On the Role of the Simplicity Heuristic in Language Processing: Evidence from Structural and Inferential Processing

Lee Osterhout[1] **and David A. Swinney**[2,3]

This paper evaluates whether or not simplicity or "minimalistic" heuristics, which are posited to account for local ambiguity resolution at the level of structural processing and for the lack of inferential elaboration during discourse processing, represent the most accurate account of language processing. Evidence from on-line studies is presented which suggests that alternative, more knowledge-based mechanisms for handling processing are brought to bear in these situations.

Perhaps the best-known (and certainly most ubiquitous) characteristic of language that needs to be confronted by language processing models is ambiguity. Ambiguity (and the decisional uncertainty that accompanies it) exists at every putative level, of language processing-phonological, lexical, structural, interpretative. Importantly, formal representations of language (such as linguistic theories of grammar) are neither intended to, nor devised in such a way as to allow for, either consideration or resolution of uncertainty over the ambiguity encountered by a processor-rather, such resolution is an issue left solely, to processing models.

[1] Northwestern University, Evanston, Illinois.

[2] Graduate Center of the City University of New York.

[3] Address all correspondence to David A. Swinney, Department of Linguistics, The Graduate School and University Center of The City University of New York, 33 West 42nd Street, New York, New York 10036.

One of the most common approaches found in processing models that attempt to deal with situations of ambiguity or uncertainty is to posit "psychologically based" heuristics motivated by notions of computational simplicity or "minimal effort." Generally, any particular simplicity heuristic is, posited to account for the resolution of uncertainty at only a particular source or level of processing, guiding the listener/reader to an initial interpretation free of uncertainty. For example, Fodor, Frazier, and their associates (e.g., Frazier, 1978; Frazier & Fodor, 1978) have claimed that readers make use of a "minimal attachment" heuristic when confronted with local syntactic ambiguities. In this, the simplest analysis (as determined by the number and depth of phrase structure nodes) is initially computed, followed by backtracking and reanalysis only when this simplicity heuristic turns out to be inappropriate.

Similar heuristics based on principles of minimal effort or simplicity have been proposed by a number of researchers to operate at various other levels of processing-lexical access, coreference assignment, inferential processing, and the like. These heuristics have appeal on a number of grounds, the most obvious being that arguments in science are deemed relatively more valuable, ceteris paribus, if they account for evidence on the basis of simple, rather than more complex, principles. In addition, simplicity is a metric which has presumed general application, and hence is often viewed as a logical candidate for guiding cognitive/psychological processing in general.

Despite the value of such heuristics, however, it seems that they are often the choice of last resort; that is, failure to find evidence supporting knowledge-based ambiguity or uncertainty resolution processes often seems to leave arguments appealing to simplicity or minimalism as the only defensible alternative. And, indeed, in a field in as much flux as language processing, simplicity arguments are hard to argue against, particularly when simplicity can be defined in terms of so many as yet empirically untested principles. It is for this reason that it seems worthwhile to constantly test predictions stemming from such heuristics, on the off chance that more sensitive tests or new experimental conditions may add to our knowledge of the bounds between such general-processing principles as compared with more knowledge-based processes operating during language comprehension and production. In what follows, we present on-line evidence concerning two areas of language processing in which simplicity or minimalistic hypotheses have held considerable influence-structural assignment and inferential elaboration-and we argue that alternative mechanisms must be considered.

ASSIGNMENT OF GRAMMATICAL ROLES AND COMPUTATION OF CONSTITUENT STRUCTURE

Local structural ambiguities, in which the proper grammatical role of a given word or phrase is temporarily uncertain, are ubiquitous in natural languages. For example, the italicized NP in (1) might be the object of the verb, as in continuation (a), or the subject of an upcoming sentential complement, as in continuation (b):

(1) The lawyer believed *the defendant* ...
     (a) and took the case.
     (b) was lying.

Considerable evidence from eye movement studies (gaze durations and regressions) has been interpreted to indicate that readers rapidly assign a single grammatical role to each word and phrase as it is encountered, even when faced with uncertainty as to the correct grammatical assignments (i.e., a serial account of dealing with structural ambiguity in parsing) (see, e.g., Frazier & Rayner, 1982; Rayner, Carlson, & Frazier, 1983; Ferreira & Clifton, 1985). Because it is argued that only a single structure is computed at a time, it is necessary for a processing model to provide principles accounting for precisely which one of the several possible analyses is initially pursued. Frazier and her associates (Frazier, 1978; Frazier & Fodor, 1978) proposed the "minimal attachment" heuristic to this end, in which the simplest structure (as defined by the number of nodes in the phrase/tree structure) is attempted first, followed by backtracking and reanalysis if this structure turns out to be incorrect.

The appeal of the minimal attachment heuristic lies in its appeal to simplicity and in the claim that it is applied in all (or most) cases of syntactic uncertainty. However, even assuming that a serial model is the correct one, other possible heuristics are conceivable. For example, Clifton, Frazier, and Connine (1984) provide evidence of the influence of verb subcategorization frame and/or argument structure biases during sentence comprehension. Such biases could be used to resolve certain types of syntactic uncertainty; the "preferred" frame or argument structure (perhaps as determined by frequency of usage) might be attempted first (see also Tanenhaus, Boland, Garnsey, & Carlson, 1989, for evidence to this effect). Therefore, as a number of authors have noted, it is critically important to determine whether or not the minimal attachment analysis of an ambiguous string is initially computed even

when other reasonable heuristics (suggesting a nonminimal analysis) could be employed.

We have investigated this question by recording event-related potentials (ERPs) elicited during the comprehension of sentences containing syntactic ambiguities. The value of this approach lies in the fact that ERPs are an on-line reflection of processing activity that do not require simultaneous performance of some decision task during language comprehension (as is the case with monitoring tasks, continuous reading, syntactic decision tasks, etc.). In particular, the use of ERPs to examine structural processing is justified by recent work demonstrating that certain syntactic anomalies (e.g., violations of subcategorizational constraints) elicit a positive-going ERP component with an onset about 500 milliseconds following the violation (Osterhout, 1989; Osterhout & Holcomb, 1989). Thus, there appears to be an ERP component sensitive to the processing of syntactic variables. We have employed it to examine the question of how the processing device deals with temporary structural ambiguity as a function of the presence or absence of verb information that might help resolve the issue. In this study subjects were presented (visually) with sentences of the following type:

> (1) The lawyer hoped <u>the defendant</u> <u>was</u> lying.
>   (pure intransitive)
> (2) *The lawyer forced <u>the defendant</u> <u>was</u> lying.
>   (pure transitive)
> (3) The lawyer knew <u>the defendant</u> <u>was</u> lying.
>   (biased intransitive)
> (4) The lawyer believed <u>the defendant</u> <u>was</u> lying.
>   (biased transitive)

In these sentences, transitivity properties associated with the main verb have been manipulated. The question of interest concerns how readers assign grammatical roles to the underlined NP in the presence of this verb-based information about the structure of the underlined NP; basically, the issue concerns whether the processor will use such information on line to reduce processing uncertainty, whether it will rely on the "simplicity metric" (minimal attachment), or whether some other (nonserial) process will be demonstrated. The intransitive verb in (1) indicates that the NP will not be a simple object and should be treated as the subject of the upcoming complement (V-S' analysis). The transitive verb in (2) indicates the appropriate structure to be NP-as-object-of-verb (V-NP analysis); note that under this analysis the subsequent material

("was lying") renders the sentence ungrammatical. The main verbs in (3) and (4) can occur with or without a direct object NP, rendering the correct grammatical role of the postverbal NP temporarily uncertain. However, the verb in (3) is one that is statistically biased toward appearing as an intransitive verb and that in (4) is biased toward appearing as a transitive verb.

The crucial comparisons concern decisions made during comprehension of sentences like (3) and (4). According to strict versions of the minimal attachment hypothesis, readers should initially construct the (incorrect) V-NP analysis for sentences like (3) and (4) (and, of course, (2)). Given prior findings of a positive ERP component associated with syntactic anomaly, if these incorrect transitive interpretations are assigned, one would expect this positive ERP component to be elicited when the subject encounters the word was in sentences of type (2), (3), and (4). In contrast, if verb subcategory biases determine the initial assignment of grammatical roles, only sentence types (2) and (4) should elicit the syntactic anomaly ERP component; sentences like (3) should not, since the verb bias is consistent with the correct V-S' analysis.

Sentences were presented visually in a word-by-word manner, with each word appearing by itself for 300 ms on a CRT screen and with a 350-ms blank screen separating words. Figure 1 shows wave forms observed at site Pz following presentation of the (double underlined) word was in each sentence type, averaged across 30 exemplars per condition and 12 subjects. Inspection of Figure 1 reveals that ungrammatical sentences like (2) elicited a late positive component relative to the other sentence types with an onset about 500 ms poststimulus and with a duration of about 300 ms $(F(1, 11) > 10, p < .05$ in all comparisons) Additionally, sentences containing transitively biased verbs also appeared to elicit a positive component relative both to sentences containing pure intransitive $(F(1, 11) = 7.46, p < .05)$ and to sentences containing intransitively biased verbs $(F(1, 11) = 10.16, p < .05)$. No significant differences in the ERPs elicited by the pure intransitive and biased intransitive sentences were observed.

Thus, an ERP component associated with syntactic anomaly was elicited following transitively biased verbs but not following intransitively biased verbs in intransitive structures. This observation is most consistent with the claim that readers initially computed a single structure (the serial hypothesis) for sentences like (3) and (4), and that information about verb transitivity preferences determined which analysis was initially pursued. Importantly, these data are clearly not consistent with claims that the language processor relies on a strict minimal attachment
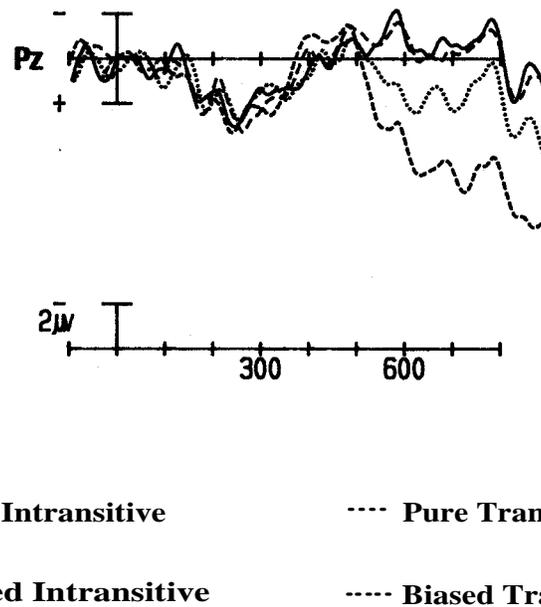
Fig. 1. Mean electrical activity recorded over site Pz during presentation of auxiliary verb in each of four sentence types.

heuristic -- the simplicity heuristic -- to resolve cases of local syntactic ambiguity.

INFERENCE GENERATION DURING DISCOURSE PROCESSING

Language comprehension viewed at the level of discourse poses a slightly different type of uncertainty to the comprehension device. One of the hallmarks of the construction of a discourse representation during comprehension is that much information is left implicit in a discourse; the listener/reader is expected to make the proper inferences concerning necessary, presupposed, and relevant information not explicitly mentioned in the discourse. However, it is clear that not all potential inferences can be; made, since even the shortest discourse can license an essentially unbounded number of inferences. Hence, the listener/reader

Table I. Sample Paragraph and Targets of the Type Used in Inference Experiment[a,b]

First sentence	
Script-suggestive context:	"John sat down to a meal of meat and potatoes."
Neutral context:	"John sat down at his usual place at the table."
Second (implied use) sentence	* He cut the juicy meat * and ate his bread.
Visual targets	Knife (Instrumental-related target)
	Frame (Control target)

[a]*Source:* Osterhout and Swinney (1989).
[b]An asterisk marks the approximate locations at which target letter strings were visually presented.

must decide which inferences among the plethora of possible inferences should be generated. In this sense, discourse-based inference generation, like the other processes mentioned above, requires the resolution of uncertainty: What information should be added to the discourse representation? What types of inferences are made, and, under what conditions?

Several authors have argued, in what has come to be the standard position in this field, that comprehenders rely on a "minimal inference" heuristic during the construction of discourse representations (Corbett & Dosher, 1978; Dosher & Corbett, 1982; McKoon & Ratcliff, 1986). Listeners/readers are argued to generate only those inferences that are minimally necessary for discourse coherence (e.g., inferences concerning causal or coreferential relations between discourse elements). Inferences that are elaborative rather than necessary for logical coherence (e.g., inferences concerning such things as implied instruments or predictable consequences) are argued not to be generated under the strong version of this hypothesis.

We have recently investigated this general claim of minimal inference generation in an experiment designed to investigate the role of scripted information and contextual cues on the generation of one type of "elaborative" inference-that involving use of an implied instrument (Osterhout & Swinney, 1989). In this study, subjects were presented auditorily with short two-sentence paragraphs such as the one displayed in Table I. Two versions of the first of the two sentences were constructed for each paragraph. One version presented a standard, scripted scenario (e.g., eating a meal), while the second version provided a neutral context omitting the scripted scenario. The second sentence of the short paragraph always implied the use of an instrument often found in the

Table II. Mean Priming Latencies (Control Target RTs, minus Instrument Targets RTs in MS) at Each of the Two Test Points

	Target test point	
Context type	1	2
Script-suggestive	9	47 [a]
Neutral	4	19

[a]$p < .05$.

suggested scenario. Inference generation was measured by a cross-modal priming task in which experimental (instrument-related) and control visual targets (to which subjects made lexical decisions) were presented at two points during paragraph presentation (to different subjects in different conditions): immediately after the context sentence (before the second sentence), and immediately after the end of the verb phrase implying use of the instrument. The premise behind use of this technique is that, if the implied instrument is activated during the ongoing (perceptual) processing of the discourse, it will cause priming (facilitation) for the lexical decision to the instrument-related target as compared with the control target at that point at which the inference is elaborated-if ever (see, e.g., Swinney, 1981, for description of the task).

Table II displays mean priming scores (response times to control words minus response times to experimental instrument-related target words) for each of the two probe points. As can be seen, no priming of instrument targets was found immediately after script-suggestive context, nor was any found immediately after the neutral context. Additionally, no priming was found following the verb phrase implying use of the instrument in the neutral context condition. However, in the condition involving *both* the script-suggestive context and the subsequent verb phrase implying use of the instrument, significant priming of the instrument target was produced, suggesting that elaboration of the instrumental inference took place at this point.

These data support our argument that, contrary to the minimal elaboration hypothesis, certain types of "elaborative" inferences *are* in fact reliably made under certain conditions. Note that the conditions under which this inferential elaboration was produced are those involving a confluence of knowledge-based information; each type of context alone did not produce the effect (and hence, appeals to simple associativity controlling this elaboration appear ruled out). While the precise mecha-

nism controlling elaboration of such inferences remains to be determined, the critical issue here is that the minimalist argument does not appear to stand up under close scrutiny.

SUMMARY

This paper has briefly reviewed two new pieces of evidence concerning the resolution of uncertainty and ambiguity during language comprehension. This evidence has disconfirmed predictions from two standard models of structural and discourse processing that rely on notions of computational or psychological simplicity (minimalism) as explanatory mechanisms. When confronted with local syntactic ambiguities, readers initially chose the path suggested by verb subcategory biases, rather than the structurally simpler path predicted by the minimal attachment heuristic. Additionally, listeners demonstrated generation of "elaborative" inferences under certain contextual conditions, a finding not in keeping with a "minimal inference" heuristic for discourse analysis. In addition to the obvious claims about the specific mechanisms involved in structural and discourse processing, we take the evidence presented here to argue for caution in reliance on pervasive general "psychologically relevant" heuristics such as simplicity in processing models. The intrinsic appeal of simplicity may, as it appears to have done in the cases cited here, act to mask the fact that its explanatory power may be almost impossible to evaluate given the ever-changing nature of the theoretical representational base on which it is applied, particularly in areas of language and linguistics. It should remain a heuristic of last resort only.

REFERENCES

Clifton, C., Frazier, L., & Connine, C. (1984). Lexical expectations in sentence processing. *Journal of Verbal Learning and Verbal Behavior, 23,* 696-708.

Corbett, A.T., & Dosher, B.A. (1978). Instrument inferences during reading. *Journal of Verbal Learning and Verbal Behavior, 17,* 479-491.

Dosher, B.A., & Corbett, A.T. (1982). Instrument inferences and verb schemata. *Memory and Cognition, 10,* 531-539.

Ferreira, F., & Clifton, C. (1985). The independence of syntactic processing. *Journal of Memory and Language, 25,* 348-368.

Frazier, L. (1978). *On comprehending sentences: Syntactic parsing strategies,* Bloomington: Indiana University Linguistics Club.

Frazier, L., & Fodor, J.D. (1978). The sausage machine: A new two-stage parsing model. *Cognition, 13,* 187-222.

Frazier, L., & Rayner, K. (1982). Making and correcting errors during sentence comprehension. *Cognitive Psychology, 14,* 178-210.

McKoon, G., & Ratcliff, R. (1986). Inferences about predictable events. *Journal of Experimental Psychology: Learning, Memory, and Cognition,* 12, 82-91.

Osterhout, L. (1989). *Event-related brain potentials elicited during sentence comprehension.* Doctoral dissertation, Tufts University.

Osterhout, L., & Holcomb, P. (1989). *Event-related potentials elicited by garden-path sentences.* Paper presented at the Ninth Conference on Event-Related Brain Potentials, Noordwijk, The Netherlands.

Osterhout, L., & Swinney, D.A. (1989). *The effects of context on elaborative inferencing during discourse processing.* Paper presented at the Annual Meeting of the Eastern Psychological Association, Boston.

Rayner, K., Carlson, M., & Frazier, L. (1983). The interaction of syntax and semantics during sentence processing. *Journal of Verbal Learning and Verbal Behavior, 22,* 358-374.

Swinney, D. (1981). The process of language comprehension: An approach to examining issues in cognition and language. *Cognition, 10,* 307-312.

Tanenhaus, M.K., Boland, J., Garnsey, S.M., & Carlson, G. (1989). Lexical structure in parsing long-distance dependencies. *Journal of Psycholinguistic Research, 18*, 37-50.