Incremental Pre-head Attachment in Japanese Parsing

Yuki Kamide
Department of Psychology, University of York, UK

Don C. Mitchell
School of Psychology, University of Exeter, UK

The present study addresses the question of whether structural analyses of verb-arguments are postponed up until the head verb has been processed (head-driven parsing accounts) or initiated prior to the appearance of the verb (pre-head attachment accounts). To explore this question in relation to a head-final language, a Japanese dative argument attachment ambiguity was examined in both a questionnaire study (Experiment 1) and a self-paced reading test (Experiment 2). The data suggested that the dative argument attachment ambiguity is resolved in the manner predicted by pre-head attachment accounts. The results were incompatible with most variants of the head-driven parsing model, and were not of the form currently predicted by constraint-satisfaction models. We end by discussing the general theoretical implications of the findings.

INTRODUCTION

Head-driven Parsing

The question of what types of information guide initial parsing decisions has been the focus of a great deal of attention in the psycholinguistic literature. Factors that have been examined include the effects of proximity and structural simplicity (Frazier, 1987), the effects of discourse context (Altmann & Steedman, 1988), effects of intonation and prosody.
(Fodor, 1998), effects of prior exposure to language (Cuetos, Mitchell, & Corley, 1996), and the effects of detailed lexical information (Ford, Bresnan, & Kaplan, 1982; MacDonald, Pearlmutter, & Seidenberg, 1994; Trueswell & Tanenhaus, 1994).

The present paper focuses on the last of these classes of information that potentially influence the course of parsing. In particular, it is primarily concerned with the role of “head information” in the early stage of parsing—an issue that has previously been explored by numerous investigators (Ford et al., 1982; Pritchett, 1991, 1992; MacDonald et al., 1994; Trueswell & Tanenhaus, 1994; Konieczny, Hemforth, Scheepers, & Strube, 1997). The “head” of a phrase or clause can be characterised as the “core” word in the constituent. For example, the head of a noun phrase (NP) (e.g., “the pretty cat”) is the noun itself (“cat”, in this example). Similarly, the head of a verb phrase (VP) (e.g., “(The pretty cat) likes cheese very much”) is the verb (“likes”). According to certain linguistic theories (Chomsky, 1981; Fillmore, 1968; Jackendoff, 1972, Grimshaw, 1990; Pollard and Sag, 1987, 1994), each and every head is associated with certain types of item-specific information (“head information”) that dictates or at least constrains the kinds of relationship it can enter into with other constituents within the sentence. If the head is a verb, for instance, this head information might include a list of permissible “argument structures” of the head (i.e., types of constituent it can be interlinked with). As an illustration, the verb “put” in Sentence (1) links with three structures.

(1) The man put the book onto the table.

One of these structures represents a person doing the “putting”, a second, the object being moved and the third the destination of the action (see Fillmore, 1968; Grimshaw, 1990; MacDonald et al., 1994, for detailed analysis of verb-argument structures).

A central concern of recent psycholinguistic research has been to assess the importance of such head information in parsing relative to other possible types of information or constraints. One, admittedly extreme, possibility is that the human parser makes use of such information to the exclusion of all other potential sources of guidance.

A system that is constrained by grammatical information stored with only heads may be referred to as a “head-driven” parser. Pritchett (1991) has proposed such a parser (see also Pritchett, 1987, 1988, 1992). This system is governed by Principle (2) as given below:

(2) A node cannot be projected before the occurrence of its head, since the relevant features which determine its categorical identity and
license both its own and its arguments’ attachment are theretofore undetermined (Pritchett, 1991; p.252).

Hence the crucial characteristic of head-driven parsing is that the parser does not make attachment decisions prior to a head. Thus, in sentences like (1), the claim is that the first Noun Phrase (NP) (“The man”) temporarily remains unattached to any other linguistic structure awaiting the arrival of the head-verb “put”. Once the verb materialises it provides “docking points” not only for the free-floating NP but also for the two other constructions to follow. Furthermore, based on the Theta Criterion (Chomsky, 1981, 1986), Pritchett (1991, 1992) argues that the head-driven parser is dominated by information stored with heads in attachment decision-making during the course of parsing. This is represented as the “Generalized Theta Attachment” principle in Pritchett (1992):

(3) Generalized Theta Attachment:
Every principle of the Syntax attempts to be maximally satisfied at every point during processing (Pritchett, 1992; p.138).

Note, however, that immediate satisfaction of licensing requirements (i.e., GTA) is not a mandatory property of a head-driven parser. For instance, one could posit a parser (a) that does not make attachment decisions until a potential head appears and (b) that does not follow the GTA principle at the head. Such a parser could still be referred to as head-driven. Thus, not every possible head-driven parser has to be governed by licensing information. On the other hand, it is also important to note here that Pritchett’s theory is rather radical compared with other licensing parsing theories (Abney, 1989; Gorrell, 1995; Weinberg, 1995) and also with constraint-satisfaction (or constraint-based) models (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994), in the sense that Pritchett (1991, 1992) rules out the use of any information other than that stored with heads, typically licensing relations. In contrast, the other theories mentioned above posit the use of several different kinds of information often on a weighted parallel basis (e.g., head information, discourse information for constraint-satisfaction models). Alternatively, they use a hierarchy of prioritised types of information (e.g., complementary attachment principles (the Verb Attachment and Low Attachment principles) are included in the account presented by Abney, 1989). Hereafter, however, we shall concentrate our discussion of head-driven parsing on parsers with both head-driven and licensing-driven properties, namely, parsers (a) that make no attachment decision until the head arrives and (b) that are governed exclusively by the premise that they have to fill licensing requirements maximally (i.e., whenever possible) as soon as they can (e.g., the head-driven parser proposed by Pritchett, 1991). Note
that such parsers attempt to fill even optional licensing requirements (e.g., optional thematic roles) in order to accomplish maximal licensing.

Some crucial features of a strictly head-driven account are set out clearly in a computational model of parsing developed by Vosse and Kempen (submitted). In this account (the Unification-Space, or U-Space, model) the arrival of successive words of a sentence causes their associated “lexical frames” to be copied in turn into a working space (“Unification Space”) that is then used as a medium for computing syntactic links between frames. Upon entering U-Space, each form is immediately linked with others already present, using sets of connections whose strengths vary over time depending on a complex set of inhibitions, and constraints designed to ensure that the links conform to a variety of types of semantic and syntactic agreement. The details of the “unification” or constraint-satisfaction procedure are critical to the workings of the implementation, but need not be spelt out for the present purpose—which is to outline an explicit head-driven model.

In the U-Space model, the structured representation of (part of) a sentence comes exclusively from (a) the lexical frames copied into U-Space as the successive words arrive and (b) the computed links between these frames. Crucially, the top-level S node is construed within the theory as being part of the lexical form for verbs (and verbs alone). It follows that before the verb arrives, nothing at all can be linked to the S node (or its constituent V node)—for the simple reason that neither yet exists at this point within the system. Because of this, the structuring of the NP must await the arrival of the verb itself. When the verb does appear and its frame joins the others in U-Space, its numerous unification constraints can finally exert their influences on the structural linkages. In a later section of the paper we shall use the explicit nature of the Vosse/Kempen statement to derive predictions for strictly head-driven parsing models of this kind. The predictions are particularly stark for head-final languages (like Japanese and Korean), but important, too, in languages with head-final subconstituents (e.g., Dutch and German). For example, consider the following Japanese sentence (Sentence 4):

(4) Chika-ga Kayo-ni koneko-o ageta.
   Chika-Nom Kayo-Dat kitten-Acc gave.
   “Chika gave a kitten to Kayo.”

This sentence features three pre-verbal arguments: “Chika-ga” (Chika-Nom), “Kayo-ni” (Kayo-Dat) and “koneko-o” (kitten-Acc). According to strictly head-driven models, the three arguments are left unattached until the head verb “ageta” appears at the end of the sentence. Hence the model posits the steady accumulation of structurally unattached constituents as the reader or listener progresses through the sentence. A system of
this kind would be expected to impose considerable demands on short-term memory.

Pre-head Attachment

Such memory demands could be held in check if, contrary to the assumptions of head-driven parsing, structural analysis is initiated before the head is heard or read. In the remainder of the present paper, we shall refer to such devices as “pre-head attachment parsers”. Pre-head attachment accounts contrast with head-driven parsers in that they impose fewer constraints on the starting point for computing structural relations. In the Japanese sentence (4) above, for example, arguments might be assigned to the as-yet-unprocessed verb as soon as each argument phrase is encountered in the sentence. Such incremental pre-head attachment parsing is premised on the assumption that representational features typically associated with heads (e.g., VPs for verbs) can be postulated before the head words themselves appear.

There are numerous different ways in which such models can be formulated. For instance, borrowing the theoretical language of the Vosse/Kempen U-Space model, one could posit a parsing system which, unlike the “official” version of the theory, allows certain kinds of non-lexical frames to be inserted into unification space in the course of sentence processing. In particular, a pre-head variant of the U-Space model might allow (say) a “generic sentence-structure frame” to be injected either just before or just after the first lexical frame is loaded. A “generic sentence-structure frame” is envisaged as being something like a standard verb-frame but lacking the lowest level information which instantiates the frame as the product of a particular verb. If this new, non-lexical frame were handled in the same way as the standard bottom-up units, this would open the way for the establishment of NP-V node links immediately after the arrival of Word 1. For example, an early NP might be marked as a potential argument for a verb slot even though the verb destined to occupy that slot had not yet made its appearance. In such a variant of pre-head attachment accounts, the exact trajectory of the build-up of link strengths would depend on the precise assumptions made about the “generic frame” and its inherent constraints. However, the broad description above is sufficient to give a flavour of models of this kind. There are several different theories that share the premise that verb-centred processes can precede the verb itself (Inoue & Fodor, 1995; Konieczny, 1996; Mazuka &

---

1 Loading the generic frame before the first word might be feasible for a reader who used the full-stop at the end of the preceding sentence to infer that a new one was about to start. In spoken comprehension, a change of speaker might provide comparable information.
Rather than examining the details of these theories, we shall refer to the general account as the pre-head attachment hypothesis. In our discussion, the pre-head attachment hypothesis is taken to cover any account that allows pre-head attachment in at least some linguistic circumstances.

The major aim of the present paper is to explore the issue of head-driven/pre-head attachment. To recap, we have defined head-driven parsers as making no attachment decisions prior to heads, and then trying to satisfy head-derived constraints as soon as heads arrive. In contrast, our definition of pre-head attachment parsers suggests that the parser does make attachment decisions before heads arrive. As already mentioned, head-final languages provide an important arena for investigating this question. In the present paper we shall concentrate upon an attachment ambiguity that occurs in the Japanese language. Before moving onto the sentence structure, we shall review one or two previous empirical studies dealing with the issue.

Empirical Studies on Head-driven/Pre-head-driven Parsing in Head-final Structures

In the past, the majority of studies of head-awaiting have been carried out in the German language (Bader & Lasser, 1994; Hemforth, 1993; Konieczny, 1996; Konieczny et al., 1997; Scheepers, Hemforth, & Konieczny, 1994). For instance, Bader and Lasser (1994) conducted a self-paced reading task using the following ambiguity in German subordinate clauses:

(5a) daß sie (Nom/Acc) nach dem Ergebnis zu fragen tatsächlich erlaubt hat
that she/her for the result to ask indeed permitted has
“that she indeed has permitted to ask for the result”

(5b) daß sie (Nom/Acc) nach dem Ergebnis zu fragen tatsächlich erlaubt worden ist
that she/her for the result to ask indeed permitted has been
“that to ask her for the result has been permitted”

The sentences in (5) contain a female singular pronoun “sie” that is ambiguous between nominative (“she” for English) and accusative case (“her” for English). Permissible argument structures of the first verb “fragen” include “fragen [NP-Nom <Agent>, NP-Acc <Goal>, nach-PP <Theme>]” and “fragen [NP-Nom <Agent>, nach-PP <Theme>]”. The temporary ambiguity in (5) is whether the ambiguously case-marked “sie” is analysed as the object of the first verb “fragen” (accusative reading), or as the subject of the second verb “erlaubt” (nominative reading). The
ambiguity can be resolved by the appearance of an auxiliary word at the end of the clause. If “hat” (has) follows, as in (5a), the nominative analysis turns out to be correct, while if “worden ist” (has been) follows, as in (5b), this implies that the correct reading is the accusative one. The head-driven parsing hypothesis predicts that the ambiguous “sie” should not be attached until the first verb (“fragen”). As soon as it is encountered, it is attached to the verb (“fragen”), and so becomes disambiguated as an accusative noun. This analysis turns out to be wrong at the end of the clause in the “hat” condition, (5a), while this is compatible with “worden ist” in (5b). Using this manipulation with two additional control conditions, Bader and Lasser (1994) found that the unambiguous “sie” tends to be analysed as the subject of the forthcoming final verb (i.e., nominative reading). They argued that this finding is incompatible with the head-driven attachment hypothesis.

There is, however, an important concern with German studies of this kind. This stems from the fact that the main manipulations were based on ambiguities of case. It is arguable that certain types of head-driven account could produce effects of the kind that Bader and Lasser (1994) and others interpret as evidence against such models. For example, it is possible to imagine a head-driven system in which case can be assigned to an NP not by the parser itself but by reference to some extraneous strategy—such as automatically assigning nominative case to the first NP in the sentence without any kind of prior higher level structural analysis. For instance, Hopf, Bayer, Bader, and Meng (1998) explicitly acknowledge, in their Footnote 6, that such case effects may be the result of statistical, as opposed to syntactic operations (also, see Hemforth, Konieczny, & Strube, 1993, for evidence of very early signs of case bias in German sentences). Such a possibility is not just a matter of ad hoc speculation: such proposals have a long history in the study of parsing (see also a related statistical proposal by Vosse and Kempen, submitted). For example, Fodor, Bever, and Garrett (1974) proposed that readers and listeners assign preliminary structure to sentences using what they called a “Canonical Order strategy”. This involves immediately treating the first NP of a sentence as its subject (and might equally well be interpreted in terms of assigning it nominative case). The crucial feature of any “canonical order” operation is that the decisions are based not on any thorough syntactic analysis of the sentences currently under scrutiny, but on the statistically most common order in which NP category assignments occur in sentences in general (see Hyönpää & Hujanen, 1997, for recent evidence that something of this kind might occur in Finnish). Another line of support for these proposals comes from observations concerning assignment of other types of licensing relations during sentence comprehension. Bates, MacWhinney and colleagues have reported numerous
studies showing that NPs can be assigned functional roles on the basis of operations that may have little or nothing to do with conventional parsing (see, for example, MacWhinney & Bates, 1989). Such investigations show that these biases are evident even when people are presented with ungrammatical word-strings—indicating that it is implausible to treat them as by-products of any routine forms of syntactic analysis. Because of the close association between case and functional roles, it seems highly likely that there is also a (non-syntactic) predisposition to assign case to at least some of the NPs within a sentence. If such NPs can receive case from something other than the results of routine syntactic analysis then the evidence for a nominative or genitive bias in the German studies might not necessarily have any implications for syntactic processing. The preferences might have occurred not because the NP has been structurally linked to a verb that is yet to appear, but because of a broad case assignment in which verb information is not involved. Given these possibilities, there are grounds for being sceptical about the implications of the German findings and, in particular, whether they provide definitive evidence against head-driven models of parsing.

In a more recent study, Koh (1997) has addressed the issue of head-driven parsing in an experiment that is not based on case assignment and is therefore not subject to the reservations raised above. This investigation was organised around a temporary attachment ambiguity in the Korean language, in which a dative-marked NP can potentially be attached either to a subordinate following a word or two later or to a matrix verb scheduled to appear at the end of the sentence. In an eye-tracking study, Koh (1997) compared reading times in sentences of this kind with sentences where there was a forced link between the dative NP and the subordinate verb. According to head-driven accounts, the matrix link should not be an option at any point prior to the appearance of the main verb. Consequently, for most of its length the sentence should be processed exactly as if it were unambiguous, leading to the prediction that there should be no systematic processing differences between the ambiguous sentence and its control. In contrast, a pre-head attachment model would allow the dative NP to be attached to the as-yet-unseen matrix verb where this is permitted (i.e., in the temporarily ambiguous condition). This would lead to processing problems in cases where the final verb rules out the interim matrix attachment. Corresponding difficulties would not, however, be expected in the control condition where matrix attachment was not posited in the first place. In line with the second hypothesis and against the head-driven account, Koh (1997) found that the total reading times were significantly longer in the region of the matrix verb. However, the interpretation of this result was complicated by the fact that different matrix verbs were used in the experimental and control conditions, raising
the possibility that there might be a *lexical* rather than a syntactic explanation for the differences obtained. To address this problem, Koh’s experiment incorporated two further conditions designed to act as lexical controls. Unfortunately, the difference was not significantly reduced in the lexical control conditions, indicating that the data could not rule out the lexical explanation of the findings. Given this pattern of results, it has to be concluded that Koh’s results fail to provide convincing support for the pre-head attachment model of parsing.

The Present Study: Japanese Argument Attachment Ambiguity

As seen, previous empirical studies (Bader & Lasser, 1994; Koh, 1997; Scheepers et al., 1994) have not succeeded in providing solid support for the hypothesis that structural attachment analyses of pre-verbal constituents are accomplished prior to the host verb. To overcome the problems of both the German and Korean studies, it is necessary to conduct experiments that avoid the theoretical and methodological objections outlined above. In the present paper, we report a series of off-line/on-line experiments in Japanese that look at whether the parser initiates structural analyses of pre-verbal constituents before the host verb. Using a Japanese ambiguity enables us to avoid the potential problem of the previous German experiments (Bader & Lasser, 1994; Scheepers et al., 1994) discussed above, namely, the possibility that the head-driven parser might pre-verbally assign cases independent of (and prior to) structural analyses. The following ambiguity will be examined (the ambiguity is based on Koh, 1997):

(6) Kyooju-ga gakusee-ni toshokansisho-ga kasita mezurasii komonjo-o miseta.
Professor-Nom student-Dat librarian-Nom lent unusual ancient manuscript-Acc showed.
“The professor showed [HA: the student] the unusual ancient manuscript which the librarian had lent [LA: the student].”

The ambiguity in (6) is that the dative NP “gakusee-ni” (student-Dat) can be attached either to the VP headed by the subordinate verb “kasita” (lent) or that headed by the matrix verb “miseta” (showed) (we shall hereafter refer to the first option as “low attachment” and the second option as “high attachment”, as well as “subordinate attachment” and “matrix attachment”, respectively). The ambiguity is “global” in the sense that the grammar does not force either attachment to be confirmed at any point in the sentence. The Japanese global ambiguity is possible since a
dative Goal argument is always optional for Japanese ditransitive verbs. Both attachment decision form a grammatical construction. In the high attachment analysis, “mezurasii komonjo” (unusual ancient manuscript) was shown to “gakusee” (student), while the Goal argument of “kasita” (lent) is assigned to a “pro” (or left empty), representing perhaps someone who is not mentioned in the sentence. In contrast, the low attachment decision leads to the interpretation that “mezurasii komonjo” was lent to “gakusee” whereas it was shown to someone else, filling the Goal of the alternative verb with a “pro”. The earliest indication of the ambiguity is the appearance of the second nominative NP “toshokansisho-ga” (librarian-Nom), since there has been only one dative NP for two clauses. However, note that it is impossible at this stage to establish whether the host verb of each clause can legitimately take a dative verb.

The two different types of parser under investigation, namely the head-driven and pre-head attachment systems, would process the Japanese sentence (especially the dative NP) in a different manner.

We start by giving detailed consideration of the step-by-step operations of a head-driven parser. By hypothesis, this would not start structural analysis of the pre-verb arguments until the verbs are encountered. When the first verb (“kasita”) appears, the dative NP would presumably be analysed as the indirect object of the verb. In Pritchett’s framework this follows from the tendency to define licensing relations as soon as possible (Generalized Theta Attachment; (3)). Within the Vosse/Kempen U-Space model the same would occur for the mere reason that the NP and “kasita”-frames co-exist for a time in U-space without any competition for their mutual connections. In either framework, the low-attached link would be retained at least until the alternative verb appears. When the matrix verb “miseta” is encountered, the parser faces the problem that both verbs can take a dative argument. At this point there are two ways in which a head-driven parser might resolve the dilemma. On the one hand it might retain the initially established low-attachment decision and avoid possible reanalysis costs (leading to a durable low-attachment preference at the end of the sentence). On the other hand, it might try to achieve a balance between the two alternatives, being guided by the competing constraints associated with the two heads. Assuming (a) that there is no additional cost associated with unravelling existing decisions (reanalysis) and (b) that there is no imbalance between the acceptabilities of the two dative roles in the two respective analyses, then this state of affairs might eventually lead to 50% low and 50% high attachment preferences. In short, head-driven accounts would predict at one extreme a 50/50 attachment bias and this would tilt in the direction of a low-attachment preference on the plausible assumption that there is some “inertia” in the system (favouring existing links over restructuring or reanalysed relations).
In contrast with this, an incremental pre-head attachment parser would be capable of making its high-level attachment decisions without having to wait for either of the verbs to appear, for example, as soon as the NP-arguments are encountered. Thus, making use of the early availability of top-down information (e.g., a “generic sentence frame” discussed earlier) the parser would link the first two NPs to the generic verb slot within the top-down frame well before a verb arrives. The later arrival of the subordinate verb (“kasita”) would provide a competing set of links for the dative NP. Potentially this could lead to a steady reduction in the strengths of the links to the matrix verb depending on the exact details of the conflict-resolution process. In some variants of the account, the procedural details might favour the links to the explicitly-realised bottom-up verb slots rather than to the postulated generic verb frames. This might lead to ultimate (though not necessarily immediate) predictions that are indistinguishable from those for the head-driven models outlined above. In other variants, initially robust links may be preserved at the expense of later-arriving competitors. Such models would predict a completely different (high-attached) preference at the end of the sentence. However, although the former type of pre-head attachment hypothesis is theoretically possible, we shall hereafter deal with only the latter account in our discussion on incremental pre-head attachment parsing for simplicity.

The study presented in the present paper consists of one questionnaire study (Experiment 1) and one self-paced reading task (Experiment 2). Experiment 1 aims to demonstrate an off-line preference for the ambiguity. Note that the two opposite accounts predict different eventual preferences as discussed above. On the other hand, Experiment 2 looks at on-line processing of the ambiguity under investigation. Additionally, before Experiment 1, a preliminary test is conducted to rate how strongly an optionally ditransitive verb requires a dative Goal argument. The rating results are to be used to counterbalance the strength of the dative requirement in the two attachment host verbs in the globally ambiguous sentences (e.g., (6)) used in the subsequent experiments. Although corpora would provide argument structure information based on data from a wider range of materials/readers, it has to be noted that no reliable corpora are available for Japanese verbs at present.

---

2 Note that the constraint-satisfaction approach (MacDonald et al., 1994; Trueswell & Tanenhaus, 1994) argues that both attachment hosts are equally preferred off-line given that the dative requirement strength is counterbalanced between the two verbs. Thus, the approach predicts no off-line attachment preference for the ambiguity given the lexical counterbalance.
EXPERIMENT 1

The present experiment examines the off-line preference for the attachment ambiguity under consideration. The eventual preference for the ambiguity has important implications for the question of whether Japanese parsing is head-final or incremental, since the two opposite accounts make different predictions about the preferences that will prevail. As already noted, head-driven accounts predict either (a) that low attachment should be eventually preferred as a consequence of immediate satisfaction of an argument structure requirement (without reanalysis), or (b) that both attachments should be equally preferred because both attachment decisions equally satisfy argument structure requirements (with reanalysis) (see the Introduction above, and also Table 2 below). In contrast, the incremental pre-head attachment hypothesis predicts that high attachment should be preferred since it is the first attachment option that the parser commits itself to.

Preliminary Test

Before Experiment 1, a preliminary questionnaire study was carried out. The purpose of this was to obtain norms in order to establish how strongly various Japanese ditransitive verbs require a Goal (dative) argument. The results of this analysis will be used to provide a basis for selecting materials for use in Experiment 1 and later also in Experiment 2. To counter against biases originating from the properties of individual verbs, we require two groups of Japanese ditransitive verbs that require a Goal argument with the same strength. Such counterbalancing should ensure that any possible attachment preference can be attributed to structural biases as opposed to lexical biases.

Eighty-eight native speakers of Japanese participated in the Preliminary Test. All of them were students (undergraduates and postgraduates) or academic staff at the Kumamoto University, Japan.

Initially, 24 globally ambiguous complex sentences such as (6) above were constructed, using 48 Japanese ditransitive verbs from Mizutani (1985). Then, the 24 complex sentences were divided into two unambiguous simplex sentences (e.g., (7a) and (7b) were made out of (6)):

(7a) Kyooju-ga gakusee-ni mezurasii komonjo-o miseta.  
"The professor showed the student the unusual ancient manuscript."

The material set of 24 ambiguous sentences is given in Appendix 2: the sentences were used as the experimental sentences in Experiments 1 and 2, as we shall discuss below.
(7b) Toshokansisho-ga gukasee-ni mezurasii komonjo-o kasita.
“The librarian lent the student the unusual ancient manuscript.”

Furthermore, versions of the sentences in which the dative arguments were omitted were constructed (e.g., (8a) and (8b) for (7a) and (7b), respectively): 4

(8a) Kyooju-ga mezurasii komonjo-o miseta.
“The professor showed the unusual ancient manuscript.”

(8b) Toshokansisho-ga mezurasii komonjo-o kasita.
“The librarian lent the unusual ancient manuscript.”

In the Preliminary Test, subjects were asked to use a numerical scale to rate how well they understood the “dative missing” sentences like (8), given the benchmark that they were called on to use the score of 100 to rate their understanding of the “full” sentences like (7). All materials were presented as pairs (e.g., (7a) and (8a) or (7b) and (8b)) in Japanese script on one-sided A4 sheets. In total, there were 48 unambiguous pairs. There was no explicit instruction as to whether or not numerical ratings could exceed 100.

The rating score for each Japanese ditransitive verb is given in Appendix 1. The score varied from 66.4 to 76.14, and the mean score was 72.21. It was revealed that the ditransitive verbs which served as a subordinate verb and those which served as a matrix verb in the initially-constructed material set (24 sentences like (6)) contained statistically equal strengths of dative requirements on average (mean scores: Subordinate Verbs: 71.40; Matrix Verbs: 71.01; $F(2,147) = 0.266, P = .609, MSe = 3.956$). Therefore, the results suggested that the original material set would provide counterbalanced dative requirement biases in the further experiments (Experiments 1 and 2).

Method

Subjects. Eighty-eight native speakers of Japanese participated in Experiment 1. All of them had previously participated in the preliminary test.

Materials and procedure. All materials were presented in Japanese script on one-sided A4 sheets. The task was to read sentences like (6) (repeated as (9a)), and to answer two subsequent questions like (9b) and

4 This somewhat reversed order of constructing materials was chosen in order to examine the same combinations of objects and verbs in both the preliminary test and further experiments.
(9c) (the (9a) versions of all sentences examined in Experiment 1 are given in Appendix 2 as the globally ambiguous (GA) condition sentences for Experiment 2):

(9a) Kyooju-ga gakusee-ni toshokansisho-ga kasita mezurasii komonjo-o miseta.
Professor-Nom student-Dat librarian-Nom lent unusual ancient manuscript-Acc showed.
“The professor showed [HA: the student] the unusual ancient manuscript which the librarian had lent [LA: the student].”

(9b) Dare-ni-kyooju-ga kononjo-o misemasitaka.
“To whom did the professor show the ancient manuscript?”

(9c) Dare-ni toshokansisho-ga kononjo-o kasimasitaka.
“To whom did the librarian lend the ancient manuscript?”

Twenty-four dative ambiguous sentences (e.g., (9a)) were tested. The subsequent questions were about Goal roles for the ditransitive verbs in the sentence. Dative NPs stood for a person or a group of people in all sentences. Subjects were explicitly instructed that they could use the same persons for both Goal roles in a sentence, and also that they could use “sonata” (someone else) when they did not find any overt Goal in the sentence. The order of the matrix question (e.g., (9b)) and a subordinate question (e.g., (9c)) was counterbalanced. Also, there were two different versions in the order of all 24 experimental sentences. Altogether, each of the four versions was presented to 22 subjects. As mentioned earlier, the ditransitive verbs used for the subordinate verb and those used for the matrix verb had statistically the same strength of dative requirements on average, based on the results from the preliminary test.

Results and Discussion

Six patterns of attachment were found in the questionnaire study. Table 1 gives those attachment patterns and percentages.

The first column indicates the name of each attachment pattern found in Experiment 1. The result (as percentage) for each pattern is shown in the second column. The alternative attachment patterns depend on which participants (constituents) were selected for the answers of the questions (9b) and (9c) above. The third column represents which constituent is the Goal role of the matrix verb in each pattern. For instance, in the MS attachment pattern (first row), the dative NP (“gakusee” (the student) for (9a)) is the answer for the question like (9b) above. On the other hand, the constituent in the fourth column indicates which constituent is the Goal role of the subordinate verb in each attachment pattern. For example,
“kyooju” (the professor) is the answer for the question like (9c) above in the MS attachment pattern. Among all the six patterns, there was a robustly significant difference across attachment pattern ($F_1(5,435) = 269.67, P < .001, MSe = 9.76; F_2(5,115) = 48.81, P < .001, MSe = 196.66). The preference for the MS analysis was significantly greater than that for any of the alternatives (e.g., MS vs HA: $F_1(1,87) = 127.42, P < .001, MSe = 34.45; F_2(1,87) = 28.90, P < .001, MSe = 547.81). In all attachment patterns, it is possible to categorise the MS and HA analyses as the high attachment group, and the SS and LA analyses as the low attachment groups, according to the attachment of the dative NP under investigation. Thus, Experiment 1 showed that readers opt for variations of high attachment significantly more frequently than those of low attachment (HA and MS vs LA and SS: 79.88% vs 4.88%; $F_1(1,87) = 1687.99, P < .001, MSe = 8.25; F_2(1,23) = 192.25, P < .001, MSe = 263.13). This is inconsistent with any of the head-driven predictions for the ambiguity (which ranged from a low-attachment preference to a 50/50 attachment preference). The results are also incompatible with versions of the pre-head driven model based on the premise that the initial high-attached preference is substantially reversed before the end of the sentence (as might occur, for example, if links to explicitly-realised verbs were favoured over those to postulated verbs). The overwhelming bias in favour of high-attachment is not consistent with competing effects of this kind. In contrast, the results are fully in line with the predictions of the pre-head driven models in which an initial attachment to a verb-slot within a hypothesised top-down structure is lodged so strongly by virtue of its early arrival that it cannot be displaced by its competitors.

**EXPERIMENT 2**

Finally, a self-paced reading task on the ambiguity was conducted. The purpose of the experiment is to demonstrate on-line preferences for
the ambiguity. There were the following three experimental conditions in Experiment 2 (the slashes show the segmentation pattern used in the self-paced experiment; the dative NP is italicised; the verb(s) in bold is the legitimate attachment host; (11) gives the name of each segmented region):

(10a) Globally Ambiguous (GA) condition:
Kyooju-ga / _gakusee-ni_ / toshokansisho-ga / _kasita_ / mezurasii komonjo-o / _miseta_.
Professor-Nom student-Dat librarian-Nom lent unusual ancient manuscript-Acc showed
“The professor showed [HA: the student] the unusual ancient manuscript which the librarian had lent [LA: the student].”

(10b) High Attachment (HA) condition:
Kyooju-ga / _gakusee-ni_ / toshokansisho-ga / yabutta / mezurasii komonjo-o / _miseta_.
Professor-Nom student-Dat librarian-Nom tore unusual ancient manuscript-Acc showed
“The professor showed the student the unusual ancient manuscript which the librarian had torn.”

(10c) Low Attachment (LA) condition:
Kyooju-ga / _gakusee-ni_ / toshokansisho-ga / _kasita_ / mezurasii komonjo-o / yabutta.
Professor-Nom student-Dat librarian-Nom lent unusual ancient manuscript-Acc tore
“The professor tore the unusual ancient manuscript which the librarian had lent the student.”

(11) Matrix Subject / Dative NP / Subordinate Subject / Subordinate Verb / Accusative NP / Matrix Verb.

(10a) contains a ditransitive verb (i.e., a potential attachment host for the dative NP) for both the subordinate and matrix verbs, while (10b) and (10c) contain one ditransitive verb for only the matrix and subordinate verbs, respectively. The other verbs (the subordinate verb in (10b) and the matrix verb in (10c)) are monotransitives that cannot take a dative indirect object. The legitimate reading: both attachments for (10a), high attachment only for (10b), and low attachment only for (10c). Therefore, each condition ((10a), (10b), and (10c)) is referred to as the “Globally Ambiguous (GA)” condition, the “High Attachment (HA)” condition, and the “Low Attachment (LA)” condition, respectively.

Examining the three conditions enables us to have a close look at attachment procedures at each part of the sentences. Table 2 shows
attachment decisions that each account predicts at three position in each condition.\(^5\)

According to head-driven accounts, the parser delays all argument-
attachment decisions until it comes across a verb. Such accounts therefore
predict that no attachment is established prior to the first verb in any of the
three conditions. However, when the first verb is encountered, head-
derived information becomes available, and the parser tries to fill the
associated requirements as soon as possible. Thus the prediction of head-
driven accounts is that, for the GA and LA conditions, low attachment
takes place at the subordinate verb—because in these two cases the verb
enables a dative NP to receive a thematic role. In the LA condition, this
structural analysis should be retained, without further problem or reanalysis, until the end of the sentence. Note that the LA sentence is comparable to the following sentence in which Pritchett (1992) predicts no
processing breakdown occurs:

(12) Frank-ni Tom-ga Guy-o syookai suru to John-wa omotteiru.
     Frank-dat Tom-Nom Guy-Acc introduce Comp John-Top thinking.
     “John thinks that Tom will introduce Guy to Frank.”

\(^5\) In Table 2, we include only predictions made by the head-driven and pre-head attachment parsing hypotheses. However, there are a number of other theoretically possible attachment patterns. The following shows such patterns excluded in Table 2:

(i) no attachment → low attachment at NP-Nom2 in all conditions
(ii) high attachment → low attachment at NP-Nom2 in all conditions
(iii) low attachment → high attachment at the subordinate verb in the GA LA conditions
(iv) no attachment → high attachment at the subordinate verb in the GA and LA conditions
(v) high attachment → low attachment at the matrix verb in the GA condition
(vi) high attachment → 50/50 at the matrix verb in the GA condition
(vii) low attachment → low attachment at the matrix verb in the ga condition.
There are two differences between the sentence structures in (10c) (the LA condition) and (12). Firstly, the subordinate clause is a sentential complement clause in (12) whereas it is a relative clause in (10c). Secondly, the matrix subject (“John-wa”) is placed just before the matrix verb in (12) whereas it (“toshokansisho-ga”) is placed three constituents before the matrix verb in (10c). Despite these differences, Sentence (12) provides a good pointer to the prediction Pritchett (1992) would make for (10c), since, in both sentences, attachment of the dative NP is in principle ambiguous until the matrix verb appears and the matrix verbs cannot take dative arguments. Pritchett (1992) argues that his head-driven parser would attach the dative NP (“Frank-ni” in (12)) to the first verb (“shookai suru” in (12)) as soon as the verb is encountered. The clause is first analysed as simplex (i.e., the matrix clause), but the subsequent appearance of the second subject “John-wa” indicates that the clause is subordinate. According to Pritchett (1992), the initial attachment of the dative NP, which would be referred to as low attachment in (10c), is retained throughout the sentence, and no reanalysis is required, as described in the following:

(13) Such sentences (N.B. the original simplex sentence analysis prior to the complementiser and the final complex sentence analysis at the matrix verb) are both fully grammatical and acceptable and yield no conscious processing difficulty whatsoever (Pritchett, 1992; p.153).

Thus, it seems safe to argue that Pritchett (1991, 1992) unambiguously predicts no reanalysis at the matrix verb in the LA condition.

In the GA condition, it is revealed that the matrix verb also holds a thematic requirement. As previously discussed, the parser might resolve this conflict in either of two different ways. The parser might retain the previously established low attachment avoiding reanalysis (low attachment). Alternatively, it could opt equally for both attachments, with competition being resolved in favour of the low attachment decision in half of the cases, while reanalysis occurs in the other half (50/50 attachment). Turning next to the HA condition, the basic prediction is that the dative NP initially remains unattached (because the head properties of the subordinate verb prevent it from taking dative arguments). At the matrix verb, matrix verb attachment is established—as this is the first, and only linkage, which licenses the dative thematic role. This process should be relatively low-cost, since there is no competition or reanalysis involved in this case.

The corresponding predictions for pre-head attachment accounts are markedly different. Such models posit that the parser attaches the dative NP to an attachment host (e.g., the matrix IP, the matrix VP, or a verb-slot
postulated as part of a top-down generic sentence frame), and that this commitment is made well before the subordinate verb is first encountered.\(^6\)

At the matrix verb, the parser has to reconcile this initial commitment with the argument structure of the final two-place matrix verb. In the GA and HA conditions the initial commitment and the final argument structures are entirely compatible with one another and so the prediction is that the analysis will run through quickly and easily in these two conditions. In the LA condition, on the other hand, the initial commitment and the matrix argument structure are in conflict with one another. Here, it is predicted that the conflict should lead to a revision, and consequently an increment in reading time for the final verb.

The predictions in Table 2 will result in particular reading time patterns in the self-paced reading task. Table 3 summarises the various reading time predictions that can be made on the basis of the two classes of parsing model.

Head-driven parsing accounts predict different attachment operations at the subordinate verb across the conditions: low attachment occurs in the GA and LA conditions, while no attachment decision is made in the HA condition. It seems reasonable to consider that the accounts (as well as other types of parsing models) predict making no attachment and making an attachment requires statistically indistinguishable reading times when no reanalysis is involved. Thus, no reading time difference is predicted across the conditions at the subordinate verb within the head-driven framework (GA = HA = LA). At the matrix verb, the accounts predict no reading time difference across the conditions given the low attachment

---

\(^6\) Given the outcome of Experiment 1, we assume that there is little or no reanalysis of the generic verb (high-attached) link when the subordinate verb is encountered. (Such reanalysis would have yielded end-of-sentence results different from those already reported and so there is little point in entertaining this possibility any further.)
prediction (GA = HA = LA). On the other head, reanalysis is involved in half of cases in the GA condition given the 50/50 prediction. It is predicted that the GA condition takes more time to read than the other conditions (GA > HA = LA). On the contrary, incremental parsing accounts predict no reading time difference across the conditions at the subordinate verb, since they predict that the previously established high attachment is retained in all conditions (GA = HA = LA). At the matrix verb, the accounts predict that the LA condition takes more time to read due to a reanalysis from high to low attachment (LA > GA = HA).  

Method

Subjects. Twenty-four native speakers of Japanese participated in Experiment 2. All of them were students (undergraduates and postgraduates) at the University of Exeter. All of them had corrected or uncorrected normal vision. None of them had participated in a similar psycholinguistic experiment before, or was aware of the purpose of the experiment. Subjects were paid £2 for their participation. The average total duration of their stay in English-speaking countries was 10.42 months (minimum: one month; maximum: eight years).

Apparatus. A NEC laptop computer (98 NOTE SX/T) was used for controlling the experiment procedure and recording date. The experiment was set up and run with an experimental package for NEC computers written in N88-BASIC. Two keys on the computer keyboard were used for Yes and No keys.

Materials and conditions. As discussed above, three experimental conditions were examined as shown in (10) (the experimental sentences of Experiment 2 are given in Appendix 2). There were 24 sets of three

---

7 Alternatively, head-driven accounts can predict that the reanalysis processes do not result in producing statistically longer reading times in the GA condition at the matrix verb, since analysis takes place only in 50% of all cases given this prediction (i.e., GA = HA = LA).

8 On the other hand, it is possible to assume that assigning a “pro” causes some measurable costs. The operation could occur at the subordinate verb in the GA and LA condition or at the matrix verb in the GA condition: the pre-head attachment parser might decide to leave the Goal role of the low attachment verb unfilled as soon as the verb is encountered (i.e., at the subordinate verb), or after all items have appeared in the sentence (i.e., during wrap-up processes at the matrix verb). Given the former, the GA and LA conditions take equally longer than the HA condition at the subordinate verb (i.e., GA = LA > HA). Given the latter, the GA condition takes longer than the HA condition at the matrix verb, while the relation between the LA and GA conditions is arbitrary (i.e., LA > HA; GA > HA).
experimental conditions. The sentences in the GA condition had been tested in Experiment 1. As fillers, 48 other sentences with irrelevant structures were included. The 24 sets of experimental sentences were divided into three material files. Each material set was read by one of three subject groups (eight subjects for each group). As shown in (10), one of three conditions contained a different lexical item from the others in the Subordinate Verb region and Matrix Verb region. Table 4 shows the mean number of characters for the Subordinate Verb and Matrix Verb region for each material file:

As indicated in (10), there were three groups of verbs in the subordinate and matrix verb positions (e.g., “kasita” (lent), “miseta” (showed) and “yabutta” (torn)). We conducted a 5-scale familiarity rating questionnaire on 46 Japanese native speakers using the 72 experimental verbs and 72 filler verbs. The mean familiarity ratings of the verbs were found to be statistically indistinguishable across the verb groups (mean rating scores (1—very unfamiliar; 5—very familiar)): “kasita” verbs: 4.55 (SD = 0.39); “miseta” verbs: 4.55 (SD = 0.40); “yabutta” verbs: 4.52 (SD = 0.38); $F_1(2,90) = 0.97, P = .424, MSe = 0.01; F_2(2,46) = 0.04, P = .961, MSe = 0.11). In particular, there was no significant familiarity difference between the “kasita” verbs and the “yabutta” verbs (see the reading time predictions for the matrix verbs in Table 3) ($F_1(1,45) = 1.49, P = .228, MSe = 0.01; F_2(1,23) = 0.06, P = .804, MSe = 0.11$).

Out of 24, four sentences were followed by a yes/no question. The number of correct yes/no answers was counterbalanced (two for each).

**Procedure.** A self-paced reading technique was used in the experiment. The stimuli were presented in a non-cumulative segment-by-segment fashion with a moving-window technique. Also, they were randomised by the computer without any constraint. In 16.7% of all trails (12 in total: 4 experimental sentences, 8 fillers), a yes/no question about

<table>
<thead>
<tr>
<th>Subordinate Verbs</th>
<th>Matrix Verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>HA</td>
</tr>
<tr>
<td>File 1</td>
<td>3.25</td>
</tr>
<tr>
<td>File 2</td>
<td>3.5</td>
</tr>
<tr>
<td>File 3</td>
<td>3.5</td>
</tr>
<tr>
<td>Mean</td>
<td><strong>3.417</strong></td>
</tr>
</tbody>
</table>
the content followed the material sentence. Two keys on the computer keyboard were used for the Yes and No keys for questions. First, the first display showing “Sentence Comprehension Experiment” (in English) appeared in the centre of the screen. Then, instructions for the experimental procedure were presented on an A4 sheet. Subjects were asked to read them silently. The instructions included that the task was to read stimuli on the screen as naturally and correctly as possible while pressing the space bar. Also, other operative instructions were included, as shown below. After subjects read the instructions, the experimenter indicated all keys to be used in the experiment. Then the Practice Trial block (consisting of four trials) and the Experimental Trial block (consisting of 72 trials) followed. Subjects were allowed to ask any question about experimental procedure before/during/after the Practice Trial block. If they were certain that they understood the experimental procedure, the Experimental Trial block was started, and the experimenter withdrew from the test-room. Each trial was presented by the instruction “Press space bar” (in English). When the subject did this, the first display appeared. In both the Practice and Experimental Trial blocks, the beginning of each trial was indicated by a display consisting of only this instruction. When subjects did this, dots (●) appeared in the centre of the screen, indicating the length of the subsequent material sentence. Further space bar pressing triggered the first segment of the sentence, replacing dots with the first fragment. When subjects read the first segment, they were asked to press the space bar, so that the first segment was replaced with dots again, and the second segment was opened. Subjects continued this operation until the end of the sentence was reached, which was indicated by no further dots on the right hand site. The dots and final segment of the sentence disappeared when the space bar was pressed again. Next, a yes/no question followed in some trials. Answering the question by pressing the allocated Yes/No keys triggered the very first display of the next trial (i.e., “Press space bar”). When there was no question, the instruction appeared just after the final space bar pressing of the previous trial. The end of the Experimental Trial block was marked by a display saying “Thank you very much!!”. Then, subjects were asked to fill in a form asking about their name, affiliation, age, sex, languages they spoke (other than English and Japanese), the duration of having lived in an English-speaking country. After answering the personal information questions, the subjects were asked to call the experimenter who had been waiting outside. Then, they were paid £2 for their participation and dismissed. All sessions were carried out in a dim silent test room in the Washington Singer Laboratories, University of Exeter. The average session took approximately 30 minutes.
Results

Table 5 gives reading times per character for all regions in each condition. Among the six regions, only the subordinate verb and matrix verb comprise different lexical items across the conditions. ANOVAs confirmed that the same materials were read with the same difficulty in the other regions: there were no statistical differences in reading times across the conditions in other regions (Matrix Subject, Dative NP, Subordinate Subject, and Accusative NP). Thus we shall discuss reading times for only the two critical verb regions. At the subordinate verb, the effect of condition was not significant ($F_1(2,46) = 0.03, P = .973, MSe = 3800.98; F_2(2,46) = 0.01, P = .991, MSe = 445.25$). There was no significant difference between any pair of the three conditions. On the other hand, at Matrix Verb, the main effect of condition was significant ($F_1(2,46) = 5.33, P = .008, MSe = 7216.96; F_2(2,46) = 3.38, P = .043, MSe = 10005.82$). Furthermore, it was shown that reading times were increased at the matrix verb in the LA condition compared to the mean reading times of the other two conditions (LA vs the mean of GA and HA: $F_1(1,23) = 6.00, P = .022, MSe = 9379.73; F_2(1,23) = 4.23, P = .051, MSe = 11625.32$). In contrast, the GA and HA conditions produced statistically indistinguishable reading times for Matrix Verb ($F_1(1,23) = 0.99, P = .33, MSe = 1927.58; F_2(1,23) = 0.43, P = .519, MSe = 4511.31$). The results, therefore, suggest that readers had a problem when it turned out that the matrix verb could not take a dative argument (i.e., in the LA condition). This suggests that the parser must have been unravelling a commitment already made prior to this point—namely a link between the dative NP and the as-yet-unread matrix verb. Hence, the prolonged reading times in the LA condition are consistent with the incremental pre-head attachment hypothesis, but not with the head-driven parsing hypothesis. Also, the lack of any difference between the GA and HA conditions throughout the whole sentence.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Matrix Subject</th>
<th>Dat NP</th>
<th>Subordinate Subject</th>
<th>Subordinate Verb</th>
<th>Acc NP</th>
<th>Matrix Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>GA</td>
<td>235.79</td>
<td>238.65</td>
<td>243.65</td>
<td>240.92 (126.45)</td>
<td>141.83</td>
<td>225.44 (108.86)</td>
</tr>
<tr>
<td>HA</td>
<td>255.97</td>
<td>237.28</td>
<td>252.03</td>
<td>237.22 (141.93)</td>
<td>133.25</td>
<td>212.83 (96.23)</td>
</tr>
<tr>
<td>LA</td>
<td>251.30</td>
<td>251.01</td>
<td>253.48</td>
<td>237.37 (108.22)</td>
<td>139.56</td>
<td>287.61 (198.91)</td>
</tr>
</tbody>
</table>

Note: The values inside the parentheses show the standard deviations.
suggests that the global ambiguity did not yield significant extra processing costs.

Discussion

Overall, the whole pattern of data was compatible with the pre-head attachment hypothesis, but inconsistent with head-driven accounts: the increased reading times for the matrix verb region were found in the condition where the matrix attachment analysis turned out to be ungrammatical (i.e., the LA condition). This is an indication of early commitment to high attachment, which can be predicted only by pre-head attachment accounts. More specifically, the head-driven model failed to predict the increased processing load at the ends of sentences resolved in favour of low-attachment. Prior to the arrival of the final verb, such models have no basis for postulating anything other than low attachment, and so when the final verb eventually turns out to be fully compatible with this analysis, the model cannot fail to predict a quick and easy completion of the syntactic analysis. Equally, there was no sign at all of the positive predictions of such models—namely that there might be a cost associated with a high-attached resolution compared with the globally ambiguous control (the numerical difference was in the opposite direction—albeit non-significant).

In contrast, the results were exactly as predicted by the variant of the pre-head driven model in which the dative NP is initially attached to a postulated head-free structure and remains attached in this way up until the arrival of the final word in the sentence.

Although the reading times results may look straightforward, however, there are some possible criticisms to the present study. One of the more obvious criticisms would be that there was no manipulation before the first verb across the conditions (see (10)). The lack of manipulation made it impossible to identify the exact point at which the preferred high attachment is established. In fact, there are three possible decision-making points that are compatible with the pattern of data from Experiment 2: at the dative NP ("gakusee-ni"), at the second nominative NP ("toshokan-sisho-ga"), and before the subordinate verb ("kasita"). Although we have pointed out that effects in any of these three regions would be compatible with the present paper’s definition of he pre-head attachment parsing hypothesis (see the Introduction), further empirical specification of the attachment point would certainly play a helpful role in clarifying the details of pre-head attachment mechanisms. Moreover, the uncertainty could make some variants of a head-driven parser consistent with the findings on certain assumptions. Specifically, one could posit a head-driven parser (a) that makes no attachment decision for the dative NP prior to he
subordinate verb, and (b) that, when the verb is encountered, decides not to attach the NP to the verb and to leave it unattached instead (see the Introduction). This variant of a head-driven parser can be reconciled with the data if we make the additional assumption that attaching a constituent to an already-closed clause (e.g., the subordinate clause in the structure under investigation) is significantly more costly than attaching it to a currently-open clause (e.g., the matrix clause) (this account was suggested to us by Altmann, pers. comm.). The second variant of head-driven parsing that could account for the data is that the parser makes a low attachment decision at the subordinate verb but reanalyses it to the alternative (high) attachment option when the matrix verb appears—perhaps because the order of “NP-Nom + NP-Dat”, as in high attachment, is canonical in Japanese whereas the order of “NP-Dat + NP-Nom” as in low attachment is non-canonical (this account was suggested to us by Kempen, pers. comm.). The observed processing cost at the matrix verb in the LA condition can be explained on the assumption that cost of making an attachment to a previously-closed clause is higher than the cost of standard reanalysis in the GA condition. Moreover, this variant of head-driven parser ought to assume that there analysis in the GA condition costs as little as in the HA condition where there should be no reanalysis at the matrix verb. Such variation in attachment costs is completely acceptable within the framework of the U-Space model (Vosse & Kempen, submitted) since, within this framework, certain attachment decisions can be reversed at a cost depending on the strength of information in support of the link (Kempen, pers. comm.). Although these two variants of the head-driven hypothesis are consistent with our results, there does not seem to be any motivation (other than that of explaining our data) for a head-driven parser to behave in this way. 

A second possible criticism to the present study may be that the experimental manipulation may involve a potential problem caused by the fact that the data points come at the end of a clause. In Experiment 2, there were two critical data points: the subordinate verb and the matrix verb. A possible criticism is that an end-of-clause critical region might not reflect only structural/lexical effects but also some others such as wrap-up effects. However, it should be noted that it is almost impossible to avoid having a critical region at the end of a clause in Japanese experiments of this kind. Moreover, even if the present results reflect wrap-up effects, it must still be explained why these wrap-up effects cause the differences we found.

**GENERAL DISCUSSION**

The experiments used in the present study were based on the Japanese ambiguity in which a dative NP can be attached to either one of two verbs:
a verb within a subordinate clause or a matrix verb. Using the only globally ambiguous version of this sentence structure (as in (6)), Experiment 1 revealed that subjects showed a very strong tendency to interpret the dative NP as being attached to the matrix verb (and not, therefore, to its subordinate competitor) (this interpretation also dominated over various other potential readings of the material, such as interpretations in which the NP was somehow “shared” by the two verbs). Experiment 2 confirmed this bias in a self-paced reading study. Where the thematic properties of the matrix verb forced the dative NP to be tied in to the alternative host, the reading time for the verb was raised relative to that in the globally ambiguous control condition (which presumably forced neither matrix nor subordinate verb attachment). In contrast there was no corresponding processing load when the combined thematic properties of the two verbs together forced the dative attachment to the matrix host. This is taken to indicate that what was being forced here was precisely the commitment that had already been made prior to this point, and as a matter of preference during earlier phases in the processing of the sentence. Nor can any of these biases plausibly be attributed to systematic differences in the dative NP attracting properties of the individual verbs themselves or to differences in their overall familiarities. Prior measurements of such tendencies (from the Preliminary Test and the familiarity-rating questionnaire in the Method Section of Experiment 2) had been used to equate these properties of the matrix and subordinate verbs in the main study.

In short, the data provided persuasive evidence that, in this Japanese ambiguity, the dative NP is preferentially attached to the matrix verb (high attachment) rather than the subordinate verb (low attachment).

As spelt out in the Introduction, this bias is in line with the predictions of pre-head attachment models and contrary to the predictions of head-driven accounts. In pre-head attachment accounts, a dative NP is attached to the only verb-structure that can confidently be projected even when the verb itself has not yet been processed—namely the structure associated with the matrix verb that has to turn at some subsequent point in the sentence. Such a link-up is compatible with the data in the sense that subsequent processing is rapid and unproblematic in all conditions where the thematic properties of the matrix verb are compatible with this arguably premature decision. Where there is conflict, however, there is also an increase in processing time—as would be expected if the preliminary analysis were being reviewed and altered, shifting the dative NP from (say) its initial slot within matrix argument structure to its final place in the subordinate verb structure.

In contrast with this, the results run contrary to the predictions of head-driven accounts. The increased processing costs associated with forcing subordinate verb attachment are not readily explained in accounts in which
this interpretation is the initial commitment. According to Pritchett’s model (Pritchett, 1991, 1992), the parser should attach the dative NP to the first verb when it is encountered (i.e., the subordinate verb). This follows from the Generalized Theta Attachment principle (Pritchett, 1992) in that this particular link-up satisfies the Theta Criterion for the dative NP by locating a verb that offers it an acceptable thematic role, thereby ending the brief interval during which it is left unattached. If this is indeed what happened in the course of processing, it is not at all obvious why the subsequent reading latencies should show that it is apparently costly to retain this linkage (in the condition forcing subordinate verb attachment), while breaking and reversing the link is cost-free (i.e., at this point there are two ways where matrix attachment is forced).

In short, the results systematically and consistently support pre-head attachment models and run against the prediction of head attachment accounts.

These conclusions are consistent with those reached on the basis of German data (Bader & Lasser, 1994; Scheepers et al., 1994). However, unlike the German studies the present results could not have been distorted by the potential influence of a non-syntactic mechanism for assigning case to otherwise ambiguous NPs. To this extent the current findings go beyond the earlier data and strengthen the findings against head-driven models.

Overall, then, the present findings seem to provide clear evidence against head-driven accounts of parsing—at least as applied to the syntactic analysis of the Japanese language. While this clearly implies that such models cannot provide a universal account of human parsing, what it does not do, of course, is establish that pre-head driven parsing is necessarily any more general than head-driven parsing. It could be that the parsing machinery is differentially parameterised in different languages in such a way that the system is configured (say) as a head-driven system when it is called upon to handle head-initial languages, whereas it takes the alternative pre-head driven form when dealing with head-final languages like Japanese. The question of whether pre-head driven accounts apply universally or whether they are restricted to a subset of languages, cannot be resolved without further research designed to examine the status of the two competing accounts in a range of different languages.

In the meantime, the present study not only raises questions about strictly head-driven accounts. It also highlights imperfections in approaches which, while not purely head-driven, place very strong emphasis on those aspects of syntactic analysis that are contingent upon the prior use of such head information (e.g., the model proposed by MacDonald et al., 1994). One consequence of this heavy emphasis on head-based processing is that such models have neglected to provide well fleshed-out accounts of
the various kinds of syntactic processing that must occur in the absence of prior lexical head information. Detailed accounts of processes that occur once the verb has been read are not typically matched by clear statements about the kinds of operations that might occur prior to the appearance of a verb in languages like German, Korean and Japanese. Arguably, viable constraint-based models could be developed for this purpose—perhaps by appealing to the “emergent” biases derived from past contact with individual verbs (e.g., by following the kinds of theoretical moves set out by Tabor, Juliano, & Tanenhaus, 1997). On such accounts it may well be feasible to tap into the emergent biases of generic verb representations well in advance in specific sentences of the kind examined here. The averaged thematic properties of such schematic representations might conceivably offer mechanisms for explaining the data—namely, in this particular case, the evidence for early matrix verb attachment. However, as far as we are aware, such extensions to the constraint-based account have not yet been explicitly developed by proponents of these theories. In the absence of such explicit refinements, it is difficult to assess the prospects constraint-based models may have in accounting for argument attachment biases in head-final languages.

REFERENCES


APPENDIX 1: RATING RESULTS OF PRELIMINARY TEST

The following table shows the results of the preliminary test. Low rating scores indicate that the verbs require a dative Goal argument strongly. The English translation of the results can be obtained from the first author.

<table>
<thead>
<tr>
<th>VERB</th>
<th>RATE</th>
<th>POSITION</th>
<th>VERB</th>
<th>RATE</th>
<th>POSITION</th>
</tr>
</thead>
<tbody>
<tr>
<td>望む</td>
<td>66.40</td>
<td>S</td>
<td>表現する</td>
<td>71.71</td>
<td>S</td>
</tr>
<tr>
<td>動める</td>
<td>68.10</td>
<td>S</td>
<td>説く</td>
<td>71.73</td>
<td>M</td>
</tr>
<tr>
<td>残す</td>
<td>68.15</td>
<td>S</td>
<td>知らせる</td>
<td>71.79</td>
<td>M</td>
</tr>
<tr>
<td>賦する</td>
<td>68.35</td>
<td>M</td>
<td>見せる</td>
<td>71.84</td>
<td>M</td>
</tr>
<tr>
<td>落とす</td>
<td>68.48</td>
<td>S</td>
<td>紹介する</td>
<td>71.88</td>
<td>S</td>
</tr>
<tr>
<td>質問する</td>
<td>68.49</td>
<td>S</td>
<td>伝える</td>
<td>71.99</td>
<td>M</td>
</tr>
<tr>
<td>説明する</td>
<td>68.78</td>
<td>M</td>
<td>語る</td>
<td>72.04</td>
<td>M</td>
</tr>
<tr>
<td>向ける</td>
<td>69.00</td>
<td>M</td>
<td>送る</td>
<td>72.07</td>
<td>S</td>
</tr>
<tr>
<td>ご意見を述べる</td>
<td>69.04</td>
<td>M</td>
<td>約束する</td>
<td>72.08</td>
<td>S</td>
</tr>
<tr>
<td>述ぶ</td>
<td>69.10</td>
<td>M</td>
<td>保証する</td>
<td>72.10</td>
<td>M</td>
</tr>
<tr>
<td>貸す</td>
<td>69.25</td>
<td>S</td>
<td>報告する</td>
<td>72.17</td>
<td>M</td>
</tr>
<tr>
<td>手渡す</td>
<td>69.27</td>
<td>M</td>
<td>支給する</td>
<td>72.53</td>
<td>S</td>
</tr>
<tr>
<td>尋ねる</td>
<td>69.74</td>
<td>S</td>
<td>通じる</td>
<td>72.84</td>
<td>M</td>
</tr>
<tr>
<td>言う</td>
<td>69.84</td>
<td>M</td>
<td>通知する</td>
<td>72.84</td>
<td>M</td>
</tr>
<tr>
<td>あげる</td>
<td>70.11</td>
<td>S</td>
<td>任せる</td>
<td>72.85</td>
<td>M</td>
</tr>
<tr>
<td>預ける</td>
<td>70.40</td>
<td>S</td>
<td>払う</td>
<td>72.85</td>
<td>M</td>
</tr>
<tr>
<td>渡す</td>
<td>70.42</td>
<td>M</td>
<td>請う</td>
<td>73.01</td>
<td>S</td>
</tr>
<tr>
<td>許可する</td>
<td>70.64</td>
<td>M</td>
<td>提案する</td>
<td>73.02</td>
<td>S</td>
</tr>
<tr>
<td>売る</td>
<td>71.01</td>
<td>M</td>
<td>申請する</td>
<td>73.52</td>
<td>S</td>
</tr>
<tr>
<td>与える</td>
<td>71.12</td>
<td>M</td>
<td>配る</td>
<td>73.62</td>
<td>S</td>
</tr>
<tr>
<td>述べる</td>
<td>71.33</td>
<td>S</td>
<td>投げる</td>
<td>73.70</td>
<td>M</td>
</tr>
<tr>
<td>処方する</td>
<td>71.42</td>
<td>M</td>
<td>言う</td>
<td>74.43</td>
<td>S</td>
</tr>
<tr>
<td>話す</td>
<td>71.44</td>
<td>M</td>
<td>支持する</td>
<td>74.87</td>
<td>S</td>
</tr>
<tr>
<td>読む</td>
<td>71.56</td>
<td>S</td>
<td>求める</td>
<td>76.17</td>
<td>S</td>
</tr>
</tbody>
</table>

Note: S, used as a Subordinate Verb; M, used as a Matrix Verb.
APPENDIX 2: EXPERIMENTAL MATERIALS FOR EXPERIMENT 2

The following shows experimental sentences for each condition in Experiment 2. The GA condition was also tested in Experiment 1. “/” indicates the segmentation pattern. All sentences were presented in one line. The English translation of the materials can be obtained from the first author.

1: GA 教授が/学生に/図書館司書が/貸した/珍しい古文書を/見せた 1: HA 教授が/学生に/図書館司書が/破った/珍しい古文書を/見せた 1: LA 教授が/学生に/図書館司書が/貸した/珍しい古文書を/破った


13: GA 智恵が、従姉妹に、両親が、尋ねた、新居での暮らしを、知らせた
13: HA 智恵が、従姉妹に、両親が、想像した、新居での暮らしを、知らせた
13: LA 智恵が、従姉妹に、両親が、尋ねた、新居での暮らしを、想像した

14: GA 大旦那が、番頭に、若旦那が、述べた、斬新なアイデアを、伝えた
14: HA 大旦那が、番頭に、若旦那が、実現した、斬新なアイデアを、伝えた
14: LA 大旦那が、番頭に、若旦那が、述べた、斬新なアイデアを、実現した

15: GA 兄が、元恋人に、叔父が、送った、とりたてのカニを、ご馳走した
15: HA 兄が、元恋人に、叔父が、嫌った、とりたてのカニを、ご馳走した
15: LA 兄が、元恋人に、叔父が、送った、とりたてのカニを、嫌った

16: GA 義姉が、子供に、来客が、残した、チョコレートケーキを、与えた
16: HA 義姉が、子供に、来客が、賞賛した、チョコレートケーキを、与えた
16: LA 義姉が、子供に、来客が、残した、チョコレートケーキを、賞賛した

17: GA 部長が、係長に、事務員が、申請した、10日間の休暇を、許可した
17: HA 部長が、係長に、事務員が、忘れていた、10日間の休暇を、許可した
17: LA 部長が、係長に、事務員が、申請した、10日間の休暇を、忘れていた

18: GA 未亡人が、子供に、死んだ夫が、もたらした、耐え難い不幸を、語った
18: HA 未亡人が、子供に、死んだ夫が、克服した、耐え難い不幸を、語った
18: LA 未亡人が、子供に、死んだ夫が、もたらした、耐え難い不幸を、克服した

19: GA 大学病院が、受付に、患者が、求めた、人間ドックの結果を、通知した
19: HA 大学病院が、受付に、患者が、憂いだ、人間ドックの結果を、通知した
19: LA 大学病院が、受付に、患者が、求めた、人間ドックの結果を、憂いだ

20: GA 物理学者が、理科教師に、生徒の一人が、質問した、ピックバンの原因を、説いた
20: HA 物理学者が、理科教師に、生徒の一人が、知っていた、ピックバンの原因を、説いた
20: LA 物理学者が、理科教師に、生徒の一人が、質問した、ピックバンの原因を、知っていた

21: GA 直子が、美樹に、涼子が、表現した、本当の気持ちを、言った
21: HA 直子が、美樹に、涼子が、尊重した、本当の気持ちを、言った
21: LA 直子が、美樹に、涼子が、表現した、本当の気持ちを、尊重した

22: GA 洗濯業者が、新人工員に、工場が、支給した、地味な作業服を、渡した
22: HA 洗濯業者が、新人工員に、工場が、汚した、地味な作業服を、渡した
22: LA 洗濯業者が、新人工員に、工場が、支給した、地味な作業服を、汚した

23: GA 祐也が、成美に、正樹が、望んだ、熱いまなざしを、向けた
23: HA 祐也が、成美に、正樹が、笑った、熱いまなざしを、向けた
23: LA 祐也が、成美に、正樹が、望んだ、熱いまなざしを、笑った

24: GA 主婦が、郵便局に、郵政省が、指示した、新規定の手数料を、払った
24: HA 主婦が、郵便局に、郵政省が、心配した、新規定の手数料を、払った
24: LA 主婦が、郵便局に、郵政省が、指示した、新規定の手数料を、心配した