

## BOOK REVIEWS

### AFFIX POSITIONS AND COOCCURRENCES: THE PARADIGM PROGRAM

(Summer Institute of Linguistics Publications in Linguistics  
No. 69)

Joseph E. Grimes

Summer Institute of Linguistics  
Dallas, Texas, 1983  
Pp. vii, 90.

The program Grimes describes in this little book represents both a significant departure from, and contribution to, the "traditional" approach to computer-assisted morphological analysis. The "traditional" approach, exemplified in programs such as those described in Koskeniemi (1983) and Wolfart & Pardo (1979), tests a user-supplied morphological analysis against a set of word-forms. The PARADIGM program is concerned with a different aspect of morphological analysis. The user supplies a list of segmented word-forms, with morphemes separated by blanks or dashes (the output from either of the previously mentioned programs could be used), and the program produces output from which the user can deduce both relative affix positions (the "slots" the morphemes occupy) and various types of cooccurrence restrictions.

Unlike the "stemming" of "lemmatization" programs, the PARADIGM program does not *test* or *apply* a pre-existent morphological analysis, but rather *performs* some limited analytic functions. The program proceeds through three distinct stages of analysis:

- The first stage produces tables of relative affix positions, from which the user can easily construct an affix slot graph.
- The second stage produces tables of "distinct affix sets", i.e., lists of affixes that do not cooccur in the set of word-forms supplied by the user. These lists include both affixes that occupy the same slot (and therefore obviously cannot cooccur) and affixes that do not cooccur for a variety of other possible reasons, including hapology, semantic restrictions, or incomplete paradigms having been supplied for analysis. It is left to the user to determine what the significance of these sets may be.
- The third and final stage of computer analysis produces **component subgraphs** of affix cooccurrences, which describe the patterns of affix combination observed for each of the affixes in the test data. From these component subgraphs the user constructs a **complete cooccurrence graph**, a task that requires some art, skill,

and practice, but is well worth the effort. Assuming that the input data are complete and accurate, the resulting graph defines all valid affix combinations.

For the most part, Grimes's description of the program, his instructions for its use, and his directions for interpreting the output are admirably complete, concise, and intelligible to the non-specialist. Two exceptions should be noted:

- 1) the section on constructing a complete cooccurrence graph from the component subgraphs is perhaps overly terse, and
- 2) a further step in analysis, in which the complete cooccurrence graph is related to a semantic analysis of cooccurrence restrictions, is mentioned but not discussed at all.

In both cases the prospective user is referred to other sources for the missing information, but the value of this book would have been greatly enhanced by a slight expansion on these two topics.

Perhaps the most serious shortcoming of the PARADIGM program is that it requires extensive "normalization" of the input data. There is no provision for allomorphic variation, so that all affixes must be represented as invariant morphemes. Homographs are not tolerated, so the user must "subscript" or otherwise distinguish between homographic affixes in the input data. Finally, if variant affix orderings are found, each ordering must be submitted for analysis in a separate run of the program.

The PARADIGM program is of interest primarily to the field linguist working on a language with extensive affixation (Grimes gives examples of Central and South American languages with up to 15 prefix slots and 28 suffix slots). Manual analysis of affix positions and cooccurrence restrictions in such languages is an intimidating, time-consuming, and error-prone chore. A modest investment of time in familiarizing oneself with the PARADIGM program should result in a substantial long-term time saving, plus a more reliable analysis.

*R.L. Brandson*

Linguistics Programme  
University of Manitoba  
Winnipeg, Manitoba, Canada

### References

- Koskeniemi, Kimmo 1983 Two-Level Morphology: A General Computational Model for Word-Form Recognition and Production. Publications of the Department of General Linguistics, University of Helsinki, No. 11.
- Wolfart, H. Christoph and Pardo, Francis 1979 Computer-Aided Philology and Algorithmic Linguistics: A Case Study. *IJAL* 45: 107-122.