Distinguishing theories of syntactic expectation cost in sentence comprehension: Evidence from Japanese*

KENTARO NAKATANI AND EDWARD GIBSON

Abstract

Previous research in the sentence comprehension literature has established that people expend resources keeping track of partially processed phrase structures during the process of comprehending sentences. An open question in this literature has been what units of syntactic expectation cost the human parser utilizes. Two viable options from the literature are (1) incomplete syntactic dependencies; and (2) predicted syntactic heads. This article provides a self-paced reading experiment from Japanese — a head-final language — that tests the incomplete dependency hypothesis. The materials in the current experiment manipulate the number of dependents of an upcoming verb, by manipulating (1) the presence/absence of a locative postpositional phrase modifier of the verb and (2) the presence/absence of a dative argument of the verb. The results failed to show any support for the incomplete dependency hypothesis, but were completely consistent with the predictions of the predicted head hypothesis. Taken with the results from the literature, these results offer support for the predicted head hypothesis.

1. Introduction

Current research in sentence comprehension has established that numerous factors affect the moment-by-moment interpretation of a sentence (Tanenhaus and Trueswell 1995; Gibson and Pearlmutter 1998). These factors include (1) the lexical frequencies of the words involved (MacDonald et al. 1994; Trueswell 1996); (2) the working memory resources that are needed to retain the current structure and integrate the upcoming words (Frazier 1979, 1987; Gibson 1998, 2000); (3) the plausibility in the world of the interpretation of these structures (Trueswell et al. 1994; Garnsey et al. 1997); (4) the discourse context in which the sentence is

Linguistics 46–1 (2008), 63–86 0024–3949/08/0046–0063
DOI 10.1515/LING.2008.003

produced (Altmann and Steedman 1988; Tanenhaus et al. 1995); and (5) the intonational properties of auditory sentences (Cutler et al. 1997; Watson and Gibson 2004). It is possible to study the effects of these different factors using either unambiguous or ambiguous sentence materials (Gibson 1991, 1998). In unambiguous materials, more complex materials give rise to slower reaction times, for example, than less complex materials. In ambiguous materials, people prefer a less complex interpretation over a more complex one.

In this article we will focus on properties of the syntactic structure of a sentence that consume working memory resources. Furthermore, we will restrict our attention to the processing of sentences independent of ambiguity, as much as possible. One type of contrast from the literature that is highly informative in the area of syntactic complexity is the contrast between NESTED (or CENTER-EMBEDDED) structures and their right- or left-branching counterparts (Yngve 1960; Chomsky and Miller 1963). For example, the right-branching English structure in (1a) is easier to understand than the nested structure in (1b), and the left-branching Japanese structure in (2a) is easier to understand than its nested version in (2b) (from Nakatani and Gibson 2003):

(1) a. Mary met the senator who attacked the reporter who ignored the president.
    b. #The reporter who the senator who Mary met attacked ignored the president.

(2) a. [syusyoo-ga utatanesita to] [syoki-ga koogisita to] [daigisi-ga hookokusita]
    [prime-minister-NOM dozed COMP] [Diet-member-NOM protested COMP] [secretary-NOM reported]
    'The secretary reported that the Diet-member protested that the prime minister dozed.'
    b. #[syoki-ga [daigisi-ga [syusyoo-ga utatanesita to] koogisita to] hookokusita]
    [secretary-NOM [Diet-member-NOM [prime-minister-NOM dozed COMP] protested COMP] reported]
    'The secretary reported that the Diet-member protested that the prime minister dozed.'

The difficulty of understanding nested structures cannot be due to the lexical content in the sentences or the meaning of the resultant propositions, because each nested sentence has the same propositional content and lexical items as its right- or left-branching counterpart. Furthermore, the nested versions have no more temporary ambiguity than their right- or left-branching controls, so the difficulty in understanding them does not have to do with ambiguity.

One framework that has been proposed to account for nested vs. non-nested contrasts is the dependency locality theory (DLT) (Gibson 1998, 2000). According to the DLT, there are two components of syntactic and semantic structure that consume working memory resources when comprehending a sentence: (a) integrating incoming words into the structure that has been built thus far; and (b) storage of expectations of up-coming syntactic heads. According to the integration component of the DLT, the difficulty of integrating a new word *w* to a syntactic head *h* in the current structure is dependent on the linear distance between *w* and *h* in terms of some function of the number of words (Hawkins 1994), the complexity of the discourse structure (Gibson 1998, 2000), the discourse accessibility of the types of NPs in the interim material (Warren and Gibson 2002) and/or the number of interfering similar NPs (Gordon et al. 2001). Processing a nested structure consumes more integration resources than processing a non-nested structure because the dependencies between words are much longer on average in nested structures than in non-nested structures, no matter what the distance metric.

For example, the integrations at the verbs in (1b) are all more complex than in (1a). Consider the verb *met* in (1b). This verb integrates with the previous NP *Mary* and the wh-pronoun *who* mediated by a wh-trace in the object position of *met*. In contrast there is only one integration at the point of processing *met* in (1a): connecting the verb to the preceding subject NP *Mary*. The integration cost difference is even greater at the verb *attacked* across the two structures. In (1a), there is a single local integration between the verb *attacked* and the preceding pronoun *who*. In (1b), the verb *attacked* is integrated with (a) its subject *the reporter*, which is a clause back in the input string and (b) with the wh-pronoun *who* (mediated by a wh-trace), which is also a clause back in the input string.

The processing results in the literature from a head-initial language like English are consistent with at least two kinds of integration processes: (a) a bottom-up integration process, such that integration consists of connecting a new word in the input to a position in the current structure by consulting the grammatical rules of the language (see [3] below); and (b) a top-down integration process, such that syntactic integration consists of matching the syntactic predictions derived from the current words in the input and the grammatical rules of the language (see [4] below). Although the English results are consistent with both (3) and (4), results from processing head-final languages like German, Japanese and Hindi (Konieczny 1996, 2000; Konieczny and Döring 2003; Nakatani and Gibson 2003;

Vasishth 2003) strongly suggest that the processor is top-down, anticipating upcoming elements, as in (4) rather than (3).

(3) *Bottom-up head-dependent distance hypothesis:*
The difficulty of integrating a new word *w* into the current structure depends on the distance back to the head *h* to which *w* connects.

(4) *Top-down incremental narrowing of predictions:*
Syntactic predictions are continually narrowed as new words in a sentence are processed (cf. the *anticipation* hypothesis of Konieczny 1996, 2000; Konieczny and Döring 2003; Vasishth 2003). The difficulty of narrowing or matching a syntactic prediction *P* when processing a new word *w* depends on the distance back to the last time that *P* was worked on in the current structure.

Consider the bottom-up and top-down hypotheses with respect to the sentences in (2) above. (2b) is nested, with one clause within another. It is therefore more complex than (2a), which is non-nested. But there is little reading time difficulty on the verbs in either (2a) or (2b) (Nakatani and Gibson 2003; cf. similar results from Konieczny [2000] for German and Vasishth [2003] for Hindi). Note that the bottom-up head-dependent distance hypothesis predicts slower RTs for longer connections between the subjects and their verbs, as in the English sentences. But no such effect occurs. The top-down incremental narrowing of predictions hypothesis in (4) is consistent with the lack of increased RTs on the verbs in a sentence like (2b), because the verbs are syntactically predictable by the occurrence of the preceding nominative subjects and verbs. In particular, the first nominative subject is consistent with almost any verb. The second nominative subject narrows the expectation to be a verb that takes a clausal complement, or possibly a verb that takes a nominative object.[1] The presence of the third nominative NP narrows the expectation even further, towards a verb that takes a clausal complement. These expectations are eventually met as the verbs are processed one at a time. But the distance back to the last point at which the expectations were narrowed for each is small in each case: the immediately preceding word. As a result, RTs are not always longer for longer distance dependencies. Rather, the difficulty of an integration is proportional to the distance back to the last location in the input string that the expectation for that word was narrowed.

Consider now the second component of the DLT, syntactic storage (or expectations). According to the expectation component of the DLT, there is an expectation cost associated with retaining the expectation of each syntactic head that is required to complete a partial input string grammatically. Thus, following the words *the man who the* ... there is a cost of four expectation units, one for each of the following heads: a subject noun for the embedded RC; a verb for the RC; an NP gap position, associated with the RC pronoun *who*; and a main verb for the sentence. In (2b), there is an expectation cost for each of three predicted verbs after processing the most embedded nominative subject.[2] As a consequence of online expectation costs, RTs during regions with more predicted verbs are read more slowly than regions with fewer predicted verbs (Chen et al. 2005; Gibson et al. 2005; Nakatani and Gibson 2003).

Consider some of the evidence for the existence of syntactic expectation costs independent of integration costs. For example, Chen et al. (2005) investigated the processing of embedded English clauses with zero, one or two further verbs pending, as in (5):

(5) a. *0 expected verbs:*
The employee realized that the boss implied that *the company planned a layoff* and so he sought alternative employment.

b. *1 expected verb, late:*
The employee realized that the implication that *the company planned a layoff* was not just a rumor.

c. *1 expected verb, early:*
The realization that the boss implied that *the company planned a layoff* caused a panic.

d. *2 expected verbs:*
The realization that the implication that *the company planned a layoff* was not just a rumor caused a panic.

The critical region in this design consists of the embedded clause *the company planned a layoff*, in italics. Because this clause has the same structure in all conditions, integration costs are identical across the four. In sentence (5a), the critical material *the company planned a layoff* is embedded as the sentential complement of the verb *implied* which is itself part of a clause embedded as the sentential complement of the matrix verb *realized*. Because both verbs *implied* and *realized* are encountered immediately after their respective subject nouns, no additional verbs are expected after the critical embedded clause. In sentence (5b), the verb *implied* is nominalized to *implication* with the result that the critical clause is a sentential complement of the noun *implication*. This change to the embedded subject noun phrase *the implication* results in the requirement for an additional verb following the critical region. Similarly, in sentence (5c) the matrix verb *realized* is nominalized to *realization*, leading to the expectation for an additional verb after the critical region. Finally, in sentence (5d), both the verbs *realized* and *implied* are nominalized and two verbs are therefore required following the critical region. As predicted by syntac-

(AutoPDF V7 24/10/07 13:15) WDG (148×225mm) TimesM J-1850 Linguistics, 46:1 PMU:I(CKN[A])24/10/2007 pp. 63–86 1850_46-1_03 (p. 67)

tic expectation costs, the critical region was read fastest in (5a), slower in (5b) and (5c), and slowest in (5d), with all predicted differences significant.

The point of this article is to attempt to distinguish what kinds of elements the human sentence processor is keeping track of in syntactic expectation. One possibility has been discussed so far: predicted syntactic heads (Gibson 1998, 2000):

(6) *Predicted syntactic head hypothesis:*
The human sentence processor is sensitive to the number of syntactic heads that are required to form a grammatical sentence at each processing state.

The evidence presented thus far is consistent with a narrower hypothesis — predicted verbs (cf. Kimball 1973) — but Chen et al. (2005) provide evidence that more then just predicted verbs are associated with online expectation costs. In particular, Chen et al. showed that the expectation of a wh-trace for a filler is also associated with a processing cost independent of other resource costs (cf. Wanner and Maratsos 1978). (See the general discussion in Section 3 for more on this issue.) Chen, Gibson and Wolf also provide pilot evidence that there is an expectation cost associated with expected prepositional phrase arguments following a verb. Thus it appears that expectation costs are not restricted to verbal expectations.

An alternative to the predicted head hypothesis in (6) is that the human sentence processor is sensitive to the number of incomplete dependencies at a processing state:

(7) *Incomplete dependency hypothesis:*
The human sentence processor is sensitive to the number of partially processed dependencies at each processing state.

Variations of (7) include sensitivity to incomplete thematic role assignments (Hakuta 1981; Gibson 1991), incomplete case-assignment relations (Lewis 1996; Stabler 1994), or partially processed phrase structure rules (Chomsky and Miller 1963).

The predicted head hypothesis and the incomplete dependency hypothesis make the same predictions on all the materials discussed thus far. For example, consider the examples in (5). For each predicted verb, there is a corresponding incomplete subject-verb dependency, in which case and thematic role need to be assigned.

It is difficult to distinguish the predicted head hypothesis from the incomplete dependency hypothesis in a head-initial language like English, because for each predicted head, there is an incomplete dependency. But this is not true in a head-final language like Japanese. In a head-final language, more than one incomplete dependency can be associated with the same predicted head, if they all depend on the same upcoming head. For example, a nominative NP, a locative NP, a dative NP and an accusative NP may all depend on a verb to follow. After processing the fourth of these NPs, there are four incomplete dependencies, one for each NP-verb connection. In contrast, under the predicted head hypothesis, only one head needs to be predicted: a verb that can take the four initial NPs as its dependents.

Notice that for the predicted head hypothesis to be plausible, people must have implicit knowledge of the existence of a variety of different argument structures for verbs, especially the most frequent argument structures. Such an assumption fits well with what we know about sentence comprehension. It is now well established that people are sensitive to fine-grained lexical frequency information, including argument-structure subcategorization information (Trueswell et al. 1993; MacDonald et al. 1994).

Some initial support is provided for the predicted head hypothesis by Miyamoto (2002). Miyamoto reported that people read a clause-initial nominative-nominative sequence slower than either a clause-initial nominative-accusative sequence or a clause-initial accusative-nominative sequence, and attributed this slowdown to the hypothesis that nominative in Japanese induces a clause-boundary. These results are as predicted by the predicted head hypothesis in (6), because a nominative-nominative sequence causes the expectation for two verbs and a linking complementizer to follow, whereas the nominative-accusative or accusative-nominative sequences require only one verb to come.[3] In contrast, these results are not predicted by the incomplete dependency hypothesis in (7). There are the same number of incomplete thematic-role assignments or incomplete dependencies after processing each of the three sequences (two incomplete dependencies in each case), and no differences are predicted. Thus these results seem to provide some initial evidence for the predicted head hypothesis. However, it is possible to account for this observation within an incomplete dependency theory by appeal to a different expectation cost function than has been implicitly assumed thus far. Until now, we have assumed that all predicted heads / incomplete dependencies are weighted equally. In contrast, Lewis (1993, 1996) proposed that syntactic expectation costs are sensitive to similarity, such that retaining similar predicted heads / incomplete dependencies may cause more of a processing load than retaining more distinct elements (cf. Lewis and Nakayama 2002; Uehara and Bradley 2002):

(8) *Interference-based incomplete dependency hypothesis:*
The human sentence processor is sensitive to the number of partially processed dependencies at each processing state.

(AutoPDF V7 24/10/07 13:15) WDG (148×225mm) TimesM J-1850 Linguistics, 46:1 PMU:I(CKN[A])24/10/2007 pp. 63–86 1850_46-1_03 (p. 69)

> Moreover, keeping track of the same kinds of incomplete dependencies is associated with a greater cost.

Applying this idea to Miyamoto's processing evidence, the reason that a nominative-nominative sequence is more complex than a nominative-accusative sequence (or an accusative-nominative sequence) may be that the two incomplete nominative-case / subject dependencies interfere with each other more than in the other two instances.

Intuitions on Japanese examples from Pritchett (personal communication in Gibson 1991), Lewis (1993) and Nakatani et al. (2000) provide similar evidence as Miyamoto's experimental contrasts. For example, structures like (9) (from Lewis 1993) below are intuitively more comprehensible than examples like (2b), in spite of the fact that there is a position in (9) in which there are as many as five incomplete syntactic dependencies:

(9) Taroo-ga Hajime-ni [Akira-ga Hanako-ni Sigeru-o syookai sita to] itta.
Taroo-NOM Hajime-DAT [Akira-NOM Hanako-DAT Sigeru-ACC introduced COMP] said
'Taroo said to Hajime that Akira introduced Shigeru to Hanako.'

After processing the NP *Sigeru-o* ('Shigeru-ACC'), there are five NPs that need to be interpreted by an upcoming verb. This example contrasts with the intuitively more complex (2b), in which there are at most three NPs that need to be interpreted by upcoming verbs. Nakatani et al. (2000) reported the results from an off-line questionnaire study, according to which singly nested structures with five initial NPs like (9) were rated significantly better than the doubly nested structures like (2b) (using the same NPs in each, as much as possible), even though the latter involved fewer initial NPs. These results are straight-forwardly accounted for by a syntactic expectation cost theory that is sensitive to the number of predicted syntactic heads during the processing of a sentence. In processing a sentence like (2b), five syntactic heads are required at the point of processing the most embedded subject NP: three verbs and two complementizers. In contrast, at most three syntactic heads (two verbs, one complementizer) are required during the processing of examples like (9). Thus the predicted head theory is consistent with the observations.

The result runs contrary to the simplified incomplete dependency hypothesis in (7), because the structure with more temporarily incomplete dependencies turns out to be less complex. But with Lewis's assumption of interference between similar incomplete dependencies in (8), the result can be explained. In particular, the reason that examples like (2b) are so difficult is that there are three incomplete nominative-case / subject dependencies, which interfere strongly with each other. There are only at most two incomplete dependencies of the same kind in processing (9) (two incomplete nominative-case / subject dependencies, two incomplete datives), and so this sentence is easier to process.

In summary, the predicted head hypothesis in (6) is consistent with processing data from both English and Japanese, using either a similarity-based interference metric as proposed by Lewis (1996), or using some other metric, such as a linear metric. The crosslinguistic evidence is also consistent with a theory that is sensitive to incomplete dependencies, but only when a similarity-based interference metric of syntactic expectation cost is used, as in (8). A study using Japanese materials was designed to test the incomplete dependency hypothesis further.

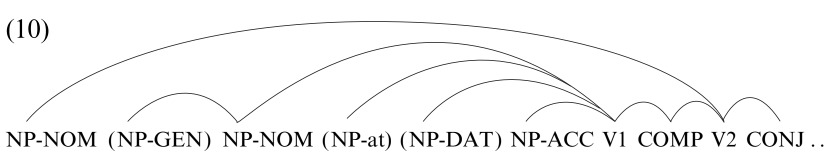
2. Experiment

The current experiment was designed to test predictions of the incomplete dependency expectation cost hypotheses, in (7) and (8). The processing of head-final languages like Japanese offers a potential way to distinguish this hypothesis from the predicted head hypothesis in (6). A strong test of the incomplete dependency hypothesis can be constructed in a head-final language by comparing structures for which the number of predicted heads is the same, but the number of incomplete dependencies differs. This is the design of the current experiment.

2.1. *Materials*

Six target conditions were prepared in a 2×3 design, crossing the presence of a dative argument (+Dative, −Dative) with the presence of a locative NP (+Locative-Adverbial [= NP-at], +Locative-Adnominal [= NP-GEN], −Locative). The form of the items is presented in (10). A full example item is presented in (11).

(10)



(11) a. +Dative, +Locative-Adverbial
NP-NOM [NP-NOM *NP-at NP-DAT* NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [sin'nyuusyain-ga *zimusyo-de kokyaku-ni* tyuumonsyo-o hassoosita to] dentatusita ato ...
telephone_receptionist-NOM [freshman-NOM *office-at client-DAT* order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman had sent the order sheet to the client while in the office, ...'

b. +Dative, +Locative-Adnominal
NP-NOM [*NP-GEN* NP-NOM *NP-DAT* NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [*zimusyo-no* sin'nyuusyain-ga *kokyaku-ni* tyuumonsyo-o hassoosita to] dentatusita ato ...
telephone_receptionist-NOM [*office-GEN* freshman-NOM *client-DAT* order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman in the office had sent the order sheet to the client, ...'

c. +Dative, −Locative
NP-NOM [NP-NOM *NP-DAT* NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [sin'nyuusyain-ga *kokyaku-ni* tyuumonsyo-o hassoosita to] dentatusita ato ...
telephone_receptionist-NOM [freshman-NOM *client-DAT* order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman had sent the order sheet to the client, ...'

d. −Dative, +Locative-Adverbial
NP-NOM [NP-NOM *NP-at* NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [sin'nyuusyain-ga *zimusyo-de* tyuumonsyo-o hassoosita to] dentatusita ato ...
telephone_receptionist-NOM [freshman-NOM *office-at* order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman had sent the order sheet while in the office, ...'

e. −Dative, +Locative-Adnominal
NP-NOM [*NP-GEN* NP-NOM NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [*zimusyo-no* sin'nyuusyain-ga tyuumonsyo-o hassoosita to] dentatusita ato ...

telephone_receptionist-NOM [*office-GEN* freshman-NOM order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman in the office had sent the order sheet, ...'

f. −Dative, −Locative
NP-NOM [NP-NOM NP-ACC V1 COMP] V2 CONJ ...
denwaban-ga [sin'nyuusyain-ga tyuumonsyo-o hassoosita to] dentatusita ato ...
telephone_receptionist-NOM [freshman-NOM order_sheet-ACC sent COMP] told after ...
'After the telephone receptionist told (somebody) that the freshman had sent the order sheet, ...'

As shown in (10), each condition consisted of the critical clause embedded within an outer clause. The outer clause was the same across all conditions (e.g., 'telephone-receptionist ... told' in [11]). The primary reason for using nested items was that all theories predict a slowdown effect on the embedded clause, at the presence of a second nominative NP. Finding such a slowdown would replicate earlier results, and provide a baseline of processing difficulty for other comparisons.

The manipulations across conditions occurred in the embedded clause. In the +Dative conditions, a dative argument of the verb was included, 'client-DAT' in (11a)–(11c). This argument was optional for all the verbs used in the items, as it always is in Japanese for dative arguments. In the +Locative-Adverbial conditions, a locative NP modified the embedded verb, 'office-at' in (11a) and (11d). In the +Locative-Adnominal conditions, the same locative NP modified the embedded subject NP, 'office-GEN' in (11b) and (11e). In the −Locative conditions (11c) and (11f), there was no locative NP. Note that, unlike in English, there is no ambiguity of attachment site for the locative in either the VP or NP modification, so ambiguity of attachment site is not a confounding factor here.[4] Each condition was further embedded as an adjunct clause (either a *because*, *when*, or *after* clause), in order to avoid wrap-up effects in online reading for the embedded verb. A full list of the materials that were used in this experiment is included in the Appendix.

2.2. *Predictions*

All syntactic expectation hypotheses predict that reading times (RTs) at the second nominative NP should be longer than on the first nominative NP. The predicted head hypothesis in (6) makes this prediction because two verbs and a complementizer are predicted following the second nominative NP, compared to only a verb being predicted after the first NP. The incomplete dependency hypotheses in (7) and (8) also make this prediction, because there are two incomplete dependencies following the second NP, compared to only one following the first. Lewis's similarity-based interference hypothesis (8) predicts a strong effect here, because the incomplete dependencies are of the same type.

The conditions differed in terms of the number of incomplete dependencies on the embedded verb, whereas the number of predicted heads before this verb was constant across the conditions. Therefore, the incomplete dependency hypotheses in (7) and (8) on the one hand, and the predicted head hypothesis in (6) on the other hand, make different predictions at the nominal positions before the embedded verb, particularly the accusative NP before the verb, which is present in all six conditions. First consider the +/−Dative factor. The incomplete dependency hypotheses predicts that RTs on the accusative NP should be slower in the +Dative conditions than in the −Dative conditions, because of the extra dependency on the verb in the +Dative conditions. Second, consider the Locative factor. The incomplete dependency hypotheses predicts that RTs on the accusative NP should be slower in the +Locative-Adverbial conditions than in the −Locative or the +Locative-Adnominal conditions, because of the extra dependency on the verb in the +Locative-Adverbial conditions.

If we make comparisons among all six conditions in terms of the number of incomplete dependencies at the preverbal position, (11a) should be the hardest, with five incomplete dependencies before the embedded verb (two nominative NPs, one locative PP, one dative, one accusative), (11e) and (11f) should be the easiest, with three incomplete dependencies (two nominatives, one accusative), and (11b), (11c) and (11d) should be in between, with four incomplete dependencies each. On the other hand, the predicted head hypothesis predicts no difference at the accusative position among any of the conditions, because all conditions have the same number of predicted heads at the preverbal position: two verbs and a complementizer.

The materials in the experiment can also be evaluated with respect to the theories of integration in (3) and (4). If integration is top-down as proposed in (4), as in previous results for head-final languages (Konieczny 1996, 2000; Nakatani and Gibson 2003; Vasishth 2003), then no differences are predicted at the embedded verb position, because the last head to be encountered before the verb is a dependent of the verb in all conditions. This is a local integration in all cases. In fact, Konieczny (1996) and Konieczny and Döring (2003) propose that people will process a word that has more predictive elements associated with it *faster* than a word with fewer predictive elements preceding it. This hypothesis predicts faster reading times at the verb for verbs with more dependents, in direct contrast to the bottom-up integration hypothesis for these materials. In support of the top-down (anticipation) hypothesis, Konieczny and Döring (2003) provide eye-tracking evidence from reading verb-final constructions in German which are very similar to the Japanese materials that were investigated here.

In contrast, if integration is bottom-up as proposed in (3), then the RTs at the verbs should vary according to (1) how many incomplete dependencies are satisfied at the verb positions and (2) how far apart the verb is from its dependent in each case. For example, in (11a), four of the five incomplete dependencies are completed at the embedded positions, with three embedded NPs and one PP being linked to the embedded verb. Three integrations are established at the embedded verb in (11b)–(11d), and two in (11e) and (11f). Therefore, according to the bottom-up head-dependent distance hypothesis, the reaction times at the embedded verb should be the largest in (11a) and the smallest in (11e) and (11f), with (11b)–(11d) in between. Furthermore, the reaction times for verbs are predicted to be slower than their arguments.

2.3. *Procedure*

The experiment was conducted using a self-paced moving-window paradigm (Just et al. 1982) in Linger 1.7, a sentence processing experimental presentation program written by Douglas Rohde, using Apple PowerBook computers on Mac OS X. Each trial began with a series of dashes marking the length and position of the words in the sentences (presented in kanji), printed approximately a third of the way down the screen. Participants pressed the spacebar to reveal each word of the sentence. Bound morphemes such as case markers and complementizers were grouped with preceding words such as nouns and verbs. As each new word appeared, the preceding word disappeared. The amount of time the participant spent reading each word was recorded as the time between key-presses. After the final word of each item, a comprehension question appeared which asked about information contained in the preceding sentence. Participants pressed one of two keys to respond "yes" or "no." After an incorrect answer, the word "INCORRECT" (in Japanese) flashed briefly on the screen. No feedback was given for correct responses. Participants were asked to read sentences at a natural rate and to be sure that they understood what they read. They were told to answer the questions as quickly and accurately as they could and to take wrong answers as an indication to read more carefully. Before the main experiment, a short list of practice items was presented in order to familiarize the participant with the task.

The 24 sets of 6 target conditions described above were distributed in a Latin Square design, resulting in 6 lists. 68 filler items were added to each list. The 92 sentences in a list were shuffled so that they were presented in a different pseudo-random order for each participant, such that no two target items were presented consecutively. 45 adult native speakers of Japanese in the Boston area participated in the experiment. They were each paid five dollars for participation in the experiment, which took about 20 minutes per session.

2.4. *Analysis*

Analyses were conducted on comprehension question response accuracies and reading times. To adjust for differences in word length across word positions as well as overall differences in participants' reading rates, a regression equation predicting reading time from word length (in terms of number of characters) was constructed for each participant, using all filler and experimental items (Ferreira and Clifton 1986; see Trueswell et al. 1994, for discussion). At each word position, the reading time predicted by the participant's regression equation was subtracted from the actual measured reading time to obtain a residual reading time. We excluded the data from one participant whose residual reading times were extremely slow (2.5 standard deviations away from the average). To remove outlier data points, the residual reading times were trimmed so that data points beyond three standard deviations from the relevant condition and position cell mean were discarded, corresponding to less than 1.8% of the data. The reading time data that we report corresponds to all the remaining data, whether or not the comprehension questions were answered correctly. The analyses without the incorrect responses showed the same statistical patterns. The analyses on raw reading times also show the same patterns, although not all comparisons reached significance.

2.5. *Results*

2.5.1. *Comprehension performance.* The comprehension question response accuracy rate for each condition is summarized in Table 1. These data were analyzed using a 2×3 ANOVA, revealing no significant difference between the +Dative and −Dative factors ($Fs < 1.2$, $ps > .28$)

Table 1. *Mean comprehension question accuracy rates with standard errors in parentheses*

	+Locative-Adverbial	+Locative-Adnominal	−Locative
+Dative	(a) 86.63% (2.60)	(b) 75.00% (3.31)	(c) 73.10% (3.40)
−Dative	(d) 75.29% (3.31)	(e) 74.12% (3.37)	(f) 76.30% (3.24)

or among the three Locative factors ($F1(2,86) = 2.84$, $p > .06$; $F2(2,46) = 1.1$, $p > .34$).

2.5.2. *Reading times.* The mean residual reading time data are summarized in Figure 1. Table 2 presents mean residual and raw times per word for all conditions.

In a 2×3 ANOVA we conducted, there were no effects of the Dative or Locative factors at the first nominative NP ($Fs < 3$). This is unsurprising, because the words are identical in all conditions in this region. At the second nominative NP, optional Locative-Adnominal NP is the only preceding optional word. In the single-factor ANOVA we conducted (Locative-Adnominal: present vs. absent), there was a tendency that the conditions with Locative-Adnominal were slower here, but the effect did not reach statistical significance ($F1(1,43) = 3.41$, $p > .07$; $F(1,23) = 3.56$, $p > .07$).

Next, we compared RTs for the first nominative NP to those for the second nominative NP. This comparison revealed that the embedded nominative NP was read significantly more slowly than the initial nominative NP position (Embedded nominative: 247.65 ms (SE 18.46) vs. Initial nominative: 0.66 ms (SE 9.73); $F1(1,43) = 62.15$, $p < .001$; $F2(1,23) = 33.60$, $p < .001$), conforming to the expectation hypotheses in general. However, after the embedded nominative position, participants tended to speed up over the sentences: the mean RTs across the conditions at any of the five regions following the embedded nominative were significantly faster than the RT at the embedded nominative ($F1 > 17$ for all comparisons, $ps < .001$; $F2 > 9$ for all comparisons, $ps < .01$). This is a typical reading time profile for a sentence: As people get more discourse context, they can generally read faster. What is most interesting is that the participants slowed down on the second nominative NP, which was either the second or the third word in the sentence.

At the dative NP region (Region 5 in Figure 1), which was only present in the +Dative conditions (11a)–(11c), the +Locative-Adverbial condition (11a) was read significantly faster than the other two together (−38.34 ms (SE 25.86) vs. 87.31 ms (SE 32.29); $F1(1,43) = 7.13$, $p < .02$; $F2(1,23) = 9.71$, $p < .01$). A direct comparison between (11a)

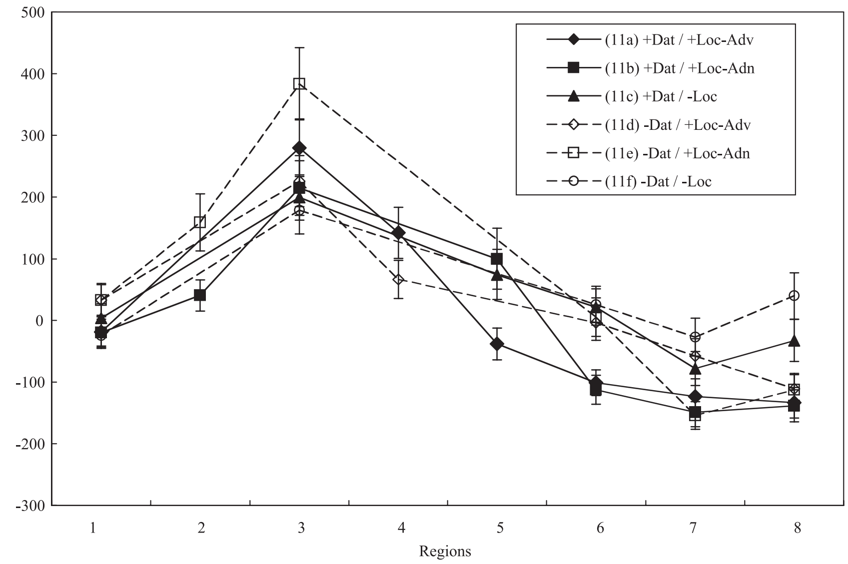


Figure 1. *Mean residual reading times and standard errors in all conditions*

Table 2. *Mean residual RTs in msec for all positions, in all conditions (raw RTs in parentheses)*

	1 NP-NOM	2 (NP-GEN)	3 NP-NOM	4 (NP-at)	5 (NP-DAT)	6 NP-ACC	7 V-COMP	8 V-CONJ
+Dat	−19.16		279.97	142.16	−38.34	−101.1	−123.46	−133.66
+Loc-Adv	(735.3)		(1062.15)	(887.59)	(711.54)	(656.41)	(698.04)	(736.56)
+Dat	−19.97	40.44	214.69		99.994	−112.62	−149	−138.67
+Loc-Adn	(737.51)	(801.35)	(1004.73)		(855.33)	(661.57)	(679.18)	(709.71)
+Dat	3.29		199.43		74.49	21.33	−78.2	−32.79
−Loc	(755.18)		(987.09)		(819.9)	(765.02)	(748.09)	(844.92)
−Dat	32.74		225.9	66.61		−3.85	−57.97	−110.95
+Loc-Adv	(784.7)		(1010.83)	(840.40)		(767.42)	(775.22)	(749.12)
−Dat	32.63	159.03	383.56			5.02	−154.23	−112.49
+Loc-Adn	(786.79)	(913.28)	(1167.37)			(766.35)	(670.3)	(738.7)
−Dat	−24.71		178.72			25.45	−27.23	39.73
−Loc	(740.89)		(963.89)			(791.07)	(796.62)	(888.76)

and (11b), in which the word position of the dative is the same, also yielded a significant advantage for the +Locative-Adverbial condition ($F1(1,43) = 4.55$, $p < .05$; $F2(1,23) = 5.19$, $p < .05$). This finding is in direct opposition to the prediction of the incomplete dependency hypothesis.

At the accusative region (Region 6 in Figure 1), a 2×3 ANOVA crossing the Dative factor and the Locative factor revealed main effects of both factors such that the +Dative conditions were read faster than the –Dative conditions ($F1(1,43) = 13.06$, $p < .005$; $F2(1,23) = 8.31$, $p < .01$) and the +Locative conditions were faster than the –Locative conditions ($F1(2,86) = 5.81$, $p < .005$; $F2(2,46) = 3.16$, $p < .06$), although no interaction was found between the two factors ($Fs < 2.6$, $ps > .09$). Again, this result runs in the direction opposite to what the incomplete dependency hypothesis predicts.

At the following two verbal regions, the effect of the Dative factor disappeared (at the first verb, $Fs < 3.6$, $ps > .06$; at the second verb, $Fs < 2.4$, $ps > .13$), while the effect of the presence of Locative (with +Locative faster) remained (at the first verb, $F1(2,86) = 8.20$, $p < .005$; $F2(2,46) = 5.85$, $p < .01$; at the second verb, $F1(2,86) = 14.16$, $p < .001$; $F2(2,46) = 16.73$, $p < .001$); no interactions were found at these regions (all $Fs < 1.2$, $ps > .1$). Similar to the results at the preceding two nominal regions, this result runs directly contrary to the prediction of the incomplete-dependency theory. It should be noted, however, that the +Locative-Adnominal conditions did not pattern like the –Locative conditions, being read as fast as or faster than the +Locative-Adverbial conditions (+Loc-Adv vs. +Loc-And vs. Loc at Region 6: −51.91 ms vs. −54.67 ms vs. 23.37 ms; at Region 7: −90.91 ms vs. −151.62 ms vs. −52.49 ms; at Region 8: −122.34 ms vs. −125.51 ms vs. 3.68 ms). Direct comparison between the +Locative-Adnominal vs. +Locative-Adverbial conditions revealed no significant RT difference at Regions 6 and 8 (all $Fs < 1$, all $ps > .8$) and a significant difference at Region 7 with the adnominal locative faster ($F1(1,43) = 6.20$, $p < .05$; $F2(1,23) = 6.76$, $p < .05$). This runs contrary to the incomplete dependency hypothesis, which predicts slower RTs for the +Locative-Adverbial condition, because of the larger number of open dependencies in this condition at this position.

In the analyses summarized in Figure 2, in which the number of initial dependencies was taken as an independent variable (5 initial NPs for (11a), 4 initial NPs for (11b), (11c) and (11d), and 3 initial NPs for (11e) and (11f)), main effects were found at the accusative-NP region (Region 6: $F1(2,86) = 10.81$, $p < .001$; $F2(2,46) = 7.33$, $p < .005$) and at the higher-V region (Region 8: $F1(2,86) = 4.80$, $p < .02$; $F2(2,46) = 6.61$, $p < .005$), but not at the embedded-V region (Region 7: $Fs < 1$, $ps > .5$). The effect at the accusative NP was again in the direction opposite to the prediction of the incomplete dependency hypothesis: having more initial NPs made processing faster. Furthermore, the results at the higher verb support the incremental prediction-narrowing integration

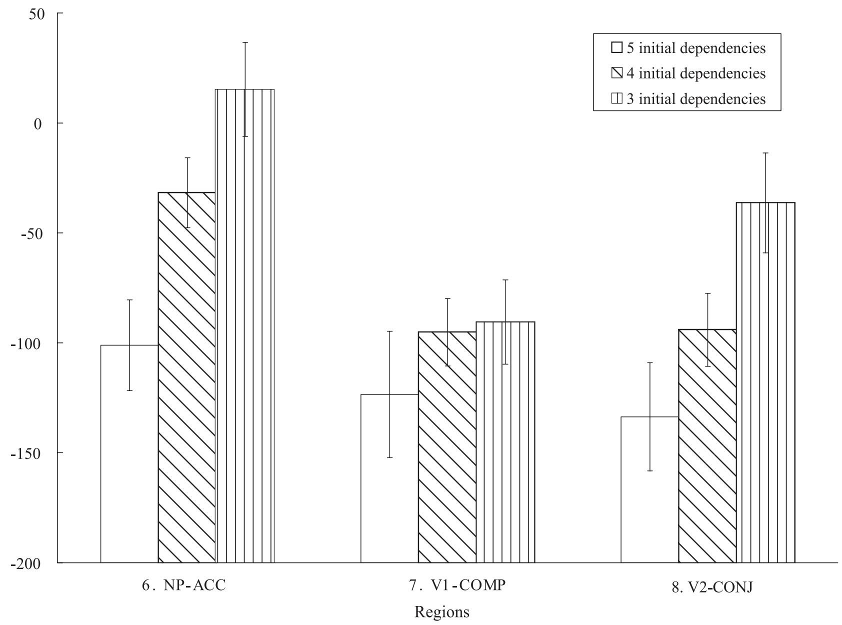


Figure 2. *Mean RTs for the last three regions according to the number of dependencies*

hypothesis over the bottom-up head-dependent distance integration hypothesis, similar to the online Japanese processing results from Nakatani and Gibson (2003). In fact, the results offer some tentative support for Konieczny's (1996) anticipation hypothesis, such that RTs should speed up at a verb, the more predictable it is depending on its preceding argument structure (see also Konieczny and Döring 2003). This hypothesis predicts differences at the most embedded verb. There are numerical trends in the predicted direction, but these trends are not significant. However, there are statistically significant trends at the outer verb in the direction predicted by Konieczny's hypothesis, which may be due to spillover from the embedded verb. Thus we see some possible support for Konieczny's hypothesis in this experiment.

3. General discussion

The main results of the current experiment were as follows. First, as in previous studies, participants slowed down significantly when they started a new embedded clause. This result is consistent with either a predicted head theory of syntactic expectation or an incomplete-dependency theory of syntactic expectation. More interestingly, participants did *not* slow down when there were more open dependencies within a clause. In particular, they did not slow down when there was an extra locative PP that was dependent on a verb to come. There was no reading time slowdown whether or not the word positions of the embedded material were matched across conditions. Furthermore, participants did not slow down when there was an additional dative argument of the verb. Again, there was no slowdown whether or not word positions were matched in this comparison. Although these latter results are both null effects, there was not even a suggestion of an effect in the direction predicted by the incomplete dependency hypothesis, in spite of the fact that the effect at the second nominative NP was highly reliable. Furthermore, the pool of subjects was substantial for a reading experiment: 45 participants. Thus the null result should probably be taken seriously. Overall, these results provide evidence against the incomplete dependency hypothesis. In contrast, these results are as expected by a predicted head theory of syntactic expectations. There was a measurable expectation cost when an additional verb and complementizer were expected, but there was no expectation effect for additional dependents of one expected verb.

Taken by themselves, the results of the current experiment are actually consistent with one version of Lewis's (1996) similarity-based incomplete dependency hypothesis in (7): a version in which there is a cost for keeping track of similar incomplete dependencies, but no cost for keeping track of different kinds of dependencies. This hypothesis explains the expectation cost effects for the second nominative, and the lack of expectation effects for the other incomplete dependencies, because they all have distinct case-marking and thematic roles. Whereas this hypothesis can successfully account for the current data, it is not successful at explaining the full range of expectation cost effects crosslinguistically. In particular, this expectation cost hypothesis does not account for two sets of results from Chen et al. (2005). In the first, it was demonstrated that keeping track of a wh-filler consumes processing cost. Specifically, in one experiment it was shown that participants read the region in italics in (12b) more slowly than the same region in (12a):

(12) a. *Sentential complement of a verb:*
The claim alleging that *the cop who the mobster attacked* ignored the informant might have affected the jury.

b. *Relative clause modifying a noun:*
The claim which *the cop who the mobster attacked* ignored might have affected the jury.

In both (12a) and (12b), a verb is expected for the NP *the claim*. In (12b), there is the added expectation of a position to be associate with the wh-filler *which*. There is no such expectation in (12a). A processing difference was found not only on the whole italicized region, but also on the initial NP *the cop*. This result demonstrates that there is a cost for keeping track of a single predicted head / incomplete dependency of a kind (in this case for a wh-filler dependency), without a second interfering one. Furthermore, a second experiment replicated this finding on simpler materials, without the additional embedded relative clause.

Some pilot work reported by Chen et al. (2005) suggests that there is an expectation cost effect for a single predicted prepositional phrase following a verb, as in the comparison between (13a) and (13b):

(13) a. *PP not predicted:*
Mary published *a book which had impressed some critics* who worked for a magazine.
b. *PP predicted:*
Mary gave *a book which had impressed some critics* who worked for a magazine to a young child.

In (13a), no PP is expected following the NP *the book*, because the verb *published* takes only an NP argument. In contrast, in (13b), the verb *gave* takes a PP argument following the NP *the book*. Chen et al. (2005) discuss pilot data that demonstrates correspondingly longer RTs for the italicized region in (13b) compared with (13a), suggesting that there is a cost associated with keeping track of the PP expectation.

In both of the cases exemplified in (12) and (13) there is a cost associated with an expectation, even when there is no similar expectation being held at the same time. Taken together with the Japanese evidence presented in the current article, these results suggest that an incomplete dependency hypothesis is not adequate to account for the crosslinguistic evidence. A predicted-head expectation account is preferred. Under a predicted head account, there is a cost associated with each predicted verb in the Japanese experiment reported here, but not for each incomplete dependency. Furthermore, there is a cost associated with a predicted empty NP position in the English wh-filler experiments, and there is a cost associated with a predicted PP position in the English argument structure experiment.

Of course, this does not mean that Lewis's hypothesis about there being interference costs associated with processing similar incomplete dependencies is incorrect. Indeed, this idea can be applied just as easily to the predicted-head expectation cost hypothesis. That is, there may be additional cost associated with interfering similar predicted heads, with the consequence that keeping track of two predicted heads of different syntactic categories may be easier than keeping track of two predicted heads of the same category. All current data that we are aware of are consistent with this hypothesis.

Finally, an interesting consequence of the current combination of results from English and Japanese is that they seem to be most parsimoniously accounted for under a theory that includes empty categories mediating wh-dependencies (Chomsky 1965, 1981; Fodor 1978). That is, if we accept that expectation costs are probably indexing predicted categories rather than incomplete dependencies (because of the evidence from head-final languages), then the only way to account for the results of Chen et al. (2005) wh-trace expectation results on materials like (12) is to assume the existence of a wh-trace, an empty category. In this experiment, that participants read the region in italics more slowly when there is a wh-filler-gap dependency pending as in (12b) than when no such dependency is pending as in (12a). If there were no wh-trace, such that the dependency were represented via a direct link between the wh-filler and the verb to come (e.g., Pickering and Barry 1991; cf. Gibson and Hickok 1993; Gibson and Warren 2004; for additional evidence of wh-filler processing in Japanese, see Miyamoto and Takahashi [2001] and Aoshima et al. [2003]), then there would be no additional expectation cost for this category prediction, because the verbal head would already be predicted by the existence of the embedded clause, and would not induce additional processing cost. Thus the results of this experiment, in conjunction with existing results from the processing of head-final languages, provide indirect evidence for the existence of empty categories in wh-filler dependencies.

Received 12 December 2003
Revised version received
11 April 2005

Konan University
Massachusetts Institute of Technology

Notes

* Correspondence address: Kentaro Nakatani, Konan University, 8-9-1 Okamoto, Higashi-Nada, Kobe 658-8501, Japan. E-mail: kentaron@konan-u.ac.jp.

1. Nominative case marking much more frequently marks subjects of verbs rather than objects of verbs. This lexical frequency effect will bias the processor in favor of the two-clause expectation over the one-clause possibility.
2. There may also be cost associated with predicting the two complementizers that mediate the verbs. It is currently an open question which kinds of predicted heads are associated with expectation costs.

3. As discussed above, there are a small number of Japanese verbs that take nominative objects. Hence a sequence of two nominative NPs is consistent with the prediction of a single verb from this class. However, the presence of a nominative marker usually suggests the subject position of a verb to come.
4. An anonymous reviewer points out that there are lexical differences before the target regions, in the form of the locative NP, which is in the locative conditions, but not in the others. Furthermore, the reviewer points out that this could give rise to differences in contingent structural frequencies in the target regions of analysis, on the following nouns, which might then provide a potential alternative account of our data. In order to work out the specific predictions of this account, it is necessary to know the structural frequencies of the materials that we compare. Because locatives are modifiers, and are hence optional, they probably do not occur along with most occurrences of most verbs, in written and spoken corpora. For example, Schütze and Gibson (1999) found that arguments frequently occur with their subcategorizing head verbs, but modifiers are generally much less common in English corpora. If head-dependent co-occurrence frequencies have some crosslinguistic generality, it is therefore likely to be the case in Japanese that instances of verbs with locative modification are less common than instances of the verbs without such modification. The contingent structural frequency hypothesis therefore predicts that there would be more difficulty in reading the structures in which the locative NP are included. But this is the opposite of the result that was observed (see the results section). Thus, the contingent structural frequency hypothesis seems to be an unlikely explanation for the observed pattern of data.

References

Aoshima, Sachiko; Phillips, Colin; and Weinberg, Amy (2003). Processing of Japanese Wh-scrambling constructions. In *Japanese/Korean Linguistics 12*, William McClure (ed.), 179–191. Stanford, CA: CSLI Publications.

Altmann, Gerry and Steedman, Mark (1988). Interaction with context during human sentence processing. *Cognition* 30, 191–238.

Chen, Evan; Gibson, Edward; and Wolf, Florian (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52, 144–169.

Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

—(1981). *Lectures on Government and Binding.* Dordrecht: Foris.

—and Miller, George A. (1963). Introduction to the formal analysis of natural languages. In *Handbook of Mathematical Psychology, Volume 2*, R. Duncan Luce, Robert R. Bush, and E. Galanter (eds.), 269–321. Wiley: New York.

Cutler, Anne; Dahan, Delphine; and van Donselaar, Wilma A. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech* 40, 141–202.

Ferreira, Fernanda and Clifton, Charles, Jr. (1986). The independence of syntactic processing. *Journal of Memory and Language* 25, 348–368.

Fodor, Janet D. (1978). Parsing strategies and constraints on transformations. *Linguistic Inquiry* 9, 427–473.

Frazier, Lynn (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*. Bloomington, IN: Indiana University Linguistics Club.

—(1987). Sentence processing: a tutorial review. In *Attention and Performance XII*, Max Coltheart (ed.), 559–585. Hillsdale, NJ: Erlbaum.

(AutoPDF V7 24/10/07 13:15) WDG (148×225mm) TimesM J-1850 Linguistics, 46:1 PMU:I(CKN[A])24/10/2007 pp. 63–86 1850_46-1_03 (p. 84)

Garnsey, Susan M.; Pearlmutter, Neal J.; Myers, Elizabeth; and Lotocky, Melanie A. (1997). The contributions of verb bias and plausibility to the comprehension of temporarily ambiguous sentences. *Journal of Memory and Language* 37, 58–93.

Gibson, Edward (1991). A computational theory of human linguistic processing: memory limitations and processing breakdown. Unpublished doctoral dissertation, Carnegie Mellon University.

—(1998). Linguistic complexity: locality of syntactic dependencies. *Cognition* 68, 1–76.

—(2000). The dependency locality theory: a distance-based theory of linguistic complexity. In *Image, Language, Brain*, Alec Marantz, Yasushi Miyashita, and Wayne O'Neil (eds.), 95–126. Cambridge, MA: MIT Press.

—and Hickok, Gregory (1993). Sentence processing with empty categories. *Language and Cognitive Processes* 8, 147–161.

—and Pearlmutter, Neal (1998). Constraints on sentence comprehension. *Trends in Cognitive Science* 2, 262–268.

—and Warren, Tessa (2004). Reading time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax* 7, 55–78.

—; Desmet, Timothy; Grodner, Daniel; Watson, Duane; and Ko, Kara (2005). Reading relative clauses in English. *Cognitive Linguistics* 16(2), 313–353.

Gordon, Peter C.; Hendrick, Randall; and Johnson, Marcus (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory and Cognition* 27, 1411–1423.

Hakuta, Kenji (1981). Grammatical description versus configurational arrangement in language acquisition: The case of relative clauses in Japanese. *Cognition* 9, 197–236.

Hawkins, J. A. (1994). *A Performance Theory of Order and Constituency.* Cambridge: Cambridge University Press.

Just, Marcel A.; Carpenter, Patricia A.; and Woolley, J. D. (1982). Paradigms and processing in reading comprehension. *Journal of Experimental Psychology: General* 111, 228–238.

Kimball, John (1973). Seven principles of surface structure parsing in natural language. *Cognition* 2, 15–47.

Konieczny, Lars (1996). Human sentence processing: A semantics-oriented parsing approach. Unpublished doctoral dissertation, Albert-Ludwigs-Universität, Freiburg.

—(2000). Locality and parsing complexity. *Journal of Psycholinguistic Research* 29, 627–645.

—and Döring, Philipp (2003). Anticipation of clause-final heads: Evidence from eye-tracking and SRNs. In *Proceedings of the 4th International Conference on Cognitive Science*, 330–335. Sydney: University of New South Wales.

Lewis, Richard (1993). An architecturally-based theory of human sentence comprehension. Unpublished doctoral dissertation, Carnegie Mellon University.

—(1996). A theory of grammatical but unacceptable embeddings. *Journal of Psycholinguistic Research* 25, 93–116.

—and Nakayama, Mineharu (2002). Syntactic and positional similarity effects in the processing of Japanese embeddings. In *Sentence Processing in East Asian Languages*, Mineharu Nakayama (ed.), 85–110, Stanford, CA: CSLI Publications.

MacDonald, Maryellen C.; Pearlmutter, Neal; and Seidenberg, Mark S. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review* 101, 676–703.

Miyamoto, Edson (2002). Case markers as clause boundary inducers in Japanese. *Journal of Psycholinguistic Research* 31, 307–347.

—and Takahashi, Shoichi (2001). A filled-gap effect without gaps in Japanese. Paper presented at the 14th Annual CUNY Conference on Human Sentence Processing, CUNY.

Nakatani, Kentaro; Babyonyshev, Maria; and Gibson, Edward (2000). The complexity of nested structures in Japanese. Poster presented at the 13th Annual CUNY Conference on Human Sentence Processing, University of California, San Diego.

—and Gibson, Edward (2003). An on-line study of Japanese nesting complexity. Poster presented at the 16th Annual CUNY Conference on Human Sentence Processing, Massachusetts Institute of Technology.

Pickering, Martin; and Barry, Guy (1991). Sentence processing without empty categories. *Language and Cognitive Processes* 6, 229–259.

Schütze, Carson T. and Gibson, Edward (1999). Argumenthood and English prepositional phrase attachment. *Journal of Memory and Language* 40, 409–431.

Stabler, Edward P. (1994). The finite connectivity of linguistic structures. In *Perspectives on Sentence Processing*, Charles Clifton Jr., Lynn Frazier, and Keith Rayner (eds.), 303–336. Hillsdale, NJ: Erlbaum.

Tanenhaus, Michael K.; Spivey-Knowlton, Michael J.; Eberhard, Kathleen M.; and Sedivy, Judie C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science* 268, 1632–1634.

—and Trueswell, John C. (1995). Sentence comprehension. In *Speech, Language and Communication*, Joanne L. Miller and Peter D. Eimas (eds.), 217–262. San Diego, CA: Academic Press.

Trueswell, John C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language* 35, 566–585.

—; Tanenhaus, Michael K.; and Garnsey, Susan M. (1994). Semantic influences on parsing: use of thematic role information in syntactic disambiguation. *Journal of Memory and Language* 33, 285–318.

—; Tanenhaus, Michael K.; and Kello, Christopher (1993). Verb-specific constraints in sentence processing: separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory and Cognition* 19, 528–553.

Uehara, Keiko and Bradley, Dianne C. (2002). Center-embedding problem and the contribution of nominative case repetition. In *Sentence Processing in East Asian Languages*, Mineharu Nakayama (ed.), 257–287. Stanford, CA: CSLI Publications.

Vasishth, Shravan (2003). *Working Memory in Sentence Comprehension: Processing Hindi Center Embeddings*. New York: Routledge.

Wanner, Eric and Maratsos, Michael (1978). An ATN approach in comprehension. In *Linguistic Theory and Psychological Reality*, Morris Halle, Joan Bresnan, and George Miller (eds.), 119–161. Cambridge, MA: MIT Press.

Warren, Tessa and Gibson, Edward (2002). The influence of referential processing on sentence complexity. *Cognition* 85, 79–112.

Watson, Duane and Gibson, Edward (2004). The relationship between intonational phrasing and syntactic structure in language production. *Language and Cognitive Processes* 19, 713–755.

Yngve, Victor H. (1960). A model and a hypothesis for language structure. *Proceedings of the American Philosophical Society* 104, 444–466.

(AutoPDF V7 24/10/07 13:15) WDG (148×225mm) TimesM J-1850 Linguistics, 46:1 PMU:I(CKN[A])24/10/2007 pp. 63–86 1850_46-1_03 (p. 86)