Why is Japanese not difficult to process?: A proposal to integrate parameter setting in Universal Grammar and parsing

Reiko Mazuka
Cornell University

Barbara Lust
Cornell University

Why is Japanese not difficult to process?: A proposal to integrate parameter setting in Universal Grammar and parsing.[1]

Reiko Mazuka and Barbara Lust

Cornell University

1. Introduction

Our concern in this paper is the relation between a veridical psycholinguistic model of natural language processing and a realistic theory of Universal Grammar. In light of this concern, we do not attempt to write algorithms for particular processing routines, but to discover general principles of processing organization which may be fundamentally related to the human biologically programmed competence for language knowledge.

In particular, we are concerned with the fact that an ultimate model of natural language processing should be universal. In fact, most processing models either explicitly or implicitly aim at universality. Kimball, for example articulated this in 1 below:

(1) "I would conjecture that Processing and all those principles that follow from it deductively are universal,..." (Kimball, 1973, 40)

North East Linguistics Society, Vol. 18 [1987], Iss. 3, Art. 5

By the term 'processing', we refer in general to "a mapping from one kind of information to another" (Marr, 1980, 24). By 'language processing', we refer to "... a sequence of operations, each of which transforms a mental representation of a linguistic stimulus into a mental representation of a different form (usually more abstract in the case of perception, e.g., a representation of meaning." (Forster, 1979, 28). We will be concerned with 'natural language parsing' defined in general as *the syntactic and semantic analysis of a surface string into constituent structure in real time. The output is a bracketed sentence.*[2]

2. Statement of problem

2.1 Basic issues in human sentence parsing

It is known that natural language processing is extremely fast and efficient. For example, Marslen-Wilson (1973, 1975) demonstrated that adults (English-speaking) restored semantic anomaly in stimulus sentences while they were shadowing them very closely, viz., within 250 to 750 msc., about one syllable behind.[3] In order to achieve this speed and efficiency, all parsing models proposed to date have either implicitly or explicitly assumed that both the efficiency principles (EP), summarized in (2) below are basic to human sentence processing.

(2) Efficiency Principle (EP) 1:

The processor should utilize any information on line as soon as it becomes available. This principle implies that keeping already available information unused during on-line processing would not be very efficient and that human beings make "parsing decisions as quickly as possible'" (Abney, to appear).

Efficiency Principle (EP) 2:

Backtracking and reanalysis will be costly.

In general, these proposed EP of human on-line processing cohere with a general assumption regarding natural language processing, viz., that both the listener and the speaker are attempting to reach a semantic level of representation as quickly as possible. They also cohere with basic assumptions regarding human memory; viz., that verbatim memory is limited and costly, and "semantic memory" is therefore to be accessed as quickly as possible.

In addition, it is generally assumed as an "obvious fact" that "humans parse sentences incrementally from left to right" (e.g. Abney, 1986); and that EP 1 and 2 characterize left to right parsing procedures.

In general, for English, in instantiating EP 1 and 2, in conjunction with the left-to-right assumption, many parsers had created a tree from the top most S and worked their way down (e.g., ATN parser, Woods, 1970; Kaplan, 1972). This organization of parsing procedures is often called "top-down" parsing as in (3) below.[4]

(3) "A tree is built from an input string by starting with the initial symbol of the grammar (that which is topmost in all trees generated by the grammar) and building a tree downward to the terminal symbols. Such procedures are called top down." (Kimball, 1973, 19)

This can be done fairly efficiently in English as the main clause generally occurs first. A 'head', which determines constituent structure (both its position and its label), generally precedes the rest of the phrase. Material occurring later in time will often not affect prior parsing, due to the right-branching (RB), head-initial nature of English. Thus both EP 1 and 2 and the left-to-right assumption are maintained for "top-down" parsing organization in English.

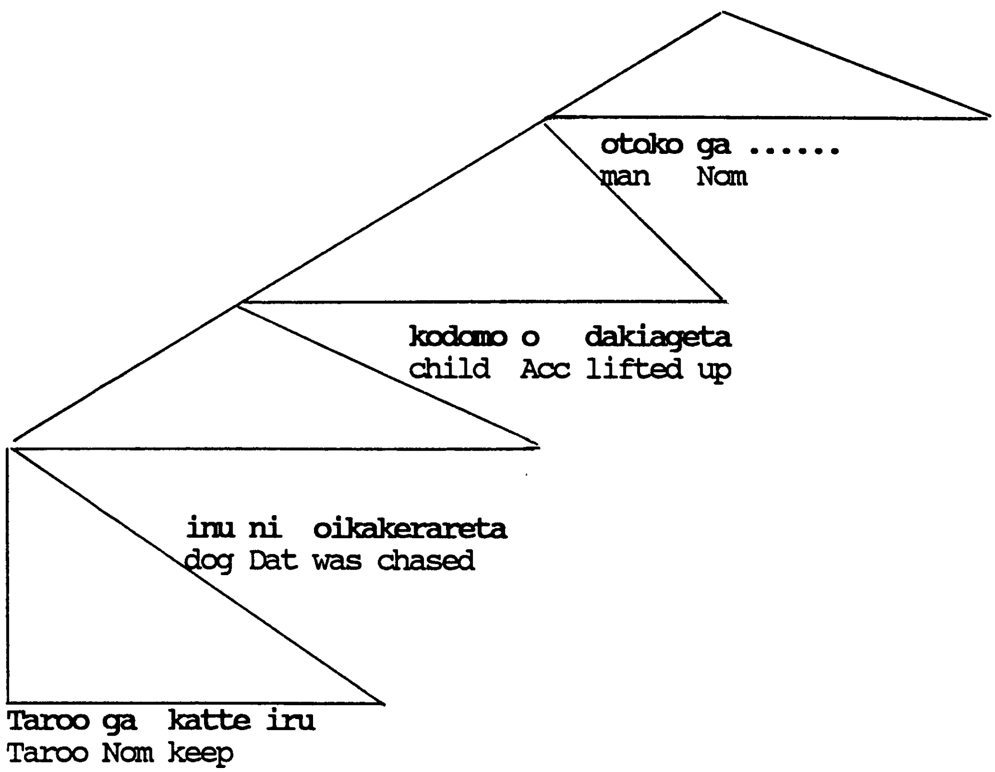
On the basis of these results in English, natural language processing has been assumed to involve a "high acceptability of RB structures" (e.g., Kimball, 1973, p.15, Yngve, 1960). Kimball (1973) wrote, for example, "Of particular interest will be those techniques which allow a string to be parsed top-down left-to-right building a phrase structure tree over the string as it is read, since this is the process employed by speakers of natural languages". (p.16). (There are areas in English where principles 1 and 2 conflict; e.g., in garden path sentences, but these are generally treated as exceptional in processing/parsing models for English (e.g., Frazier & Fodor, 1978; Marcus, 1980)).

2.2 Japanese

In Japanese, however, the linguistic facts show that if these same procedures of language parsing are applied, then this will require either extensive backtracking or possibly infinite verbatim memory. Since these results would be inconsistent with the assumed EP 1 and 2, this would mean that the same procedures which maximize processing efficiency for English do not do so for Japanese. If we assume that the parsing procedures assumed for English are universal, this would then predict the obviously false conclusion that Japanese is 'difficult to process'. We explicate this problem briefly below.

Typologically, Japanese is a SOV, left-branching (LB) and head-final language. The head of a phrase (e.g., NP, VP, AP & PP) generally comes at its end. In addition, in complex sentences, a subordinate clause precedes a main clause, and in complex NP, a relative clause precedes its head noun. Thus, when clauses are embedded recursively, the language branches out leftward as in (4) (cf. Kuno 1973).

(4)



[[$\emptyset_i$ [[$\emptyset_j$ [[Taroo ga $\emptyset_k$ katte iru]$_S$ inu$_k$ ni]$_{NP}$ oikakerareta]$_S$
name Nom keeps dog Dat was chased

kodomo$_j$ o]$_{NP}$ dakiageta]$_S$ otoko$_i$ ga]$_{NP}$ ...]
child Acc lifted up man Nom

"The man who lifted up the child who was chased by the dog Taroo keeps ..."

Japanese allows its arguments to be elliptic in not only subject, but also other positions. Japanese does not have overt relative pronouns and often no overt complementizer. Relative clauses, for example, involve a relativized noun to the right of a verb. As there is no marking on the verb, and any argument could be elliptic, a relative clause is indistinguishable from other finite clauses until the relativized noun is encountered.

2.3 Problem of Japanese parsing

The essential problem which characterizes a theory of Japanese parsing is that, in principle, Japanese is indeterminate in the amount and the type of leftward embedding it allows. Thus, when an initial item (or items) is parsed in a surface string of Japanese, it is impossible to determine how deeply, or in what way, this item is embedded under the topmost S.

This problem can be seen with simple examples where the initial surface string involves two nouns or noun phrases, as in (5). For an initial string of two NPs, for example "Hirosi ga Masao o...", an indeterminate number of parsings are possible, viz, a SOV simple sentence as in (5)a, an adverbial subordinate clause, as in (5)b, a recursive subordinate clause, as in (5)c, a relative clause, as in (5)d, a recursive relative clause, as in (5)e, or a relativized coordination, as in (5)f, a sentential subject, as in (5)g, or a recursive sentential subject as in (5)h. As the bracketings in (5) show, each of these analyses involves distinct embeddings under the topmost S. Each of these may allow infinite recursion.

Indeterminacy in Japanese also exists at the point of parsing a clause. Given a simple clause like (5)a, for example, the depth of its embedding and the structure of its embedding type are indeterminate, as the examples in (5)b-(5)h show. Simply delaying parsing decisions until a verb is encountered will not fundamentally resolve this indeterminacy, since a current clause could be recursively embedded in an indeterminable number of different ways, as the examples in 5b-h suggest:

(5) **Hirosi ga Masao o ...**
name Nom name Acc

a) [**Hirosi ga Masao o mita**]
name Nom name Acc saw
"Hirosi saw Masao"

b) [[**Hirosi ga Masao o mita**] **toki ...**]
name Nom name Acc saw when
"When Hirosi saw Masao, ..."

c) [[**Hirosi ga Masao o mita] toki Takasi ga**
name Nom name Acc saw when name Nom
hutari o yonda] node ...]
the two Acc called because
"Because Takasi called the two when Hirosi saw Masao, ..."

d) [**Hirosi ga [[Ø Masao o mita] otoko o] ...]**
name Nom name Acc saw man Acc
"Hirosi (does/did something) to the man who saw Masao"

e) [**Hirosi ga [Ø [Ø Masao o mita] otoko o] mituketa]**
name Nom name Acc saw man Acc found
tantei ni ..]
detective Dat

"Hirosi (does/did something) to the detective who found the man who saw Masao"

f) [**Hirosi ga [Ø [Ø Masao o mita] [Ø mati de**
name Nom name Acc saw town at
hurahura siteita] otoko o] mituketa] tantei ni ..]
wandering around man Acc found detective Dat

"Hirosi (does/did something to) the detective who found the man who saw Masao and was wandering around town"

g) [[[**Hirosi ga Masao o mita] koto ga]**
name Nom name Acc saw that Nom
keisatu ni mitukatta.]
police Dat was found

"That Hirosi saw Masao was found out by the police"

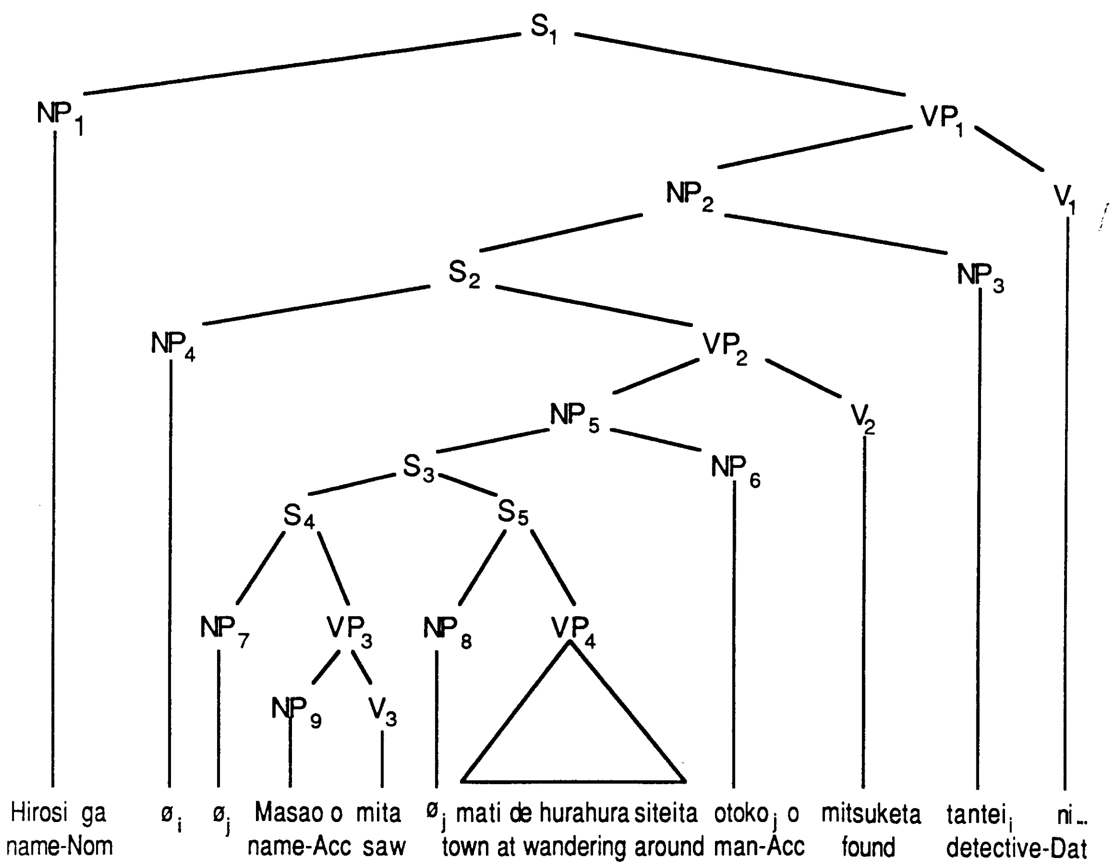
h) [[[[**Hirosi ga Masao o mita] koto ga] keisatsu ni**
name Nom name Acc saw that Nom police Dat
mitukatta] koto ga] kazoku o sinpais aseta]
was found that Nom family Acc worry made

"That the fact that Hirosi saw Masao was found out by the police made the family worry".

The head-final nature of Japanese adds further difficulty for top-down parsing procedures, in creating and labelling a new phrase node in Japanese, although this difficulty is probabalistic.[5] In general, the head of a phrase in English carries essential information both as a leading edge of a phrase (opening a bracket) and in identifying the position of the phrase with regard to the rest of the tree. For example, in the English relative clause, the 'COMP' as head of S-bar under NP is reflected in the 'wh' pronoun. Since in Japanese, the head is not available until after the phrase is completed, this critical information for phrase identification will not be available initially in on-line processing in Japanese.

The indeterminacy in Japanese left to right on-line processing is further compounded by the productive occurrence of empty categories in Japanese. For example, the tree shown in (6) represents (5)f. As can be seen from this tree, gaps can occur, both at constituent boundaries, and within constituents. Because of the possibility for infinite recursion in each of these structures, the possibility for such gaps is also in principle infinite.

6.



"Hirosi (does/did something to) the detective who found the man who saw Masao and was wandering around the town."

Due to the LB nature of Japanese, resolution of gaps by antecedent location is often only accomplished by backward anaphora, i.e., by identifying an antecedent which follows the gap, possibly indefinitely distant from it. Well-formedness of gap determination and interpretation will depend on both immediate and non-immediate dominance relations between the gap and its following antecedent. At the point of on-line processing, therefore, gap resolution is also indeterminate. In addition, if gap resolution is essential to the segmentation of a surface string into parsing units, then the productive existence of backward anaphora, given recursive LB embedding in Japanese, will also make such segmentation indeterminate.

In sum, simple application of top-down type procedures, such as may be possible for English in order to satisfy the EP 1 and 2, will fail for Japanese. They will fail to attach constituents on line to topmost S because of the indeterminacy in amount and type of leftward embedding; they will fail to initiate and label new nodes successfully because of the head final nature of the language. They will fail to account for both the location of gaps and their resolution, and thus for the on-line parsing that depends on such gap resolution.

On-line top down organization of parsing procedures in Japanese will inevitably lead to massive backtracking (thus offending EP 2) if decisions are made quickly in left-to-right order in accord with EP 1. This dilemma is not probabilistic. No amount of look-ahead (e.g., 'windows', or 'buffer cells')in the parser will solve the fundamental problem here, since the possibility for leftward recursion licensed by the LB nature of Japanese, is infinite. A model which proposes parallel processing of alternative structures/interpretations (such as proposed by Kurtzman 1984, or Gorrell 1987) will also not solve the fundamental problem here, since, as we have shown, the number and type of alternative structures/interpretations is not determinate.

Alternatively, if Japanese parsing consistently delays making decisions in order to avoid extensive backtracking, thus maintaining EP 2, this will offend EP 1; and it can require an infinite verbatim memory. Both of these possibilities, either extensive backtracking or extensive delay of parsing decisions, would thus predict that languages like Japanese should be more difficult to process than languages like English. We assume this cannot be true.

3. Possible Solutions to the Japanese Parsing Dilemma

3.1 Alternative approaches

3.1.1 Two initial alternatives

The above results require the conclusion that what constitutes "processing efficiency" in one language need not be identical to that which constitutes it in another. At first glance, there are two possible approaches to this result. One is to abandon the commitment to a universal parsing model which we began with, and to develop individual parsing models for specific languages. Another is to maintain the commitment to a universal parsing model, but to revise current models for English so that they are compatible with the Japanese facts.

The first approach is undesirable for several obvious reasons. (i) It would render parsing models basically descriptive and unconnected with UG (ii) It would accentuate the first language acquisition problem: viz., it would propose that the child must develop specific processing procedures differently dependent on which language it is acquiring; thus it is not clear how the child could initially process the evidence available to it in language acquisition. (iii) Occam's razor. Unless there is no other solution, then, we do not adopt this first approach, and maintain our commitment to the development of a universal parsing model.

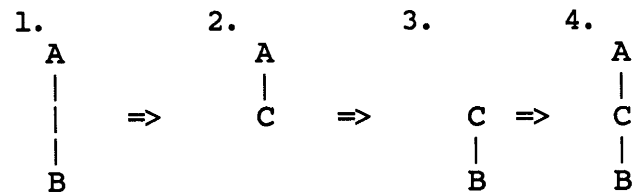
The alternative approach, i.e., maintaining the commitment to a universal parsing model by revising current assumptions regarding English processing, is actually being taken by several scholars in the field today. Two examples of such approaches are discussed briefly here. We suggest that both ultimately result in a prediction for greater processing difficulty in Japanese.

3.1.2 D-theory

Marcus and Hindle's, (Marcus, Hindle, & Fleck,1983; Marcus & Hindle, 1986), "D-theory" attempts to solve the problem of indeterminacy we have described by postulating "pseudo attachments" of nodes during on-line processing.

In this model, the parser's task is to Describe a tree, rather than build one. The parser computes dominance relations rather than immediate dominance relations. Thus, after it has been asserted that A dominates B, other dominance relations such as A dominates C and C dominates B can be inserted without reanalyzing the original dominance relation between A and B.

(7)



Because on-line pseudo-attachments are tentative and nonspecific, this theory allows subsequent insertion of additional nodes between mother and daughter nodes.

This 'D-theory,' however, essentially relies on the assumption that inserting additional dominance relations between a mother and daughter node is equivalent to the initial assertion of a dominance relation. This is equivalent to abandoning that part of EP 2 which says that 'reanalysis is costly'. If an insertion of a node between nodes whose dominance relations have already been drafted was even slightly more costly than asserting dominance relation without insertion, Japanese then would be harder to process than English since LB structures will require more such insertion than RB structures will.

Also, in the 'D-theory' model, parsing actions are triggered by matching linear templates with the content of a three cell buffer. Probabilistically, identification of a phrase should be done more accurately when a head initiates a phrase. Therefore, a head initial language such as English should have an advantage over a head final language such as Japanese.

Finally, as a consequence of abandoning EP 2, parsing strategies proposed by the D-theory do not themselves make predictions for perceived psychological differences in accessibility of processing right or LB structures, or garden path phenomena within a language. Such psychological phenomena, which have been traditionally viewed as resulting from proposed parsing strategies, must now be accounted for by an independent theory. We assume that a theory which can account for both linguistic and psycholinguistic phenomena is stronger than one that can not.

3.1.3 Licensing model

Similarly, it might appear that a 'licensing' model might resolve the dilemma we have identified, since this model is intended to be order free. In this approach, the syntactic structure which the parser recovers is represented as a licensing structure. "Every element in the structure is licensed by performing a particular function in the structure; the structure is well formed only if every element in it is licensed." (Abney, 1). Instead of computing a phrase structure directly from phrase structure rules, well-formedness of the structures derived from various principles of the grammar are computed.

As a tree is not directly computed in this parser, it cannot be said that the tree is built top down or bottom up. However, when we examine the direction of licensing in relation to left to right parsing in this model, we find that this parsing model also predicts that a LB (head final language) such as Japanese will be more difficult to parse than English. This is because it is the lexical head which is the licenser. Thus, if the head comes last, the licensee must be kept in the buffer until it can be licensed by the head. In cases where head and complement are not necessarily adjacent, e.g., where a verb is said to 'license' an adverb, the problems we raise in 2.3 will also apply to this model.[6] Compared to head-initial languages like English, Japanese will then involve a heavier processing load in this model also, if this formal model relates to psycholinguistic phenomena. Abney and Cole propose that it does: ". . .the number of processes which are waiting on right context at any point in the parse gives a rough measure of the psychological complexity of the parse" (1985, 4).

Thus, the parsing 'problems' we have identified hold whether a parser actually computes individual phrase structure rules or merely uses the head of phrases to 'license' subsequent structure.[7]

3.2 Proposed solution

In our current approach to the problem we raise in 2.3, we hypothesize that we can maintain universality of a parsing model, by considering parsing for a language to be organized in a principled way with regard to Universal Grammar. In 8, we formalize this possibility in one way by considering that parsing of a specific language may be parameterized as a deductive consequence of parameter setting of Universal Grammar for that language. By this, we mean that although there may be a universal set of common procedural components for a parser, these components may be assembled or organized differently for different languages in a way which both reflects parameter setting in UG and allows maximal processing efficiency for specific languages, i.e., maintains the two EP above.

(8) The organization of parsing procedures for a specific language may be parameterized as a deductive consequence of parameter setting in Universal Grammar.

Specifically, we assume that the linguistic differences discussed above between Japanese and English can be explained in terms of a binary valued grammatical parameter in UG, e.g., 'head direction' or 'branching direction' as has been proposed in the literature (e.g., Huang 1982, Chomsky, 1982, Lust, 1983, in preparation). We see in 2.3 above that the LB and/or head-final properties of Japanese in conjunction with assumed left-right on line processing, are one source of the apparent massive indeterminacy involved in Japanese processing. The LB or head final parameter-setting for Japanese is thus implicated in these effects.[8]

We hypothesize that such principled grammatical differences between LB and RB languages may lead to principled differences in the organization of procedures of parsing. In particular, they may lead to a choice between principally "top-down" or "bottom-up" organization of parsing procedures for RB and LB languages respectively.

We use the term "bottom-up" as in (9).

(9) **Definition of bottom-up parsing organization**

In general: "The first action of such a parser is to assign the first _m_ input symbols to some node, which is then placed at the top of a stack." (Kimball, 1973, 19)

Specifically in our proposal: In parsing organized according to 'bottom-up' principles, hypotheses about constituent structure of lower constituents are sequenced before hypotheses about the relation of such constituents to higher constituents.

In particular by this type of parsing organization, on line parsing of individual constituents is not determined by their relation to the topmost S.

If we posit principally bottom-up parsing procedures (in this sense) for LB languages, then on-line processing of a left recursive structure would not force the parser to make an initial decision as to how it should be attached to the topmost S. A bottom up principle of organization would thus enable the parser to avoid massive backtracking and reanalysis in these languages. Parameterization would link this organization of processing specifically to LB (RH), but not RB (LH) languages where such procedural organization would not be maximally efficient.

3.3 _Empirical predictions of this proposal_

At present, this proposal for parameterized parsing, in accord with parameter setting in UG, stands as a hypothesis. It makes several predictions, however, in several domains.

3.3.1 Psycholinguistic evidence.

Several studies of English processing have demonstrated differential accessibility of main and subordinate clauses during on-line processing (e.g., Kornfeld, 1973; Bever & Townsend, 1979; Kurtzman, 1985; Von Eckardt & Potter, 1985). These results are consistent with our hypothesis. If the English speaker is processing clauses top-down, s/he can identify the type of the clause as main or subordinate by its relation to the topmost S. However, a priori, the same cannot be true for the Japanese speaker, as the LB (right-headed) nature of the language does not permit the Japanese speaker to differentiate types of clauses until their very end. Therefore, this initial experimental evidence in English is consistent with our hypothesis in that it would argue that on-line processing procedures must be differentially organized in Japanese and English. This hypothesis is being further tested experimentally in Japanese.

In addition, if processing procedures are organized differentially across languages, that is, if they are organized in accord with their grammatical parameter setting, then this proposal makes a further critical prediction. Namely, for a particular structure which may occur in either of two languages which differ in their parameter setting as either RB or LB, then the on-line processing of this same structure should differ across these languages. For example, a LB structure may occur in RB languages like English, just as certain RB structures may occur in LB languages. Our proposal predicts that the processing of such LB structures will differ in principle dependent on whether the language in which they occur is RB or LB. In RB languages like English, they may cause additional complexity while in LB languages like Japanese, they may not.

Previous psycholinguistic studies of English, have in fact argued that in many ways language processing of LB structures in English is more complex and less efficient than processing of RB structures (Smith & MacMahon, 1970; Bever, 1970; Bever & Townsend, 1979; Levin, Grossman, Kaplan, & Yang, 1972; Kornfeld, 1973.)

We also know that semantically complex relations which can be conveyed with complete facility as LB in one language, can only be conveyed with complete facility as RB in another, even when the specific grammar of the language actually allows LB in both. For example, as in (10) below, the natural Japanese sentence involving LB recursive sentential subjects in (10)a, is blocked in the parallel LB English structure in (10)b or (10)c, even though the specific language grammar for English does allow sentential subject embedding. In English, the representation corresponding to (10)a must be RB as in (10)d.

(10)

a. [[[[gakko o kiratte iru koto ga] titioya ni sirareta
school Acc hate that Nom father Dat was known

koto ga] John o hidoku okoraseta
that Nom name Acc very made angry

koto ga] hahaoya o yuutu ni saseta]
that Nom mother Acc depressed

b. */? [[[[That John hated school] was known to his father] made him very angry] distressed his mother]

c. * [That[that[that John hated school] was known to his father] made him very angry] distressed his father]

d. [His mother was distressed [[that John was angry] [because his father knew [that he hated school]]]

These initial results are consistent with our critical hypothesis.

3.3.2 Linguistic evidence

In certain cases, specific languages may have grammaticized certain phenomena related to the parameterized parsing we have hypothesized here. For example, it is well known that in English, a clause initial 'that' complementizer for finite sentences may not be omitted in LB structures, e.g., (11)b, although it may be in RB structures as in (11)c.

(11) a [[That she was smart] was clear]
b * [[Ø She was smart] was clear]
c [It was clear [that/Ø she was smart]]

If organization of processing procedures for RB languages like English favors initial on-line determination of the relation of a new node to the main clause S, this information would be lacking in LB structures without an overt complementizer, but not in RB structures. If processing organization is differentiated as we have hypothesized, this would favor initial complementizer marking for LB structures in English, although not necessarily in RB structures. (As we have noted above, such complementizers are often not overt in LB languages and are clause final, when they do occur.)

The presence and absence of WH-movement in English and Japanese respectively might also be related to the different organization of processing procedures which we have hypothesized across these languages. In top down processing, a fronted WH-comp plays a significant role in identifying type of clause as well as in indicating the existence of a gap in the following clause. On the other hand, as bottom-up parsing procedures do not expect such information at the beginning of a the clause, it is not as efficacious to have a WH-word fronted in a LB language such as Japanese, as in a RB language.[9]

3.3.3 Acquisition evidence

The hypothesis we have proposed states that in first language acquisition, one deductive consequence of grammatical parameter setting involves the establishment of principles for parsing organization.

We do know from previous experimental studies of first language acquisition of English and Japanese, that very young Japanese children find LB structures significantly more accessible than RB structures in language production and comprehension, while English children demonstrate the reverse (e.g., Lust, Wakayama, Snyder & Mazuka, in preparation; Lust & Mazuka, 1986). Japanese children find backward anaphora to be natural in the LB structures (Lust et al., in preparation; Lust & Mazuka, 1986). However, if parameter setting has consequences for actual on-line language processing as predicted by our hypothesis, and if this differential processing is an essential and deductive consequence of grammatical parameter setting, then we predict that Japanese and English children would demonstrate adult-like on-line processing differences from early stages. This prediction is being experimentally tested now (Mazuka, in preparation).

3.3.4 Relation between parsing and language acquisition

Our proposal may help to resolve a fundamental paradox of current linguistic theory as a theory of the Initial State (cf. Fodor, Bever & Garrett, 1974, Valian 1986 for example). Namely, if parsing primary language data (PLD) depends on specific grammatical information, then how is it possible for the child to parse PLD if it does not know this grammar? But if it knows the grammar first, how can the child then be using the data in order to acquire its language?

By our proposal, UG is innately programmed, and initial PLD processing is linked to UG, not to specific language grammars. In addition, a universal 'parser' in the sense of a possibly unordered set of mechanisms or procedures for speech and language processing may be universally innately predetermined. This 'parser' is innately intrinsically linked to UG. The procedures which the parser makes available, however, are not initially organized until parameter-setting in UG occurs and determines this organization. Acquisition of specific language grammar will thus cohere with acquisition of processing efficiency in non-circular fashion.

It might appear that one paradox has been replaced by another here. If P-setting is required for parsing, then how can P-setting ever occur unless data to set it by is processed. Yet if such data can be processed without P-setting, this would appear to vitiate our proposal. Our proposal acknowledges, however, that some form of individual procedures or mechanisms for processing all possible structures in all possible languages may pre-exist. A child must be born with the ability to effectively parse both types of structure corresponding to both values on a binary parameter. What we do suggest is that these procedures are not systematically organized at first, i.e., that processing efficiency does not exist until these various procedures are organized in accord with the UG-determined organization of this language.

This aspect of our proposal is necessitated also by adult grammar. Languages which instantiate a value in parameter setting (e.g., which are 'principally' LB, or 'right-headed') will very often also allow RB or left-headed structures, just as RB languages may allow the reverse. The theory of 'parameter setting' does not entail that movements, scrambling or markedness of some form may not occur within a language to provide structures which do not cohere with the unmarked parameter setting value. Our proposal does not predict that adults cannot in some way process the structures which do not cohere directly with the parameter-setting of their language. It does predict, however, that processing of these alternative structures is not maximized by the general organization of processing which has been derived for that language as a deductive consequence of its grammatical parameter-setting. Thus our proposal predicts the psycholinguistic processing phenomena which we discuss above.

4 Conclusion

In conclusion, we have argued that Japanese (and other languages like it in LB structure) are not parsable by models developed to date for English without loss of efficiency, and without the false conclusion that 'Japanese should be more difficult to process'. We argue that this problem holds even in the case of more current 'principle-based' parsers. We have identified one source of this dilemma in the LB property of such languages. We have hypothesized as one possible solution to this dilemma that the organization of parsing procedures may be parameterized. One possible parameterization may lie in a branching direction (or head direction) parameter of UG.

We have proposed a model in which a universal set of parsing procedures may be organized differently across languages, in accord with the grammatical parameter-setting of those languages. Principles of such procedural organization, e.g., whether or not they are ordered to prioritize 'bottom-up' procedures are deductively determined by parameter-setting of Universal Grammar.

If this proposal is confirmed, it would strengthen the theory of UG as a realistic theory of the human mind. Here, UG would have consequences not only on the structure of language competence but on the systematic organization of language performance. This proposal would contrast with models which propose 'functional' explanations for certain aspects of language structure. In this proposal, the grammar itself, that is UG, in part would explain certain properties of processing procedures.[10]

It would also strengthen the theory of UG as a model of the Initial State. In general, if our hypothesis is confirmed, this would suggest that in first language acquisition, early data representation, and consequently induction from it, is to a significant degree determined by deduction from a child's grammatical knowledge. This would question the very distinction between induction and deduction often assumed in the language acquisition literature.

5. Addendum

Since our initial construction of this paper, it has come to our attention that other scholars have independently proposed that parsing parameterization may be linked to the branching direction parameter, although the specifics of these proposals differ from ours, *cf*., Frazier and Rayner, to appear.

Several studies in computer science in Japan (Matsumoto et al., 1983; Takakura, 1986; Okumura, 1986) as well as in Taiwan (Lin, 1985; Lin et al., 1986a, 1986b) independently propose that bottom up parsing procedures are more efficient for parsing LB languages such as Japanese and Chinese.

We also thank Robert Berwick for drawing our attention to several 'parallel' models in the current computational literature which directly confront the left recursion processing problems we raise in this paper. The linguistic, psycholinguistic and learnability predictions of these models must now be studied in relation to our proposal in this paper.

FOOTNOTES

1 This paper is partially supported by NSF dissertation grant #3312 to Barbara Lust for Reiko Mazuka; and to a Cornell College of Human Ecology small grant to Reiko Mazuka. We thank Howard Kurzman, James Gair, James Huang, John Bowers, John Kingston, Marion Potts, and Jennifer Cole for discussion of the issues raised in this paper.

2 We do not assume that phrase structure is generated by phrasal rules, only that it must exist as ultimate parsing output. In particular, it is necessary for execution of structural constraints and principles relevant to anaphora.

3 Frazier (1980) refers, for example, to "how very quickly the parser makes even higher order decisions about the analysis of linguistic material" (178).

4 We distinguish the terms "top down" and "bottom up" parsing from "hypothesis driven" and "data driven" parsing. In the former, we refer to the way a tree is built, either from higher node downward or lower node upward.

5 As opposed to the logical impossibility of determining the depth of embedding in LB structures such as in 5, it is not impossible for a top down procedure to determine the label of a phrase before a phrasal head is reached. In fact, Marcus and Hindle's "D-theory" parser (which we discuss below) attempts to determine a phrase label by examining three items in a buffer. Whether the leading edge of a particular phrase can be identified before the head is reached in this way is an empirical issue. However, as it is apparent that the existence of a head at the beginning of a phrase makes identification of the phrase more quickly, given left-right parsing, the head final nature of Japanese also makes the prediction that, overall, Japanese should be more difficult to process than English.

6 It is not clear how extensions of $\overline{X}$ theory to non-lexical heads (e.g., Chomsky, 1986) will be captured in this model, nor how adjunct clauses will be.

7 Similar issues will characgerize an $\overline{X}$ theory-based proposal like that of Berwick and Weinberg 1985, which initially 'underdetermines' structure and feature assignments.

8 Precise formulation of the parameter at issue here is currently under study. It is generally assumed that a "head direction" parameter or a "branching direction" parameter should be able to represent the relevant linguistic variation (or an alternative formulation). However, as

suggested above, branching direction and head direction contribute qualitatively different types of processing difficulty for Japanese and it is the LB property that is the source of the essential problem we have identified. These facts appear to evidence an independence of the two parameters.

Notice too that it is not impossible that other attempts to formalize essential parametric differences across English and Japanese (e.g., Kuroda, 1985, or Fukui, 1986)underlie the essential grammatical distinctions includng the ones we have identified above. In this case, given our hypothesis in this paper, it may be possible to rationalize the relation of these alternative parameter formulations to processing phenomena such as we have discussed.

9 Certain grammatical options such as heavy NP shift or extraposition, which appear to occur in RB languages, but not in LB languages may also be motivated by such parametric differences in processing organization as we have hypothesized.

10 In turn, these different organizations of processing procedures may then have specific language grammatical consequences.

BIBLIOGRAPHY

Abney, S. (1986). Functional elements and licensing. Paper presented to GLOW, Gerona Spain.

Abney, S. (to appear). Licensing and parsing. S. Berman, J.-W. Choe, & McDonough (Eds.), Proceedings of NELS, 17, Amherst, MA: University of Massachusetts at Amherst.

Abney, S., & Cole, J. (1985). A Government-Binding parser. In S. Berman, J.-W. Choe, & McDonough (Eds.), Proceedings of NELS, 16, 1-17, Amherst, MA: University of Massachusetts at Amherst.

Berwick, R., & Weinberg, A. (1985). Deterministic parsing and linguistic explanation. Language and Cognitive Processes, 1, 109-134.

Berwick, R., & Weinberg, A. (1985). Deterministic parsing: A modern view. In Berman, S., Choe, J-W., & McDonough, J. (Eds.), Proceedings of NELS 15. Amherst, MA: University of Massachusetts at Amherst.

Bever, T. (1970). The cognitive basis for linguistic structures. In J. R. Hayes (Ed.), Cognition and the Development of Language. New York: John Whiley and Sons.

Bever, T., & Townsend, D. (1979). Perceptual mechanisms and formal properties of main and subordinate clauses. In W. Cooper and E. Walker (Eds.), Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett, 159-226, Hillsdale, NJ: Lawrence Earlbaum Associates.

Chomsky, N. (1982). Some Conceptual and Consequences of the Theory of Government and Binding. Cambridge, MA: MIT Press.

Chomsky, N. (1986). Barriers. Cambridge, MA: MIT Press.

Fodor, J., Bever, T. and Garrett, M. (1974). The Psychology of Language. New York: McGraw Hill.

Forster, K. I. (1979). Level of processing and the structure of the language processor. In W. Cooper and E. Walker (Eds.), Sentence Processing: Psycholinguistic Studies Presented to Merrill Garrett, 27-85, Hillsdale, NJ: Lawrence Earlbaum Associates.

Frazier, L. (1980). Parsing and Constraints on Word Order. In J. Lowenstamm (Ed.), University of Massachusetts Occasional Papers in Linguistics, 177-198.

Frazier, L., & Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. Cognition, 6, 291-325.

Frazier, L., & Rayner, K. (to appear) Parameterizing the language processing system: Left- vs. right-branching within and across languages. In Hawkins, J. (Ed.), Explaining language universals. Oxford: Basil Blackwell.

Fukui, N. (1986). A theory of category projection and its application. Unpublished Ph.D. dissertation. MIT.

Gorrell, P. (1987). Theoretical and methodological issues in natural language processing. Unpublished Ph.D. Dissertation, University of Connecticut.

Huang, J. (1982). Logical relations in Chinese and the theory of grammar. Unpublished Ph.D. Dissertation. MIT.

Kaplan, R. M. (1972). Augmented transition networks as psychological models of sentence comprehension. Artificial Intelligence, 3, 77.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. Cognition, 2, 15-47.

Kornfeld, J. (1973). The influence of clause structure on the perceptual analysis of sentences. unpublished PhD dissertation, MIT.

Kuno, S. (1973). The structure of the Japanese language. Cambridge, Massachusetts: MIT Press.

Kuno, S. (1974). The position of relative clauses and conjunctions. Linguistic Inquiry, 1, 117-136.

Kuroda, S.-Y. (1985). Whether you agree or not: Rough ideas about the comparative grammar of English and Japanese. Unpublished manuscript. San Diego: University of California.

Kurtzman, H. (1984/1985). Studies in Syntactic Ambiguity Resolution, Ph.D. Dissertation, MIT (Bloomington, IN: Indiana University Linguistics Club).

Levin, H., Grossman, J., Kaplan. E., & Yang, R. (1972). Constraints on the eye-voice span in right and left embedded sentences. Language and Speech, 15, 30-39.

Lin, L.-J. (1985). A syntactic analysis system for Chinese sentences. Unpublished MA thesis, Taiwan: Taiwan National

Lin, L.-J. (1985). A syntactic analysis system for Chinese sentences. Unpublished MA thesis, Taiwan: Taiwan National University.

Lin, L-J, Huang, J., Chen, K-J, & Lee, L-S. (1986a). A Chinese natural language processing system based upon the theory of empty category. Proceedings of 5th National Conference on Artificial Intelligence of American Association of Artificial Intelligence. pp. 1059-1062.

Lin, L-J, Chen, K-J, Huang, J. & Lee, L-S. (1986b). SASC: A syntactic system for Chinese sentences. Proceedings of 1986 International Conference on Chinese Computing. (Sponsored by Chinese Language Computer Society.)

Lust, B. (1983). On the notion "Principle Branching Direction", a parameter of Universal Grammar. In Y. Otsu, H. VanRiemsdijk, K. Inoue, A. Kamio, & N. Kawasaki (Eds.), Studies in generative grammar and language acquisition, 137-151, Tokyo: International Christian University.

Lust, B., & Mazuka, R. (1986). Universals in First Language Acquisition: Grammatical Competence on Processing Strategies. Paper presented at Biannual Conference of Society for Research in Child Development. Baltimore.

Lust, B., Wakayama, T., Snyder, W., Mazuka, R., & Oshima, S. (in preparation). On the first language acquisition of null anaphora in Japanese. Unpublished manuscript Cornell University.

Lust, B. (in preparation). Principles of first language acquisition: Cross-linguistic study of pronouns and their antecedents.

Marcus, M. (1980). A Theory of Syntactic Recognition for Natural Language. MA: MIT.

Marcus, M., & Hindle, D. (1986). Manuscript handed out in class at the 1986 LSA Summer Institute. CUNY/Bell Laboratory.

Marcus, M., Hindle, D., & Fleck, M. (1983). D-theory: Talking about talking about trees. Proceedings of the 21st meeting of the Association of Computational Linguistics, 129-136.

Marr, D. (1980). Vision, New York: W.H. Freeman and Company.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. Nature, 244, 522-523.

Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. Science, 189, 226-228.

Matsumoto, Y., et al. (1983). BUP: A bottom-up parser embedded in Prolog. New Generation Computing, 1 (2), 145-158.

Mazuka, R. (in preparation). Parameterized parsing: Consequences for first language acquisition of Japanese and English. Ph.D. Dissertation, Cornell University.

Mazuka, R. (in preparation). Processing differences in Japanese and English speaking children's first language acquisition: Consequences of parameter-setting. Ph.D. dissertation, Cornell University.

Mazuka, R., & Lust, B. (in preparation). A cross-linguistic test of the main clause effect in young children's syntactic processing: A study of Japanese. Paper presented at Boston University Conference on Child Language. 1985.

Okumura, M. (1986). Nihongo rikai system ni kansuru kisoteki kenkyu. (Basic research on Japanese comprehension system). PROLOG ni yoru shizen gengo shori gijutsu ni kansuru kisoteki kenkyu [I]. (Basic research in natural language processing technics using PROLOG). (In Japanese). Tokyo Kogyo Daigaku Johokogakuka, Tanaka Kenkyushitsu Kenkyuseika TR-1 (Technical Report TR-1, Tanaka Research Lab., Dept. of Information Technology, Tokyo Institute of Technology).

Smith, K., & McMahon, L. (1970). Understanding order information in sentences: Some recent work at Bell Laboratories. In G.B. Flores d'Arcais and W.J.M. Levelt (Eds.), Advances in Psycholinguistics, North-Holland, Amsterdam.

Takakura, S. (1986). Einichi kikai honyaku shisutemu ni okeru henkan jisho no jido sakusei ni kansuru kisoteki kenkyu (Basic research on automatic generation of a translation dictionary for an English-Japanese machine translation system). PROLOG ni yoru shizen gengo shori gijutsu ni kansuru kisoteki kenkyu [I] (Basic research in natural language processing technics using PROLOG). (In Japanese) Tokyo Kogyo Daigaku Johokogakuka, Tanaka Kenkyushitsu Kenkyuseika TR-1 (Technical Report TR-1, Tanaka Research Lab., Dept. of Information Technology, Tokyo Institute of Technology).

Valian, V. (1986). Parsing and Positive Evidence in Language Acquisition. Unpublished manuscript. Wellesley College.

Von Eckhardt, B., & Potter, M. (1988). Clauses and the semantic representation of words. Memory and Cognition, 13, 371-376.

Woods, W. (1970). Transition network grammars for natural language analysis. Communications of the ACM, 13, 591-606.

Yngve, V. (1960). A model and a hypothesis for language structure. Proceedings of the Americal Philosophical Society, 104, 444-466.