

# The FINITE STRING Newsletter

## Announcing a New Section

Many aspects of computational linguistics research and development are extremely labor-intensive. These include realistic grammars of a given natural language, collections of semantic interpretation rules and case frames, corpora, lexicons, and software tools, to name just a few.

One of the ways **The FINITE STRING** can help is in publishing catalogs of such resources. The following submission by Gerald Gazdar on sources of parsers within the paradigm of generalized phrase structure grammars inaugurates this section, called **SOURCES**. Note that we are interested in publishing catalogs – not descriptions of a single example – nor descriptions of items associated with a single site.

This section will only be as good as the items submitted to it. Potentially it can contain some of the most valuable information available. We welcome and encourage your submissions.

## Sources

### Recent Computer Implementations of Phrase Structure Grammars

**Gerald Gazdar**  
Center for Advanced Study in the Behavioral Sciences  
202 Junipero Serra Boulevard  
Stanford, CA 94305, USA

[Address after 1 September 1985:

The University of Sussex  
School of Social Sciences  
Arts Building  
Falmer, Brighton BN1 9QN England ]

This note lists twelve recent implementations of phrase structure grammars. Each entry begins with the name and address of the persons or institution primarily responsible. This is followed by summary details of the systems, together with information about the programming language(s) they were written in and the machines they run on. The entries conclude with a list of any publications or working papers that have resulted from the projects described. Many of the systems presuppose the grammatical formalism for phrase structure grammars outlined in Gazdar and Pullum (1982), and the terminology used in the entries below reflects that fact.

Gazdar, Gerald and Pullum, Geoffrey K. 1982 Generalized Phrase Structure Grammar: A Theoretical Synopsis. University Linguistics Club, Bloomington, Indiana. Also distributed as University of Sussex Cognitive Science Research Paper 7 (CSRP 007).

**John Bear**  
Linguistics Research Center  
P.O. Box 7247, University Station  
University of Texas  
Austin, TX 78712 USA

An all paths, left corner chart parser that uses features for agreement and unbounded dependencies. It handles relative clauses, questions, and existentials. No semantics is implemented. The parser is written in Interlisp and runs on a DEC20.

Bear, John 1981 Gaps as Syntactic Features. MA dissertation. University of Texas at Austin. [Published by IULC, Bloomington, Indiana, in 1982.]

Bear, John and Karttunen, Lauri 1979 PSG: A Simple Phrase Structure Parser. *Texas Linguistic Forum* 15: 1-46.

**Roger Evans**  
Cognitive Studies Programme  
University of Sussex  
Brighton BN1 9QN, UK

ProGram is a grammar development system intended for use by linguists and computational linguists developing grammars for significant fragments of natural languages. It incorporates all aspects of the 1982 version of the Generalized Phrase Structure Grammar (GPSG) framework: features, metarules, immediate dominance and linear precedence (ID/LP) rules, feature instantiation, coordination, etc. The parser uses ID/LP format directly and does not create the induced set of PS rules. ProGram allows the user to 'direct' the parsing if he wishes. Designed for portability, the system is written in Prolog and runs on VAXes under VMS or UNIX<sup>1</sup>, and on other UNIX systems.

<sup>1</sup>UNIX is a trademark of Bell Laboratories.

Evans, Rogert and Gazdar, Gerald 1984 The ProGram Manual. University of Sussex Cognitive Science Research Paper 35 (CSRP 035).

Keller, Bill 1984a Generating Logic from ProGram Parse Trees. University of Sussex Cognitive Science Research Paper 39 (CSRP 039).

Keller, Bill 1984b A Lexicon Handler for the ProGram Grammar Development System. University of Sussex Cognitive Science Research Paper 40 (CSRP 040).

**Mark Johnson**  
Department of Linguistics  
University of California at San Diego  
La Jolla, CA 92093 USA

A suite of small programs including a sentence generator employing features, instantiation, and ID/LP; a feature package defining unification, increment, etc.; and an

LR(1) parser. The programs are written in FranzLisp and run on a VAX under UNIX.

---

**James Kilbury**  
**Technische Universität Berlin**  
**Fachbereich Informatik (20)**  
**Institut für Angewandte Informatik**  
**Projekt KIT, Sekr. FR 5-8**  
**Franklinstrasse 28/29**  
**D-1000 Berlin 10, West Germany**

A modified Earley-Shieber parser using a "first" relation for the ID/LP formalism is currently being extended to cover all aspects of the GPSG framework including direct parsing with metarules. The system is implemented in Waterloo Prolog and runs on an ITEL AS/5-7031 (equivalent to IBM 370).

Kilbury, James 1984a A Modification of the Earley-Shieber Algorithm for Direct Parsing of ID/LP Grammars. Unpublished paper, Technische Universität Berlin.

Kilbury, James 1984b GPSG-based Parsing and Generation. To appear in: Rollinger, Claus-Rainer, Ed., *Probleme des (Text)-Verstehens - Ansätze der Künstlichen Intelligenz*. Max Niemeyer, Tübingen.

---

**Francis Jeffrey Pelletier**  
**Department of Philosophy**  
**University of Alberta**  
**Edmonton, Canada T6G 2H1**

A recursive descent parser that incorporates metarules and slash categories and provides intensional logic translations. The parser does not incorporate features or the ID/LP format for rules. The system is written in SNOBOL and runs on an Amdahl 470.

---

**Stephen G. Pulman**  
**Computer Laboratory**  
**University of Cambridge**  
**Corn Exchange Street**  
**Cambridge CB2 3QG, UK**

A recursive transition network (RTN) based parser operating either depth or breadth first. The system compiles metarules and phrase structure rules into an RTN and then optimises it. Slash categories are included, but not other featural information. A minimal semantics is associated with one test grammar. The implementation is written in PDP11 and runs on a VAX under the VMS version of the POPLOG system.

Pulman, Stephen 1983 Generalised Phrase Structure Grammar, Earley's Algorithm, and the Minimisation of Recursion. In: Sparck-Jones, K. and Wilks, Y, Eds., *Automatic Natural Language Parsing*. Ellis Horwood, Chichester, England: 117-131.

Pulman, Stephen 1984 Limited Domain Systems for Language Teaching. *Proceedings of Coling84*.

**Lenhart K. Schubert**  
**Department of Computing Science**  
**University of Alberta**  
**Edmonton, Canada T6G 2H1**

A left corner parser that prunes syntactically or semantically unusual alternatives. It incorporates features and morphological analysis, coordination and slash categories, but not metarules. Semantically, it provides first order logic translations. The system is intended as a front end for a question-answering system with access to a logic-based semantic net. Lisp and Pascal versions exist, both of which run on an Amdahl 470.

Pelletier, Francis Jeffrey and Schubert, Lenhart 1984 Two Theories for Computing the Logical Form of Mass Expressions. *Proceedings of Coling84*. Stanford, California: 108-111.

Schubert, Lenhart 1982 An Approach to the Syntax and Semantics of Affixes in 'Conventionalized' Phrase Structure Grammar. *Proceeding of the 4th Biennial Conference of the Canadian Society for Computational Studies of Intelligence*: 189-195.

Schubert, Lenhart 1984 On Parsing Preferences. *Proceedings of Coling84*. Stanford, California: 247-250.

Schubert, Lenhart and Pelletier, Francis Jeffrey 1982 From English to Logic: Context-Free Computation of 'Conventional' Logic Translation. *American Journal of Computational Linguistics* 8(1): 27-44.

---

**Hidetoshi Shirai**  
**Department of Mathematical Engineering**  
**and Instrumentation Physics**  
**Faculty of Engineering**  
**University of Tokyo**  
**Hongo 7-1-2, Bunkyo-ku**  
**Tokyo 113, JAPAN**

A deterministic parser based on PARSIFAL that incorporates metarules, raising constructions, and unbounded dependencies. A Montague semantics is associated with the parse trees. The system is written in Lisp on a Hitac M200H.

Shirai, Hidetoshi 1983 Deterministic Parser. In *Proceedings of the Workshop on Non-transformational Grammars*. ICOT, Tokyo: 57-61.

---

**SRI International**  
**[Stuart Shieber and Stan Rosenschein]**  
**333 Ravenswood Avenue**  
**Menlo Park, CA 94035 USA**

The PATR-I system employs a CKY parser and the feature system allows Boolean combinations of feature equalities to be interpreted on the fly. There are no metarules. The semantics is converted to first-order logic and passed to a theorem prover. The system is written in Interlisp and runs on a DEC20.

Rosenschein, Stanley and Shieber, Stuart M. 1982 Translating English into Logical Form. *Proceedings of the 20th Annual Meeting of the Association for Computational Linguistics*: 1-8.

**SRI International**  
**[Stuart Shieber et al.]**  
**333 Ravenswood Avenue**  
**Menlo Park, CA 94035 USA**

The PATR-II system exists in three implementations, two in Lisp (Interlisp on a DEC20 and Zetalsip on a Symbolics 3600) and one in Prolog (on a DEC20). The Lisp versions use a CKY parser whilst the Prolog version uses Earley's algorithm. The feature system uses directed acyclic graph structures and the semantics is embedded in that feature system. Morphological analysis is achieved by the method of Kimmo Koskenniemi (implemented in Lisp by Lauri Karttunen).

- Karttunen, Lauri 1984 Features and Values. *Proceedings of Coling84*. Stanford, California: 28-33.
- Koskenniemi, Kimmo 1983 A Two Level Model for Morphological Analysis. *Proceedings of the 8th International Joint Conference on Artificial Intelligence*: 683-685.
- Pereira, Fernando and Shieber, Stuart 1984 The Semantics of Grammar Formalisms Seen as Computer Languages. *Proceedings of Coling84*. Stanford, California: 123-129.
- Shieber, Stuart 1983 Sentence Disambiguation by a Shift-reduce Parsing Technique. Technical Note 281. SRI International, Menlo Park, California. Also in *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*: 113-118, and in *Proceedings of the 8th International Joint Conference on Artificial Intelligence*: 699-703.
- Shieber, Stuart 1984a Direct Parsing of ID/LP Grammars. *Linguistics and Philosophy* 7: 135-154.
- Shieber, Stuart 1984b The Design of a Computer Language for Linguistic Information. *Proceedings of Coling84*. Stanford, California: 362-366.
- Shiebert, Stuart; Stucky, Susan; Uszkoreit, Hans; and Robinson, Jane 1983 Formal Constraints on Metarules. Technical Note 283. SRI

International, Menlo Park, California. Also in *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*: 22-27.

- Stucky, Susan 1983 Metarules as Meta-Node-Admissibility Conditions. Technical Note 304. SRI International, Menlo Park, California.

---

**Henry Thompson and John Phillips**  
**Department of Artificial Intelligence**  
**University of Edinburgh**  
**Hope Park Square**  
**Edinburgh EH8 9NW, UK**

MCHART is a chart parser intended for grammar testing. It incorporates all aspects of the 1982 GPSG framework: features, metarules, feature instantiation, coordination, etc. The semantics is currently being implemented. The system is written in Lisp and runs on a DEC10 and a VAX under UNIX.

- Thompson, Henry 1981 Chart Parsing and Rule Schemata in PSG. *Proceedings of the 19th Annual Meeting of the Association for Computational Linguistics*: 167-172.
- Thompson, Henry 1982 Handling Metarules in a Parser for GPSG. Edinburgh D.A.I. Research Paper No. 175. Also, in: Barlow, M.; Flickinger, D.; and Sag, I.A., Eds., *Developments in Generalized Phrase Structure Grammar: Stanford Working Papers in Grammatical Theory, Volume 2*. Indiana University Linguistics Club, Bloomington, Indiana: 26-37.
- Thompson, Henry 1983 Crossed Serial Dependencies: A Low-Power Parseable Extension to GPSG. *Proceedings of the 21st Annual Meeting of the Association for Computational Linguistics*: 16-21.
- Thompson, Henry and Phillips, John 1984 An Implementation of GPSG within the MCHART Parsing Framework. Unpublished paper, Department of Artificial Intelligence, University of Edinburgh.