies between syntactic elements, such as those found in relative clauses and in coordinate constructions. The combinatory operations are related to some of the simplest combinators which have been used to provide a foundation for applicative systems such as the lambda calculus and the related programming languages. The research addresses a number of questions of practical importance. The weaknesses of most current theories of grammar in the face of the full range of coordination phenomena means that existing computational grammars have the characteristics of unstructured programs – that is, they are non-modular and hard to modify, placing practical limitations on the size and portability of the systems that include them. The standard theories show a similarly bad fit to a number of other phenomena of practical importance, notably phrasal prosody and intonation. Most of my current work is in this latter area, in particular in the problem of synthesising contextually appropriate intonation in limited conversational domains.

My second principal research interest concerns a computationally-based semantics for tense and temporal reference, and exploits the advantages of computational models for capturing phenomena which are presupposition-laden and involve interactions with non-sentence-internal knowledge. The work shows that the primitives involved in this domain are not solely (or even primarily) temporal, but rather are concerned with "contingent" relations between events, such as causation. This project also addresses a practical concern, for any database that is to be interrogated or updated in natural language making use of tense and related categories is certain to require structuring in the same way. A number of domains are under investigation, including certain problems in the graphical animation of action sequences.

References:

(1987) Combinatory Grammars and Parasitic Gaps. *Natural Language and Linguistic Theory*, 5, 403-439.

(1988a) Temporal Ontology and Temporal Reference. (with Marc Moens), *Journal of Computational Linguistics*, 14, 15-28.

(1988b) Interaction with Context During Human Sentence Processing. (with Gerry Altmann), *Cognition*, 30, 191-238

(1990a) Gapping and Constituent Coordination. *Linguistics and Philosophy*, 13, 207-264.

(1991a) "Structure and Intonation", *Language*, 68.2,260-296.

Marilyn A. Walker

Department of Computer and Information Science
lyn@linc.cis.upenn.edu

1.1 Japanese Discourse and the Process of Centering

Keywords: discourse, centering, zero pronouns, japanese

This research has two aims: (1) to generalize CENTERING theory, a computational account of discourse processing[12, 7, 6, 3, 4, 13], and apply it to discourse processing in Japanese, and (2) to provide some insights on the effect of syntactic factors in Japanese on discourse interpretation.

We argue that while discourse interpretation is an inferential process, syntactic cues constrain this process[11], and demonstrate this argument with respect to the interpretation of ZEROS, unexpressed arguments of the verb, in Japanese. The syntactic cues in Japanese discourse that we investigate are the morphological markers for grammatical TOPIC, the post-position *wa*, as well as those for grammatical functions such as SUBJECT, *ga*, OBJECT, *o* and OBJECT2, *ni*. In addition, we investigate the role of speaker's EMPATHY, which is the perspective from which an event is described. This is morphologically indicated through the use of verbal compounding, i.e. the auxiliary use of verbs such as *kureta, kita*. We have also investigated the interaction of speaker's EMPATHY with explicit TOPIC marking, with having previously been the CENTER and with other discourse factors such as INDEFINITENESS[18].

Our results are based on a survey of more than 30 native speaker's interpretations of short discourses, consisting of minimal pairs, varied by one of the above factors. We demonstrate that these syntactic cues do indeed affect the interpretation of ZEROS, but that having previously been the TOPIC and being realized as a ZERO also contribute to an entity being interpreted as the TOPIC. We propose a new notion of TOPIC AMBIGUITY, and show that CENTERING provides constraints on when a ZERO can be interpreted as the TOPIC[15, 16].

1.2 Redundancy in Dialogue and Reasoning about Plans

Keywords: dialogue, mixed initiative, collaborative plans

I am investigating of the use of LOGICALLY REDUNDANT utterances (LRUs) in dialogue, in order to elucidate their role in establishing the beliefs necessary for mutual understanding and joint action[14]. Discourse participants begin a dialogue with different beliefs and agreement as to the current state of the world, what goals are to be achieved, what constraints must be met, and how goals can be achieved is often established over the course of the conversation[19, 5, 1].

LRU's include cases where conversants REPEAT or PARAPHRASE their own or another conversant's previous utterances. They also include utterances that make an INFERABLE that follows from previous utterances explicit. This account is based on an analysis of two types of naturally occurring advisory dialogues in which two agents are constructing a plan through the dialogue. One set are from a radio talk show called *Speaking of your Money* on WCAU in Philadelphia[1]. The other set of dialogues result from a client phoning an

[1]This corpus was collected and transcribed by Martha Pollack and Julia Hirschberg[10].

expert to help them diagnose and repair various software faults[2]. My hypotheses about the functions of logical redundancy include:

- LRU's function to limit the processing required by resource-bounded agents. Agents may reason about whether other agents will be able to make certain inferences or retrieve certain facts from mutual beliefs. Making inferables explicit ensures that the inferable fact is part of the current context, and with low processing cost for other conversants. Similarly, explicitly stating a proposition that was added to the mutual beliefs in a prior segment of the dialogue, but which is not in the current segment, saves search/retrieval time for other conversants. In addition, LRU's, such as summaries, mark segment boundaries[19, 17]. Segmentation in discourse, also known as GLOBAL FOCUSING[2], is an important example of modularization of relevant knowledge with its concomitant payoff in reduced processing. Summaries serve to coordinate subdialogue transitions by functioning as an agreement between two agents to close a segment, and therefore contribute to agents' achievement of reduced processing through focusing.

- LRU's may be a product of an agent's own resource bounds. In dialogue, timely action is crucial, since a delay in a response actually carries meaning in and of itself. A simple response strategy that allows more time for inferences and maintains discourse coherence is to paraphrase or repeat what someone else said (with the appropriate intonation). In addition, agents may want to ensure understanding or agreement before a dialogue moves on; misunderstandings result in revising beliefs and belief revision is costly. Paraphrasing and making inferences explicit decrease the likelihood of misunderstanding.

This analysis includes an investigation into the intonational marking of redundancy in the radio talk show dialogues. Intonation is one of the main ways that the information status of discourse entities is indicated. Preliminary results include the fact that it is common for given(as believed) information to receive a pitch accent, but that the type of pitch accent may be qualitatively different than that for brand-new information[9]. Furthermore informationally redundant utterances are realized with sustained levels[8] or downstepping tones in cases where the proposition realized by the utterance is discourse salient.

Future work includes the development of an explicit computational model that demonstrates the function of these utterances in constraining processing.

References

[1] Julia R. Galliers. Cooperative interaction as strategic belief revision. In M.S. Deen, editor, *Cooperating Knowledge Based Systems*, page 1. Springer Verlag, 1991.

[2] Barbara J. Grosz. The representation and use of focus in dialogue understanding. Technical Report 151, SRI International, 333 Ravenswood Ave, Menlo Park, Ca. 94025, 1977.

[3] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Providing a unified account of definite noun phrases in discourse. In *Proc. 21st Annual Meeting of the ACL, Association of Computational Linguistics*, pages 44–50, Cambridge, MA, 1983.

[2]These dialogues were taped at one of Hewlett-Packard's customer response centers[19, 17].

[4] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Towards a computational theory of discourse interpretation. Unpublished Manuscript, 1986.

[5] Barbara J. Grosz and Candace L. Sidner. Plans for discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press, Cambridge, MA.*, 1990.

[6] Aravind K. Joshi and Steve Kuhn. Centered logic: The role of entity centered sentence representation in natural language inferencing. In *Proc. International Joint Conference on Artificial Intelligence*, page pp., 1979.

[7] Aravind K. Joshi and Scott Weinstein. Control of inference: Role of some aspects of discourse structure - centering. In *Proc. International Joint Conference on Artificial Intelligence*, pages pp. 385–387, 1981.

[8] Cynthia A. McLemore. *The pragmatic Interpretation of English Intonation: Sorority Speech.* PhD thesis, University of Texas, Austin, 1991.

[9] Janet Pierrehumbert and Julia Hirschberg. The meaning of intonational contours in the interpretation of discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press, Cambridge, MA.*, 1990.

[10] Martha Pollack, Julia Hirschberg, and Bonnie Webber. User participation in the reasoning process of expert systems. In *Proc. National Conference on Artificial Intelligence*, 1982.

[11] Ellen F. Prince. Fancy syntax and shared knowledge. *Journal of Pragmatics*, pages pp. 65–81, 1985.

[12] Candace L. Sidner. Toward a computational theory of definite anaphora comprehension in English. Technical Report AI-TR-537, MIT, 1979.

[13] Marilyn A. Walker. Evaluating discourse processing algorithms. In *Proc. 27th Annual Meeting of the Association of Computational Linguistics*, pages 251–261, 1989.

[14] Marilyn A. Walker. Redundancy in collaborative dialogue. In *AAAI Fall Symposium on Discourse Structure*, 1991.

[15] Marilyn A. Walker, Masayo Iida, and Sharon Cote. Centering in japanese discourse. In *COLING90: Proc. 13th International Conference on Computational Linguistics, Helsinki*, page 1, 1990.

[16] Marilyn A. Walker, Masayo Iida, and Sharon Cote. Japanese discourse and the process of centering. *In submission*, 1992.

[17] Marilyn A. Walker and Steve Whittaker. Mixed initiative in dialogue: An investigation into discourse segmentation. In *Proc. 28th Annual Meeting of the ACL*, pages 70–79, 1990.

[18] Bonnie Lynn Webber. *A Formal Approach to Discourse Anaphora.* PhD thesis, Harvard University, 1978. Garland Press.

[19] Steve Whittaker and Phil Stenton. Cues and control in expert client dialogues. In *Proc. 26th Annual Meeting of the ACL, Association of Computational Linguistics*, pages 123–130, 1988.

Bonnie Lynn Webber

Professor, Department of Computer and Information Science
bonnie@central.cis.upenn.edu

Animation from NL Instructions

Keywords: Instructions, Planning, Animation

Along with Professors Norm Badler and Mark Steedman, I have been co-supervising members of the Animation and Natural Language group (AnimNL) on research aimed at creating *animated task simulations* automatically from *natural-language instructions*. It is work that has great practical value for high-level control of human-figure animation [2, 3]. But it is work that has also led to new insights about natural-language understanding and the relationship between that understanding and behavior. For the research being done by members of the group, please see statements by **Baldwin**, **Di Eugenio**, **Geib**, **Jung**, **Levison**, **Moore**, and **White**. Here I will present an overview.

First, the type of instructions we are concerned with are those that an agent would have for carrying out a *complicated maintenance or assembly procedure* for the first time or infrequently. Such instructions are quite different in form and intent from instructions given as advice to an agent already engaged in a task, when the agent is at an impasse or decision point in what to do [1, 5, 8]. Advisory instructions are usually brief, specifying a way of carrying out a goal that the agent has already adopted. Maintenance and assembly instructions are much longer, describing a partially-ordered sequence of goals, means for achieving them and constraints on those means. While the issues involved in understanding and acting upon these various types of instructions overlap, longer instructions emphasize the role of understanding and reasoning *prior* to the agent's commitment to action. Since we view an agent as continually committing to further action (even while it is acting), we take the agent to be continually reassessing and elaborating its understanding of already given instructions, in response to its actions.

(For our initial prototype, we have decided to restrict ourselves to single-agent procedures, even though Penn's Graphics Laboratory is able to simulate and animate environments with more than one agent acting together. The reason is that we have discovered a sufficient number of new language, planning and plan inference problems that require solutions, even without considering the communication actions that multi-agent planning and coordination require.)

Instructions are given to the AnimNL system in *steps* consisting of one or more utterances. A step specifies a continuous behavior that the agent must attend to. Following a step, the agent can again direct its attention to the instructor and the next step in the procedure. (By breaking instructions into steps *a priori*, the system is relieved of the problem of deciding how much of the instructions to understand before simulating the agent acting in accordance with them.)

Steps are processed by a parser that uses a combinatory categorial grammar (CCG) [9] and produces an action representation based in part on Jackendoff's *Conceptual Structures* [7]. This representation captures general features shared by whole classes of verbs and the inferences that follow from them. We retain however the actual lexical items, since they contribute to identifying the particular positive and negative intentions that an agent is meant to adopt. The representation of each subsequent instruction step is then used

to elaborate a structure that represents the animated agent's *beliefs* and *intentions* and the relationships between them. (The structure is called a *plan graph* and is described in greater detail in [6].) Beliefs in the plan graph include *inter alia* beliefs about the consequences of actions and about entities discovered or changed as a consequence of those actions. Intentions in the plan graph include positive intentions to satisfy a given goal or act in a given way at a given time (or in a given situation), and negative intentions (to not act in a particular way during a particular situation or while satisfying one or more goals).

The processes that are being developed to elaborate the plan graph in response to the output of the parser and internal changes in the plan graph itself include:

Situation Visualization - makes existential assumptions explicit, specifying entities assumed to be present prior to, during and following actions, as well as changes to those entities as a consequence of actions. (The structure of entities and situations constitutes the system's discourse model.) See **Baldwin**.

Modal Reasoning - reasons about relationships between action and knowledge. Used to verify existential assumptions about changes in the agent's knowledge that may be brought about by planned actions and to assist in planning to find things whose location is not known. See **Moore**.

Goal–Action Elaboration - links goal specifications and action specifications and, in doing so, may further specify those actions and/or augment the plan graph with additional intentions. See **Di Eugenio**.

Sub-goal Planning - reasons about how to satisfy goals, using limited look-ahead to allow future intentions to affect decisions of how current goals should be satisfied. This view mirrors that of Goal–Action Elaboration, where current goals are seen to inform decisions about how current actions are realized. See **Geib**.

Object-specific Planning - reasons about how to carry out a particular physical action on a particular object, based on geometric and functional properties of the object. It is this process that recognizes, for example, that the "open" action in "open the shoe box" and "open the cereal box" should be elaborated to different behavior specifications (grasping the top of the box and lifting vs. undoing the top flaps). Object-specific planning follows referent identification, since an object's relevant aspects are not always conveyed linguistically ("open that"). See **Levison**.

(Other processes will be developed and incorporated in future versions of the system as the need for them becomes evident.)

When an intention becomes sufficiently specified for the agent to be ready to commit to it and temporal dependencies permit such commitment, the intention is gated, triggering another, low-level planning process (called *posture planning*, see **Jung**). The output of this process may either be an indication that the agent's "body" is unable to carry out the behavior that its "mind" would like it to or a collection of *behaviors* to be executed (simulated) in parallel.[1] Actions change the world, as well as the agent's knowledge. Such changes trigger further elaboration of the agent's intentional structure and further commitments to action. Thus AnimNL's architecture supports the integration of goal-directed and situated activity that is characteristic of current approaches to rational agency [4].

[1] Previous actions need not be completed before a new action is committed to: an agent can be (and usually is) doing more than one thing at a time.

References

[1] Alterman, R. and Zito-Wolf, R. Planning and Understanding: Revisited. In J. Hendler (ed.), *Planning in Uncertain, Unpredictable, or Changing Environments.* Technical Report 90-45, Systems Research Center, University of Maryland.

[2] Badler, N. and Webber, B. Task Communication through Natural Language and Graphics: Workshop Report. *AI Magazine.* To appear Winter 1990.

[3] Badler, N., Webber, B., Kalita, J. and Esakov, J. Animation from Instructions. In N. Badler, B. Barsky and D. Zeltzer (eds.), *Making Them Move: Mechanics, Control and Animation of Articulated Figures*, Los Altos CA: Morgan Kaufmann Publishers, 1990, pp.51-93.

[4] Bratman, M., Israel, D., and Pollack, M. Plans and Resource-bounded Practical Reasoning. *Computational Intelligence* 4(4), November 1988, pp.349-355.

[5] Chapman, D. Vision, Instruction and Action. Cambridge MA: MIT Press, 1991.

[6] Di Eugenio, B. and White, M. On the Interpretation of Natural Language Instructions. Submitted to *1992 Int. Conf. on Computational Linguistics* (COLING-92), Nantes, France, July 1992.

[7] Jackendoff, R. *Semantic Structures.* Cambridge MA: MIT Press, 1990.

[8] Martin, C. and Firby, R.J., An Integrated Architecture for Planning and Learning. *Sigart Bulletin* 2(4), August 1991, p.125-129.

[9] Steedman, M. Gapping as Constituent Coordination. *Linguistics and Philosophy* 13, 1990, pp. 207-263.

Michael White

Department of Computer and Information Science
mwhite@linc.cis.upenn.edu

Yet Another Study of Tense and Aspect

Keywords: Aspect, Lexical Semantics, Tense, Perspective

Perhaps nowhere in the study of language is it more evident that everything is related to everything else than in the study of tense and aspect. My research on aspect directly reflects this, as my goal is to provide an integrated treatment of the wide range of aspect-related phenomena within the rubric of Jackendoff's Conceptual Structures. My research on tense, in contrast, is more limited in scope, as my goal is to examine in detail the little studied phenomenon of shifted temporal perspective in narrative.

Aspect and Conceptual Structures

Research in lexical and conceptual semantics has been regarded with much deserved skepticism. Despite this healthy pessimism, however, interest in lexical semantics has been renewed. Central to this resurgence has been the assumption that subtle distinctions in word meaning correspond to otherwise puzzling syntactic variations, leading to the hope of a principled basis for identifying lexical semantic representations (cf. [5]).

The syntax of these semantic representations has been investigated by Jackendoff [3], Levin and Rappaport [9], and Pinker [7], who have all argued that the neo-Davidsonian notion of theta roles or thematic relations (cf. [2]) cannot account for observable connections between syntax and meaning. In light of these arguments, these researchers have concluded that structured semantic representations should be used instead. However, as Dowty and Zwarts and Verkuyl [11] have pointed out, the *semantics* of these semantic representations is in dire need of study.

I intend to study the semantics of Jackendoff's Conceptual Structures [3, 4] within a computational framework, drawing upon the formal treatment of Jackendoff's representations found in Zwarts and Verkuyl [11] and the modifications suggested by Pinker [7]. My primary objective is to demonstrate that Conceptual Structures provide an adequate basis for an integrated treatment of the wide range of aspect-related phenomena, including lexical aspect, incremental theme arguments (cf. [2]), aspectual verbs (eg. *start*), auxiliaries, temporal modifiers, and aspectual type coercion (cf. [6]).

Tense and Perspective in Narrative

The role of perspective in narrative has received comparatively little study within computational paradigms. Recent work by Caenepeel [1], however, has shown the importance of perspective for the well-known problem of determining temporal ordering in discourse, especially in the case of narrative. I intend to extend her work by examining in detail the phenomenon of shifted temporal perspective, found for example in temporally sequenced pluperfects. My aim is to demonstrate that these cases are instances of the general phenomenon of indirect discourse, and furthermore that they necessitate extensions to Reichenbachian [8] frameworks for the discourse processing of tense (cf. [10]).

References

[1] Mimo Caenepeel. *Aspect, Temporal Ordering and Perspective in Narrative Fiction.* PhD thesis, University of Edinburgh, 1989.

[2] David Dowty. Thematic proto-roles and argument selection. *Language*, 67(3):547–615, 1991.

[3] Ray Jackendoff. *Semantic Structures.* MIT Press, 1990.

[4] Ray Jackendoff. Parts and boundaries. *Cognition*, 41:9–45, 1991.

[5] Beth Levin and Steven Pinker. Introduction to special issue of *cognition* on lexical and conceptual semantics. *Cognition*, 41:1–7, 1991.

[6] Marc Moens and Mark Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, June 1988.

[7] Steven Pinker. *Learnability and Cognition: The Acquisition of Argument Structure.* MIT Press, 1989.

[8] Hans Reichenbach. *Elements of Symbolic Logic.* Macmillan, New York, 1947.

[9] Malka Rappaport and Beth Levin. What to do with theta-roles. In Wendy Wilkins, editor, *Thematic Relations*, volume 21 of *Syntax and Semantics*, pages 7–36. Academic Press, Inc., 1988.

[10] Bonnie Lynn Webber. Tense as discourse anaphor. *Computational Linguistics*, June 1988.

[11] Joost Zwarts and Henk Verkuyl. An algebra of conceptual structure; an investigation into Jackendoff's conceptual semantics. To appear in *Linguistics and Philosophy*, 1991.

Appendix

A CLiFF Talks

Spring 1991

Feb 1, 1991 David Magerman CS – Stanford
Pearl: A Probabilistic Chart Parser

Mar 4, 1991 Massimo Poesi CS – Rochester
Towards An Inferential Characterization Of Reference
And Scope Disambiguation

Mar 18, 1991 Breck Baldwin CIS – UPenn
On Definite Reference

Mar 25, 1991 Michael Niv CIS – UPenn
On Parsing Without Strategies

Apr 1, 1991 Owen Rambow CIS – UPenn
Long-Distance Scrambling and TAGs

Apr 8 , 1991 Dan Hardt CIS – UPenn
VP Ellipsis

Apr 15, 1991 Megan Moser LING – UPenn
Incremental discourse semantics for sentential negation

Apr 22, 1991 Young-Suk Lee LING – UPenn
Scrambling and the Adjoined Argument Hypothesis

Apr 29, 1991 Yves Schabes CIS – UPenn
Polynomial time and space shift-reduce parsing
of arbitrary context-free grammars

May 13, 1991 Eric Brill CIS – UPenn
Zellig Harris meets the Brown Corpus

May 20, 1991 Owen Rambow CIS – UPenn
Lexical Semantics for Lexicalized Grammars:
Meaning-Text linguistics and the linguistic relevance of TAGs

Summer & Fall 1991

June 10, 1991 Barbara Di Eugenio CIS – UPenn
Action Representation for Natural Language Instructions

June 10, 1991 Christine Nakatani CIS – UPenn
Resolving a pragmatic prepositional phrase attachment ambiguity

Sept 17, 1991 Dr. Sangal Indian Institute of Technology
Parsing strategies in the Paninian grammatical theory

Sept 27, 1991 Dan Hardt CIS – UPenn
A Discourse Model Approach to VP Ellipsis

Oct 3, 1991 Jamie Henderson CIS – UPenn
A Model of Real Time Syntactic Parsing in Bounded Memory

Oct 10, 1991 Michael Niv CIS – UPenn
A discourse-based explanation of apparent Right Association effects

Oct 24, 1991 Mark Steedman CIS – UPenn
A Theory of "Phonological Form"

Oct 31, 1991 Philip Resnik CIS – UPenn
Left-corner parsing and psychological plausibility

Nov 7, 1991 Marilyn Walker CIS – UPenn
Redundancy in Collaborative Dialogue

Dec 5, 1991 Eric Brill CIS – UPenn
Some Statistical Properties of Natural Language
and their Application to Cryptology and Part-of-Speech Tagging

Dec 11, 1991 Michel DeGraff CIS – UPenn
The Syntax of Predication in Haitian

B LINC Lab Technical Reports, 1991

Investigating A Proof-Theoretic Meta-Language For Functional Programs (Dissertation)
John Hannan
MS-CIS-91-09
LINC LAB 191

Type-Raising and Directionality In Combinatory Grammar
Mark Steedman
MS-CIS-91-11
LINC LAB 192

Surface Structure, Intonation and Meaning In Spoken Language
Mark Steedman
MS-CIS-91-12
LINC LAB 193

A Simple, Yet Probabilistically Tractable Algorithm For First Principles Diagnosis
Ron Rymon
MS-CIS-91-13
LINC LAB 194

Common Knowledge: A Survey
Marilyn A. Walker
MS-CIS-91-14
LINC LAB 195

Dynamic Binding Communication Mechanism
Jeffrey S. Aaronson
MS-CIS-91-16
LINC LAB 196

Tree-Adjoining Grammars and Lexicalized Grammars
Aravind K. Joshi
Yves Schabes
MS-CIS-91-22
LINC LAB 197

Unification Of Simply Typed Lambda-Terms As Logic Programming
Dale Miller
MS-CIS-91-24
LINC LAB 198

Unification-Based Tree Adjoining Grammars
K. Vijay-Shanker
(University of Delaware)
Aravind K. Joshi
(University of Pennsylvania)
MS-CIS-91-25
LINC LAB 199

Action Composition For The Animation Of Natural Language Instructions
Libby Levison
MS-CIS-91-28
GRAPHICS LAB 40
LINC LAB 200

Combining A Type Hierarchy With A Rule-Based Reasoner
Lokendra Shastri
D.R. Mani
MS-CIS-91-33
LINC LAB 201

Generation and Synchronous Tree-Adjoining Grammars
Stuart M. Sheiber
(Harvard University)
Yves Schabes
(University of Pennsylvania)
MS-CIS-91-42
LINC LAB 202

Synchronous Tree-Adjoining Grammars
Stuart M. Sheiber
(Harvard University)
Yves Schabes
(University of Pennsylvania)
MS-CIS-91-43
LINC LAB 203

Using Lexicalized Tags For Machine Translation
Anne Abeillé
(University of Paris)
Yves Schabes
(University of Pennsylvania)
Aravind K. Joshi
(University of Pennsylvania)
MS-CIS-91-44
LINC LAB 204

Flexible Support For Trauma Management Through Goal-Directed Reasoning and Planning
Bonnie L. Webber
Ron Rymon
John R. Clarke
MS-CIS-91-54
LINC LAB 205

Surface Structure, Intonation, and "Focus"
Mark Steedman
MS-CIS-91-63
LINC LAB 206

Computational Accounts of Music Understanding
Daniel Hardt
MS-CIS-91-66
LINC LAB 207

Towards Goal-Directed Diagnosis (Preliminary Report)
Ron Rymon
(University of Pennsylvania)
Bonnie L. Webber
(University of Pennsylvania)
John R. Clarke
(Medical College of Pennsylvania)
MS-CIS-91-67
LINC LAB 208

Abstract Syntax and Logic Programming
Dale Miller
MS-CIS-91-72
LINC LAB 209

Unification Under A Mixed Prefix
Dale Miller
MS-CIS-91-81
LINC LAB 210

E-kernel On The IBM Victor V256 Multiprocessor–An Experimental Platform For Parallel Systems
Dennis G. Shea
MS-CIS-91-101
LINC LAB 211