is the constant of the pumping lemma, the pumped substring cannot be longer than \( k \), and therefore the only other place we might be able to pump would be in the middle as and the middle cs. But this would result in violating the condition that \( l \) may not be greater than \( i \). Thus, \( z \) cannot be pumped without violating the pumping lemma, and hence (7) is not context-free. Since context-free languages are closed under intersection with regular sets, it follows that (6) is not context-free either. Since context-free languages are also closed under substitution, this means that (5) is also not context-free. Finally, since (5) is the intersection of English, with the regular set (4), it follows that English, is not context-free. Q.E.D.

**Alexis Manaster-Ramer**  
Computer Science Department  
Wayne State University  
Detroit, MI 48202

## REFERENCES

Kac, M.B.; Manaster-Ramer, A.; and Rounds, W.C. 1987 Simultaneous-Distributive Coordination and Context-Freeness. *Computational Linguistics* 13(1-2): 25-30.

Langendoen, D.T. 1977 On the Inadequacy of Type-3 and Type-2 Grammars for Human Languages. In Hopper, P.J., Ed., *Studies in Descriptive and Historical Linguistics: Festschrift for Winfred P. Lehman*. John Benjamins, Amsterdam: 159-172.

Manaster-Ramer, A. 1983 The Soft Formal Underbelly of Theoretical Syntax. *CLS* 19: 256-262.

Manaster-Ramer, A. In press. Dutch as a Formal Language. *Linguistics and Philosophy* 10.

Pullum, G. and Gazdar, G. 1982 Context-Free Languages and Natural Languages. *Linguistics and Philosophy* 4: 471-504.

## NOTES

1. Pullum and Gazdar (1982) state that they “can tolerate” examples like (2), and Langendoen (personal communication) agrees.
2. In other terms, we must be able to tell which verb would agree with which subject given the chance, and disallow just those combinations where the result would be a marked singular verb.
3. Ignoring, for the sake of simplicity, the arguments advanced in Manaster-Ramer (1983; in press) about the need to state formal results about natural language in other than weak generative capacity terms.
4. Ignoring, for the sake of readability, the punctuation that would normally be required in written English and the suprasegmental features that would occur in the spoken language.
5. In the discussions of formal properties of natural languages, substitutions have not figured at all prominently, whereas homomorphisms, which are just a special case of substitutions, have. It may be helpful, therefore, to point out that a substitution is a mapping like a homomorphism except that it is usually denoted by \( \sigma \) rather than \( h \) and that it may associate each element in the vocabulary of a language with a whole set (possibly infinite) of strings rather than with just one string, as in the case of a homomorphism. In the present case, we needed to employ a (slightly more general kind of) substitution in order to be able to associate the women with sing as well as danced. It should also be noted that, while the man and the women are linguistically analyzable, we have for technical convenience treated them as single elements of the terminal vocabulary in defining the substitutions.
6. \( \lambda \) denotes the empty string.

**A NOTE ON A STUDY OF CASES**

This note describes and illustrates a study of deep cases using a large sample of sentences. The purpose of the note is to draw attention to the value of the source material used for those interested in case-based representations of sentence meaning, and to indicate the potential utility of the study results.

The purpose of this note is to draw attention to the utility of a specific source of data relevant to the use of case-based meaning representations in language processing, by illustrating the way we have exploited this source.

Like many others, we have used a language analyser that builds meaning representations expressing semantic case roles; specifically, Boguraev's (1979) analyser builds dependency trees with word senses defined by semantic category primitive formulae, and with case labels, i.e., semantic relation primitives, on the constituents of verb (and some other) structures.

Using the analyser for more varied and demanding purposes than Boguraev's original tests (see, e.g., Boguraev and Sparck Jones 1983) left us dissatisfied with the original set of case relations. We therefore carried out a detailed analysis of a large sample of English sentences to evaluate our proposals for a better-founded and more comprehensive set of case relations. This study exploited F.T. Wood's "English prepositional idioms" (Wood 1967), which provides a careful account, supported by extensive examples, of the uses of English prepositions and preposition-like terms. For instance,

**WITHIN**

(1) Inside
Within the house all was quiet.
The Kingdom of God is within you.

(2) Amongst the members of a group.
Opinion within the profession is divided.

(3) Inside specified bounds or limits.
They were ordered to remain within the precincts of the college.
The scholarship is open to anyone residing within fifty miles of the university.
He always strove to live within his income.

... Our study was intended to establish both the justification for each case relation individually, by reference to a range of sentences, and the plausibility of the set of relations as a whole, by reference to the complete set of sentences. Looking at Wood's description of a preposition's sense, and its accompanying illustration(s), we tried to assign a case label to the link between the sentence elements made by the preposition which we felt captured the essential nature of that link, at the level of generality represented by a set of 20-30 cases. Thus "location" would be the label associated with a number of specific space-relation prepositions, e.g., above, at, by. The study was primarily concerned with prepositionally-
based cases, but we considered other case instantiations, though not systematically. While additional cases might therefore be motivated by non-prepositional sources we have not examined, we believe that a complete set would have to include something like our list, to cover prepositions adequately.

The study was a purely investigative one; we were interested in the legitimacy of the cases as characterisations of sentence relations, and did not address the question of how the specific assignments, for the individual sentences, could be achieved automatically. However, as we had already demonstrated that a quite refined set of cases could be applied by Boguraev's existing analyser, and have since, as practical need has arisen, implemented further cases, we feel some confidence in the feasibility of automatic assignment of the cases in the present set. (This of course accepts that some individual sentences may present considerable difficulties.) We are nevertheless not offering our results here with any great claims about the especial novelty or merits of our case set, but rather as a practical contribution to the utilisation of case labels, justified by the large and varied sentence sample studied. Though domain-specific strategies may reduce the need to handle lexical ambiguity, especially in nouns and verbs but even in prepositions, it may nevertheless be necessary even in the domain-specific case to provide for distinct semantic functions in prepositions.

Wood's complete set of sentences is substantial; we took one sentence for each straightforward sense defined by Wood's numbering, omitting items labelled "phrases" and with occasional modifications, e.g., to shorten sentences. This gave us 421 sentences altogether, for 86 prepositions. As a result of our analysis we emerged with 28 cases. These are listed below with notes on the types of source from which they may be derived, and (our) examples showing both prepositional and non-prepositional uses. This list is followed by illustrations from the Wood sample.

In the list each case, or tag, name is prefaced by its abbreviation. The meaning of the case is not explicitly defined, but is assumed to be ostended by the subsequent examples (and more particularly by the fuller sample derived from Wood). Possible sources for the cases include the lexicon and syntactic structures like embedded clauses identified by the parsing program (and in our analyser elements of structure identified by semantic pattern matching). The normal linkage marked by a case is between verb and noun group, but two nominal groups may also be case-linked, and also more than two items (though this is not common). In the illustrations we have adopted the convention of marking the head words of the two linked constituents by –. Thus the tag "after" labels the relation between left and breakfast in *John left following breakfast*, written as “John –left following –breakfast”. In the cases where complete constituents like embedded clauses fill case roles, they are bracketed with [ ], and the whole item is marked with –. The cases are alphabetically ordered here on their abbreviations.

(aec) ACCOMPANIMENT
dictionary: e.g., “with”; can link nominals
John –went to the zoo (along) with –Mary.  
John –went everywhere with his –violin.

(act) ACTIVITY
dictionary: e.g., “at”
John –beat Mary at –chess.  
Aloysius –beat Sebastian –running.

(adest) ABSTRACT-DESTINATION
dictionary: e.g., “to”
I –reached my –conclusion.  
When heated, water –turns into –steam.

(aft) AFTER
dictionary: e.g., “after”, program: ing-phrases, adverbials; can link nominals
John –left after –Bill.  
John, –[having bought the book], –took it home.

(ag) AGENT
program: contents of “subj” register
John was –hit by –Bill.  
–Malaria –killed the girl.

(aloc) ABSTRACT-LOCATION
dictionary: e.g., “in”
Parry was –fixed in his –mind that the Mafia was out to get him sometime.

(asour) ABSTRACT-SOURCE
dictionary: e.g., “from”
I –got them from –Bill.  
She –makes dresses of –silk.

(attr) ATTRIBUTE
dictionary: e.g., “with”; can link nominals
The –girl in –blue was happy.  
The –girl with the pink –hat was sad.

(bef) BEFORE
dictionary: e.g., “before”, program: adverbials; can link nominals
John –left before –noon.  
The stock market –was very active ahead of the –Budget.

(comp) COMPARISON
dictionary: e.g., “as”, program ?; can link nominals
John –passed as –Bill.  
He was given –cash instead of –kind.  
Time –flies like an –arrow.

(dest) DESTINATION
dictionary: e.g., “to”
John –went to –Paris.  
Bill –walked up to –John.
Our treatment of the sample of sentences taken from Wood is illustrated by the two selections below. These show first the complete analyses for two prepositions, and then the complete set of sentences for two cases. The average number of sentences per case in the sample is 15, with low variation. However, “location” has 80 sentences, reflecting the very large number of different space-relation prepositions there are. It is possible that more specific space-location cases would be desirable, though the right level of discrimination and particular set would not be easy to establish. We are, in any case, not suggesting that the use of a case tag in the representation of a sentence delivered by the analyser makes it unnecessary to indicate the specific lexical sense of the preposition.

**AT**

/loc/ Mr Brown is at the office.
/loc/ The concert starts at half past seven.
/loc/ At the second roundabout, turn left.
/loc/ We have arrived at our destination.
/act/ The two tribes were constantly at war.
/act/ She loved to watch the children at play.
/man/ He left at a moment’s notice.
/man/ He vaulted over the gate at a single leap.
/force/ I have called to see you at the request of a friend.
/man/ Cars are parked at their owners’ risk.

(poss) POSSESSED-BY

program: noun phrases; only links nominals
The daughter of the Mayor was blonde.
The pocket of my uncle’s coat was empty.

(quant) QUANTITY

program: noun phrases; can link nominals
There were four students including a dwarf.
It was too long by twelve inches.

(reas) REASON

dictionary: e.g., “of”, “because of”, program: embedded clauses
John is afraid of [being apprehended by the police].
John is happy about school.
The car—whizzed along at sixty—miles an hour. She felt sure she would—faint at the—sight of blood.

As we passed the gate the dog—flew out at—us. That boy—is clever at—mathematics.

The new library was—opened by the—Mayor. I shall—go by the 10:30—train.

He—seized me by the—arm. He—succeeded by hard—work.

I—know him by—sight. We—lost the match by one—goal.

She—sat by the—pillar. We—went by the—shop without realising it.

They should have—been here by—now. They decided to—travel by—night.

She—did the work to the—best of her ability. She—lay back with her—eyes closed.

The car—whizzed along at sixty—miles an hour. The child—screamed in—terror.

The last few years of his life were—spent in—poverty. The water—shot up in a—fountain.

He—is, without—doubt, a very able person. He—left at a moment’s—notice.

He said nothing in—reply to my questions. I—did it without—thinking.

He—succeeded by hard—work. I—know him by—sight.

I only—said it in—fun. It—is all right in—theory.

She—did the work to the—best of her ability. She—lay back with her—eyes closed.

The car—whizzed along at sixty—miles an hour. The child—screamed in—terror.

The last few years of his life were—spent in—poverty. The water—shot up in a—fountain.

Our complete sample is available in listings, with a fuller description. We acknowledge the immense value of Wood’s work as a source, and are grateful to the publisher Macmillan for permission to reproduce and use Wood’s material.

References

Boguraev, B.K. 1979 Automatic Resolution of Linguistic Ambiguities. Technical Report 11, Computer Laboratory, University of Cambridge. Boguraev, B.K. and Sparck Jones, K. 1983 How to Drive a Database Front End Using General Semantic Information. Conference on Applied Natural Language Processing, Santa Monica: 81-88. Wilks, Y. 1977 Good and Bad Arguments about Semantic Primitives. Communication and Cognition 10: 181-221. Wood, F.T. 1979 English Prepositional Idioms. Macmillan, London, England.