Having Your Cake and Eating It Too: Autonomy and Interaction in a Model of Sentence Processing

Kurt P. Eiselt*  
College of Computing  
Georgia Institute of Technology  
Atlanta, Georgia 30332-0280  
eiselt@cc.gatech.edu

Kavi Mahesh*  
College of Computing  
Georgia Institute of Technology  
Atlanta, Georgia 30332-0280  
mahesh@cc.gatech.edu

Jennifer K. Holbrook  
Department of Psychology  
Albion College  
Albion, Michigan 49224  
jen@cedar.cic.net

Abstract

Is the human language understander a collection of modular processes operating with relative autonomy, or is it a single integrated process? This ongoing debate has polarized the language processing community, with two fundamentally different types of model posited, and with each camp concluding that the other is wrong. One camp puts forth a model with separate processors and distinct knowledge sources to explain one body of data, and the other proposes a model with a single processor and a homogeneous, monolithic knowledge source to explain the other body of data. In this paper we argue that a hybrid approach which combines a unified processor with separate knowledge sources provides an explanation of both bodies of data, and we demonstrate the feasibility of this approach with the computational model called COMPERE. We believe that this approach brings the language processing community significantly closer to offering human-like language processing systems.

The Big Questions

Years of research by linguists, psychologists, and artificial intelligence specialists have provided significant insight into the workings of the human language processor. Still, fundamental questions remain unanswered. In particular, the debate over modular processing versus integrated processing rages on, and experimental data and computational models exist to support both positions. Furthermore, if the integrated processing position is correct, just what exactly is integrated? And if the modular position is the right one, what are the different modules? Do they interact, and if so, to what extent and when? Or are those modules entirely autonomous?

Wrestling with these questions induces considerable frustration in researchers. This frustration stems not only from the research community's apparent inability to answer them satisfactorily, but also from the overwhelming importance of the answers themselves—these answers, once uncovered, undoubtedly will impact thinking in all areas of artificial intelligence and cognitive science research, including visual processing, reasoning, and problem solving, to name just a few. In this paper, we intend to provide the reader with answers to some of these questions—answers based on nearly ten years of our own interdisciplinary research in sentence processing, and built upon the work of many others who went before us. In brief, we propose a model of language understanding (or, more specifically, sentence processing) in which all linguistic processing is performed by a single unified process, but the different types of linguistic knowledge necessary for processing are separate and distinct. This model accounts for conflicting experimental data, some of which suggests an autonomous, modular approach to language processing, and some of which indicates an integrated approach. Because it is a closer fit to the experimental data than any model which has gone before, this model consequently points the way to more human-like performance from language processing systems.

Background

Our new model of sentence processing has its roots in work begun nearly ten years ago. That research effort started as an attempt to explain how the human language understander selected the most context-appropriate meaning of an ambiguous word, and then was able to correct both the choice of word meaning and the surrounding sentence interpretation, without reprocessing the input, when later processing showed that the initial choice of word meaning was erroneous.

The resulting computational model, ATLAST (Eiselt, 1987; Eiselt, 1989), resolved word sense ambiguities by activating multiple word meanings in parallel, selecting the meaning which matched the previous context, and deactivating but retaining the unchosen meanings for as long as resources were available for retaining them. If later context proved the initial decision to be incorrect, the retained meanings could be reactivated without reaccessing the lexicon or reprocessing the text. ATLAST proved to have great psychological validity for lexical processing—its use of multiple access was well grounded in the psychological literature (e.g., Seidenberg, Tanenhaus, Leiman, & Bienkowski, 1982), and, more importantly, it made psychological predictions about the retention of unselected meanings that were...
experiments inherent in backtracking by postponing any de-
avoiding erroneous decisions was offered by parallel parsers
of recovery from erroneous syntactic decisions, nor did it
decision was discovered, and then reinterpret the input (e.g.,
a correct decision, thereby avoiding erroneous decisions
cision until enough input had been processed to guarantee
which maintained all plausible syntactic interpretations at
the same time (Kurtzman, 1985). ATLAST, however, was a
model of semantic processing and did not address the issue
showing that in dealing with syntactic ambiguity, the sen-
tence processor accesses all possible syntactic structures
the decision is delayed. Furthermore, the work suggests
an interaction of the various knowledge types, as in some
cases semantic information influences structure assignment
or triggers reactivation of unselected structures. This model
of limited delayed decision in syntactic ambiguity resolution
had much in common with the ATLAST model of semantic
ambiguity resolution. Both models proposed an early commit-
ment where possible. Both models had the capability to
pursue multiple interpretations in parallel when ambiguity
made it necessary. Both models explained error recovery
as an operation of switching to another interpretation main-
tained in parallel by the sentence processor. Finally, both
models made decisions by integrating the preferences from
syntax and semantics.

One explanation for this high degree of similarity between
the syntactic and semantic error recovery mechanisms is that
there are two separate processors, one for syntax and one for
semantics, each with its corresponding source of linguistic
knowledge, and each doing exactly the same thing. A more
economical explanation, however, is that there is only one
process which deals with syntactic and semantic information
in the same manner. We have chosen to explore the latter
explanation, as others have done, but we have also chosen to
maintain the separate knowledge sources for reasons which
will be explained below. (See also Holbrook, Eiselt, &
Mahesh, 1992.)

Overview of COMPERE

Our new model of sentence processing, called COMPERE
(Cognitive Model of Parsing and Error Recovery), consists of
a single unified process operating on independent sources
of syntactic and semantic knowledge. This is made possible
by a uniform representation of both types of knowledge. The
unified process applies the same operations to the different
types of knowledge, and has a single control structure which
performs the operations on syntactic and semantic knowl-
edge in tandem. This permits a rich interaction between the
two sources of knowledge, both through transfer of control
and through a shared representation of the interpretations of
the input text being built by the unified process.

An advantage of representing the different kinds of
knowledge in the same form is that the boundaries between
the different types of knowledge can be ill-defined. Often
it is difficult to classify a piece of knowledge as belonging
to a particular class such as syntactic or semantic. With a
uniform representation, such knowledge lies in between and
can be treated as belonging to either class.

Syntactic and semantic knowledge are represented in sep-
ate networks in which each node is a structured representa-
tion of all the information pertaining to a syntactic or
semantic category or concept. A link, represented as a slot-
filter pair in the node, specifies a parent category or concept
of which the node can be a part, together with the condi-
tions under which it can be bound to the parent, and the
expectations that are certain to be fulfilled should the node
be bound to the parent. In addition, nodes in either network
are linked to corresponding nodes in the other network so
that the unified process can build on-line interpretations of
the input sentence in which each syntactic unit has a corre-
sponding representation of its thematic role and its meaning.
In addition, there is a lexicon as well as certain other minor
heuristic and control knowledge that is part of the process.

COMPERE's architecture and knowledge representation
are displayed graphically in Figures 1 and 2.)

The unified process is a bottom-up, early-commitment
parsing mechanism integrated with top-down guidance
through expectations. The operators and the control struc-
ture that constitute the unified process are described briefly
in the algorithm shown in Figure 3.

The COMPERE prototype has been implemented in
Common LISP on a Symbolics LISP Machine. At this time,
it's unified process can perform on-line interpretations of its
input, and can recover from erroneous syntactic decisions
when necessary. COMPERE is able to process relatively
complex syntactic structures, including relative clauses, and

Natural Language Sentence Analysis 381
can resolve the associated structural ambiguities.

**Autonomy and interaction effects from one process**

COMPERE is able to exhibit seemingly modular processing behavior that matches the results of experiments showing the autonomy of different levels of language processing (e.g., Forster, 1979; Frazier, 1987). It is also able to display seemingly integrated behavior that matches the results of experiments showing semantic influences on syntactic structure assignment (e.g., Crain & Steedman, 1985; Tyler & Marslen-Wilson, 1977). For example, consider the processing of the following sentence:

(1) The bugs moved into the new lounge were found quickly.

This sentence has a lexical semantic ambiguity at the subject noun *bugs* that could mean either insects or electronic microphones. In addition, it is also syntactically ambiguous locally at the verb *moved* since there is no distinction between its past-tense form and its past-participle form. In the simple past reading of *moved*, it would be the main verb with the corresponding interpretation that “the bugs moved themselves into the new lounge.” On the other hand, if *moved* is read as a verb in its past-participle form, it would be the verb in a reduced relative clause corresponding to the meaning “the bugs which were moved by somebody else into the new lounge....” Parse trees for the two structural interpretations and the corresponding thematic-role assign-
Null Context: When sentence (1) is presented to COMPERE in a null semantic context, one where there is no bias for either meaning of the noun bugs, COMPERE reads ahead without resolving the lexical ambiguity at the word bugs. When it encounters the structural ambiguity at the verb moved, COMPERE does not have the necessary information to decide which of the two structures in Figures 4 and 5 is the appropriate one to pursue.

However, COMPERE has a syntactic preference for the main-verb interpretation over the relative clause one. Though this preference can be explained by the minimal attachment principle (Frazier, 1987), COMPERE offers a more general explanation. Extrapolating from Stowe's model, we have endowed COMPERE with the pervasive goal of completing an incomplete item at any level of processing. In syntactic processing, it has a goal to complete the syntactic structure of a unit such as a phrase, clause, or a sentence. COMPERE prefers the alternative which helps complete the current structure (called the Syntactic Default) over one that adds an optional constituent leaving the incompleteness intact. For instance, in (1), a VP is required to complete the sentence after seeing The bugs. Since the main-clause interpretation helps complete this requirement and the relative-clause interpretation does not, the main-clause structure gets selected. In other words, COMPERE would rather use the verb to begin the VP that is required to complete the sentence structure than treat it as the verb in a reduced relative clause which would leave the expectation of the VP unsatisfied. This behavior is the same as the one explained by the “first analysis” models of Frazier and colleagues (Frazier, 1987) using a minimal-attachment preference. COMPERE can produce this behavior by applying structural preferences independently since it maintains separate representations of syntactic and semantic knowledge.

As a consequence of choosing the main-clause interpretation, the lexical ambiguity is also resolved. The electronic bug meaning is now ruled out since there is a selectional restriction on the verb moved that is not satisfied by electronic bugs (namely, they cannot move by themselves).

Thus, until seeing the word were, the verb moved is treated as the main verb since it satisfies the expectation of a VP that is required to complete the sentence. However, at this point, the structure is incompatible with the remaining input. COMPERE recognizes the error and now tries the alternative of attaching the VP as a reduced relative clause so that there is still a place for a main verb. This results in a garden-path effect upon reading this sentence. That is, the sentence processor is led up a garden path and has to backtrack when later information shows that it was the wrong path to take. This behavior is not influenced by semantic or conceptual preferences and can be perceived as a modular behavior. COMPERE's error recovery method was first developed in the ATLAST model (Eiselt, 1987). It was also experimentally validated (Eiselt & Holbrook, 1991).

As a consequence of switching to the new syntactic interpretation, COMPERE makes corresponding changes to thematic role assignments and also "unresolves" the lexical ambiguity. There is no longer any reason to eliminate the electronic bug meaning since either kind of bugs can be moved by others.

Semantically Biasing Context: Now consider sentence (1) in a semantically biasing context such as the one in (2).

3COMPERE's program does not resolve lexical semantic ambiguities at this time. We are currently rectifying this by incorporating lexical ambiguity resolution strategies from our earlier model ATLAST (Eiselt, 1989) in COMPERE.

4At present, COMPERE is not capable of using context effects in its ambiguity resolution process. However, its architecture supports the inclusion of such effects and we are working on providing context information to the unified process.
The semantic context in (2) resolves the lexical ambiguity by choosing the electronic bug meaning. This decision helps COMPERE resolve the structural ambiguity at the verb moved. Using its conceptual knowledge, represented as a selection restriction, that only animate agents can move by themselves, COMPERE decides that moved cannot be a main verb and goes directly to the reduced relative clause interpretation (Fig. 5), thereby avoiding the garden path. This shows how the same unified process that previously exhibited modular processing behavior can also produce interactive processing behavior when semantic information is available. Syntax and semantics interact in COMPERE to help resolve ambiguities in each other.

COMPERE can also use independent syntactic preferences in other types of sentences such as those with prepositional attachment ambiguities. The COMPERVER prototype thus demonstrates that the range of behaviors that the interactive models account for (Crain & Steedman, 1985; Tyler & Marslen-Wilson, 1977), and the behaviors that the “first analysis” models account for (Frazier, 1987), can be explained by a unified model with a single processor operating on multiple independent sources of knowledge.

**Comparative evaluation**

There is certainly nothing unique about a unified process model of language understanding—the integrated processing hypothesis has been visited and revisited many times, for good reason, and with significant results (e.g., Jurafsky, 1992; Lebowitz, 1980; Riesbeck & Martin, 1986). Yet each of these models labors under the assumption that the integration of processing necessarily goes hand in hand with the integration of the knowledge sources. While this design decision may make construction of the corresponding computational model easier, it also makes the model incapable of easily explaining the autonomy effects demonstrated by Forster (1979), Frazier (1987), and others. As shown above, COMPERE’s unified processing mechanism combined with its separate sources of linguistic knowledge offers an explanation for observed autonomy effects as well as the interaction effects reported by Marslen-Wilson and Tyler (Tyler & Marslen-Wilson, 1977). Furthermore, the integrated models noted above cannot capture syntactic generalizations.

Another form of the modularity debate concerns the effect of context on syntactic decisions—does context affect structure assignment, or are context effects absent until later in language processing (Taraban & McClelland, 1985)? Though we do not have a model of context effects in COMPERE, we believe that contextual information can be incorporated as an additional source of preferences in COMPERE’s architecture.

An added benefit of COMPERE’s sentence processing architecture is that it offers an explanation for the effects of linguistic aphasias. In reviewing the aphasia literature, Caramazza and Berndt (1978) concluded that the evidence pointed strongly to the functional independence of syntactic and semantic processing. COMPERE suggests an alternate explanation—the different aphasic behaviors are not due to damage to the individual processors, but are instead due to damage to the individual knowledge sources or, perhaps, to the communications pathways between the knowledge sources and the unified processor.

We believe that COMPERE’s architecture accounts for the wide variety of seemingly conflicting data on linguistic behavior better than any previously proposed model of sentence processing. Yet COMPERE is not the first sentence processing model to be configured as a single process interacting with independent knowledge sources. The localist or punctate connectionist models of Pollack (1987; Waltz and Pollack, 1985) and Cottrill (1985; Cottrill and Small, 1983) resemble COMPERE at a gross architectural level, but these models did not offer the range of explanation of different behaviors that COMPERE does; for example, these models do not recover from errors, nor can they deal with complex syntactic structures such as relative clauses.

Despite all its theoretical advantages over other models, the prototype implementation of COMPERE is not yet fully developed and suffers from some weaknesses. Its role knowledge is fairly limited, and its conceptual knowledge is even more so. Also, the implementation currently diverges slightly from theory. The divergence appears in the process itself: the theoretical model has a single unified process, while the prototype computational model consists of two nearly-identical processes—one for syntax and one for semantics. These two processes share identical control structures, but they are duplicated because we have not yet completed the task of representing the different types of information in a uniform format. Some readers may take this as an indication that we are doomed to failure, but the connectionist models mentioned earlier serve as existence proofs that finding a uniform format for representing different types of linguistic knowledge is by no means an impossible task.

**Conclusion**

Is the human language understander a collection of modular processes operating with relative autonomy, or is it a single integrated process? This ongoing debate has polarized the language processing community, with two fundamentally different types of model posited, and with each camp concluding that the other is wrong. One camp puts forth a model with separate processors and distinct knowledge sources to explain one body of data, and the other proposes a model with a single processor and a homogeneous, monolithic knowledge source to explain the other body of data. In this paper we have argued that a hybrid approach which combines a unified processor with separate knowledge sources provides an explanation of both bodies of data, and we have demonstrated the feasibility of this approach with the computational model called COMPERE. We believe that this approach brings the language process-
ing community significantly closer to offering human-like language processing systems.

Acknowledgement: We would like to thank Justin Peterson for his comments on this work and his help in finding good examples.

References


