

# Incremental Processing of Coreference and Binding in Japanese

*Sachiko Aoshima, Masaya Yoshida, and Colin Phillips*

*Abstract.* This article presents two on-line self-paced reading studies and three off-line acceptability judgment studies on the processing of backward anaphoric dependencies in Japanese in which a pronoun precedes potential antecedent noun phrases. The studies investigate the real-time formation of coreference relations and operator-variable binding relations to determine whether speakers of head-final languages are able to construct grammatically accurate syntactic structures before they encounter a verb. The results of the acceptability rating studies confirm previous claims that backwards anaphoric dependencies in Japanese are more acceptable in configurations where a pronoun has been fronted via scrambling from a position where it would be c-commanded by its antecedent. The results of the on-line studies demonstrate that these acceptability contrasts have an immediate impact on parsing. Reading-time results showed immediate sensitivity to the semantic congruency between an NP and a preceding pronoun in noncanonical (“scrambled”) word orders, and no immediate effect of semantic congruency otherwise. This contrast was found both for coreference relations involving the personal pronouns *kare/kanajo* (experiment 1) and for operator-variable relations involving the demonstrative pronoun *soko* (experiment 3). These findings go beyond previous evidence for incremental parsing in head-final languages by showing that Japanese speakers build compositional structures (such as anaphoric relations) in a grammatically constrained fashion in advance of encountering a verb in the input.

## 1. Introduction

Word order differences between head-initial languages like English and head-final languages like Japanese mask fundamental commonalities in the structural representations used in both types of language. A great deal of work has documented that Japanese is subject to many of the same syntactic constraints as English, despite its head-finality and relatively liberal word order (e.g., Harada 1972; Harada 1977; Hoji 1985; Miyagawa 1989; Saito 1983, 1985, 1989, 1992; Saito & Hoji 1983; Takezawa 1987; Ueda 1990). From the perspective of real-time language processing, on the other hand, it has been less clear whether head-final languages are parsed and interpreted in the same way as head-initial languages. In particular, it has been difficult to establish whether the parsing of head-final languages proceeds with the same

We are grateful to two anonymous reviewers and to Janet Fodor, Tomohiro Fujii, Hajime Hoji, Norbert Hornstein, Ayumi Ueyama, and Amy Weinberg for valuable feedback at various stages of this research, and to John Mathews, Hiromu Sakai, Hajime Ono, Megumi Yoshimura, and Chao Zhang for assistance in conducting the experiments described here. This work was supported in part by grants to Colin Phillips from NSF (#BCS-0196004) and the Human Frontiers Science Program (#RGY-0134), and to Masaya Yoshida from an ESRC postdoctoral fellowship (PTA-026-27-1379).

degree of incrementality and accuracy that has been demonstrated for head-initial languages. In this article we present a series of experimental studies on the processing of coreference relations in Japanese that together indicate a high degree of incremental, grammatically accurate structure building before a verb is reached.

There is much evidence for highly incremental structure building and interpretation in head-initial languages like English. Evidence for rapid interpretive mechanisms that integrate each incoming word into an ongoing structure within a few hundred milliseconds comes from speech shadowing (Marslen-Wilson 1973, 1975), eye-tracking measures (e.g., Tanenhaus et al. 1995, Sussman & Sedivy 2003), event-related brain potentials (e.g., Kutas & Hillyard 1980; Garnsey, Tanenhaus & Chapman 1989; Neville et al. 1991; Friederici, Pfeifer & Hahne 1993), and speed-accuracy tradeoff measures (McElree & Griffith 1995), among many others. Much evidence also indicates that real-time structure building is sensitive to grammatical constraints, including constraints on argument structure (e.g., Boland et al. 1995, Pickering & Traxler 2001), long-distance dependencies (e.g., Stowe 1986, Traxler & Pickering 1996, Phillips 2006), and coreference constraints (e.g., Nicol & Swinney 1989, Sturt 2003, Kazanina et al. 2007).

Findings of grammatical accuracy in real-time processing imply that the representations that are constructed in real-time include sufficient detail for implementation of grammatical constraints. If incremental interpretation is achieved by means of more superficial heuristics and strategies (e.g., Townsend & Bever 2001, Ferreira 2003), then it is less expected that structure building should reflect detailed grammatical constraints. This argument forms the basis of the current study, where we use tests of real-time application of constraints and biases in pronoun interpretation as a probe for incremental structure building in Japanese. Evidence for real-time application of grammatical constraints is also important for theories that attempt to minimize the difference between models of grammatical structure and models of real-time processes (e.g., Bresnan 1978; Phillips 1996, 2003; Steedman 2000; Kempson, Meyer-Viol & Gabbay 2001; Kempen & Harbusch 2002). An important prediction of such theories is that real-time language processes should be grammatically accurate. To the extent that this prediction is borne out, it poses a challenge for well-known grammatical models that explicitly eschew claims about real-time computation (e.g., Chomsky 1965).

Whereas the evidence for incremental, accurate structure building in head-initial languages is extensive, and numerous explicit models are available to capture how this is achieved (e.g., Johnson-Laird 1983, Steedman 1989, Resnik 1992, Stabler 1994, Sturt 1997, Sturt & Crocker 1996, Vosse & Kempen 2000), less is known about real-time interpretation in strongly head-final languages like Japanese. Many studies on English-type languages have emphasized how incremental interpretation is guided by detailed information about the combinatorial possibilities of verbs (e.g., Ford, Bresnan & Kaplan

1982; Trueswell, Tanenhaus & Kello 1993; Boland et al. 1995), but this is of limited value in Japanese, where comprehenders may encounter a long sequence of argument and adjunct phrases before encountering a verb, as in (1).

- (1) John-ga denwa-de Mary-ni Tom-ga asa rokuji-ni  
John-NOM phone-by Mary-DAT Tom-NOM morning six-at  
inu-ni esa-o ageta ka kiita.  
dog-DAT food-ACC gave if asked  
'John asked Mary by phone if Tom gave his dog food at six  
in the morning.'

A central concern in research on the processing of head-final languages has been to determine whether sentences like (1) are incrementally parsed using similar mechanisms to those used in languages like English. One view is that verbs and other heads of phrases play a similarly pivotal role in head-initial and head-final languages, and that constituents are combined to form larger units only when the head of a phrase is encountered. For Japanese the consequence of this *head-driven* approach is that structure building is predicted to be delayed until a clause-final verb is reached (Pritchett 1988, 1991, 1992; Mulders 2002). Although this approach has had relatively few advocates, it is compatible with much existing evidence, and it has the advantage that it can be straightforwardly implemented using familiar algorithms used to parse head-initial languages. In contrast, incremental attachment models posit that Japanese sentence structures are gradually expanded as each new word is encountered, with no delay until a verbal head is reached (e.g., Fodor & Inoue 1994, 1998; Inoue & Fodor 1995; Mazuka & Itoh 1995; Kamide & Mitchell 1999; Mazuka, Itoh & Kondo 2002; Miyamoto & Takahashi 2002; Fodor & Hirose 2003; Kamide, Altmann & Haywood 2003; Nakatani & Gibson 2003). These models align well with the intuition of incrementality reported by most speakers of Japanese, but there are few explicit models of incremental, accurate structure building in strongly head-final structures (Crocker 1996, Sturt & Crocker 1996, Sturt 1997, Schneider 1999, Aoshima 2003). More important for our current purposes, existing evidence for incremental parsing in head-final languages shows that speakers are sensitive to the sequencing of preverbal phrases, but provides less information on how preverbal phrases are combined in accurate compositional interpretive processes. There is little doubt that case-marking information provides useful cues that help to guide interpretation in these languages, but it is less clear specifically what representations are constructed on the basis of case information (Inoue 1991, Miyamoto 2002).

A number of different types of evidence have been presented in support of incremental accounts of the processing of head-final languages such as Japanese. Mazuka and Itoh (1995) present one such argument based on sentences like (2), in which a sentence-initial sequence of NPs could initially

be understood as belonging to a single ditransitive clause but must ultimately be parsed as belonging to a main clause and a relative clause.

- (2) a. John-ga Mary-ni atta syoonen-o hometa.  
 John-NOM Mary-DAT met boy-ACC praised  
 ‘John praised the boy who met Mary.’  
 b. John-ga Mary-ni mita syoonen-o syookaisita.  
 John-NOM Mary-DAT saw boy-ACC introduced  
 ‘John introduced to Mary the boy who he met.’

Mazuka and Itoh observe that Japanese speakers experience surprise upon reaching the transitive verb *mita* ‘saw’, and reason that this is due to an unfulfilled expectation for a ditransitive verb. The formation of this expectation is attributed to preverbal structure building. However, the surprise effect is open to different interpretations. Because the effect occurs at the verb position, it is also compatible with a head-driven account in which the parser delays structure building until a verb is encountered, at which point it attempts to satisfy the thematic requirements of all NPs at once. Under this account the transitive verb causes surprise because it does not assign a thematic role to the dative NP. Alternatively, the surprise effect may indeed reflect expectations formed in advance of the verb, but this could be achieved by a mechanism that analyzes sequences of preverbal arguments as a list, rather than as a compositional hierarchical structure. Thus, a limitation of this type of evidence for incrementality is that it relies on effects at the verb position to draw inferences about the processing of preverbal material.

Bader and Lasser (1994) present a related argument, based on patterns of ease and difficulty at sentence-final auxiliaries in German (3). Examples (3a,b) are identical and ambiguous until the sentence-final auxiliaries, due to the fact that the feminine pronoun *sie* is ambiguous between nominative and accusative. Bader and Lasser reason that if the parser delays structural decisions until a verb is reached, then it should analyze the pronoun *sie* as an accusative direct object of the verb *fragen* ‘ask’, as this allows the pronoun to immediately receive a thematic role. This predicts a preference for the completion of the sentence in (3b).

- (3) a. ...daß sie<sub>NOM</sub> [CP nach dem Ergebnis zu fragen]  
 That she for the results to ask  
 tatsächlich erlaubt hat.  
 indeed permitted has  
 ‘...that she indeed has given permission to ask for the result.’  
 b. ...daß [CP sie<sub>ACC</sub> nach dem Ergebnis zu fragen]  
 That her for the results to ask  
 tatsächlich erlaubt worden ist.  
 indeed permitted been is  
 ‘...that permission indeed has been given to ask her for the result.’

In fact, German speakers found it easier to process (3a), in which *sie* is the subject of the second verb *erlaubt* ‘permitted’. Bader and Lasser argue that this preference is a consequence of an incremental parser that shows a bias to treat *sie* as a nominative subject pronoun as soon as it is encountered, which makes it unavailable as the object of the first verb. However, a limitation of this evidence for incrementality is that it focuses on resolution of the case ambiguity of the pronoun *sie* and does not directly implicate compositional structure building (i.e., formation of syntactic or semantic relations among preverbal NPs).

A third type of argument for incremental structuring comes from studies on the processing of noncanonical word orders derived by “scrambling.” The logic is that if noncanonical word orders are derived by transformational operations then those operations may be associated with increased processing difficulty. If increased processing difficulty is observed prior to a verb, this may reflect the preverbal application of the transformational operation. For example, Miyamoto and Takahashi (2002) showed in a self-paced reading study on Japanese that the noncanonical NOM-ACC-DAT word order in (4b) is read more slowly than the canonical NOM-DAT-ACC word order in (4a).

- (4) a. Ofisu-de syokuin-ga kakarityoo-ni otya-o dasita  
office-at employee-NOM manager-DAT tea-ACC served  
josei-o teineini hometa-to Aiharasan-ga hanasiteita.  
woman-ACC politely praised-COMP Aihara-NOM said
- b. Ofisu-de syokuin-ga otya-o kakarityoo-ni dasita  
office-at employee-NOM tea-ACC manager-DAT served  
josei-o teineini hometa-to Aiharasan-ga hanasiteita.  
woman-ACC politely praised-COMP Aihara-NOM said  
‘At the office, Aihara said that the employee politely praised the woman who had served tea to the manager.’

Miyamoto and Takahashi’s finding demonstrates that scrambling is associated with additional processing difficulty before the verb is reached (see also Mazuka, Itoh & Kondo 2002), but this does not necessarily implicate incremental construction of detailed syntactic structures, as Miyamoto and Takahashi point out. The noncanonical order may be more difficult simply because it is less frequent than the canonical order. Similarly, reading times may have been faster on the preverbal NP in (4a) than in (4b) because a NOM-DAT sequence is frequently followed by an accusative-marked NP, whereas a NOM-ACC sequence does not strongly predict a following dative-marked NP.

An interestingly different argument for incremental interpretation in Japanese comes from studies in the Visual World paradigm by Kamide and her colleagues (Kamide, Altmann & Haywood 2003). Kamide and colleagues showed that Japanese speakers can use information from a pair of sentence-initial NPs to anticipate an upcoming referent, as indexed by increased looks

to a target object in a visual scene. For example, in a scene containing a waitress, a customer, and a hamburger, Japanese speakers show an increased proportion of looks to the hamburger following a sequence *waitressu-ga kyaku-ni* 'waitress-NOM customer-DAT...' than following a corresponding sequence in which the second NP bears accusative case. This shows that, at least in highly suggestive visual contexts, Japanese speakers can use information from early NPs to anticipate subsequent NPs in the input, but as Kamide and colleagues acknowledge, this leaves open the question of the quality of the preverbal representations. More is needed to show that anticipatory eye movements are based on the combined interpretation of the preverbal NPs, and these results do not show preverbal application of constraints on structural combinations of NPs.

In this article we investigate the nature of preverbal structure building in Japanese using studies of anaphoric relations between preverbal noun phrases. Following a logic suggested by Bader (1994) and Schneider (1999) we reason that if Japanese speakers are able to form anaphoric relations between noun phrases before reaching a verb then this implies some form of compositional structure building—that is, the parser builds a structure that combines the preverbal noun phrases and interprets the semantic relations between them. Furthermore, if on-line formation of anaphoric relations applies only in those configurations where anaphoric relations are judged as acceptable in off-line measures, then this implies that the same detailed representations that underlie the acceptability ratings also underlie real-time structure building.

In what follows we distinguish two types of anaphoric relations: *coreference* relations that assign to a pronoun the same referent as its antecedent (5a), and *binding* relations, in which a pronoun is interpreted as a variable bound by its quantificational antecedent (5b) (Reinhart 1983). It is well known that a pronoun may not c-command its antecedent, whether referential or quantificational (6) (Chomsky 1981). Bound-variable interpretations of pronouns are typically available only when the quantificational antecedent c-commands the pronoun (7b), but c-command is not normally a requirement for coreference relations (7a). Nevertheless, cases of *backwards anaphora* (or *cataphora*) in which a pronoun precedes an antecedent within the same clause are generally judged to be marked or degraded (8a) (cf. Kuno 1986). Corresponding cases involving quantificational antecedents are, unsurprisingly, unacceptable (8b).

- (5) a. The student<sub>i</sub> lent some money to his<sub>i</sub> friend.  
 b. Every student<sub>i</sub> lent some money to his<sub>i</sub> friend.
- (6) \*He<sub>i</sub> thinks that John<sub>i</sub> is smart.
- (7) a. A story about John<sub>i</sub> impressed his<sub>i</sub> mother.  
 b. \*A story about every student<sub>i</sub> impressed his<sub>i</sub> mother.

- (8) a. ?His<sub>i</sub> friend lent some money to the student<sub>i</sub>.  
 b. \*His<sub>i</sub> friend lent some money to every student<sub>i</sub>.

Similar contrasts obtain in Japanese, although they are sometimes obscured by the slightly different pronominal repertoire of Japanese. The personal pronouns *kare* 'he' and *kanojo* 'her' participate in coreference relations but typically resist bound-variable readings (Kuroda 1965; 1985, 1991; Kitagawa 1981; Nakayama 1982; Noguchi 1997; Saito & Hoji 1983). *Kare* may corefer with a c-commanding antecedent, such as the nominative subject NP in (9a), but coreference is less acceptable when *kare* precedes its antecedent, as in (9b), where no c-command obtains since *kare* is embedded inside the subject NP. Meanwhile, the Japanese demonstrative pronoun *soko*, which refers to an organization or a place, can be used as a bound-variable pronoun (Hoji 1991, 1995, 2002; Hoji et al. 2000; Kuroda 1979; Takubo & Kinsui 1998; Ueyama 1998, 2000). As in English, the bound-variable reading is possible when the quantificational antecedent c-commands *soko* (10a), but it is disallowed when c-command does not obtain and *soko* precedes the antecedent (10b).

- (9) a. Daigakusei-ga kare<sub>i</sub>-no tomodachi-ni okane-o kasita.  
 college student-NOM he-GEN friend-DAT money-ACC lent  
 'The college student lent his friend some money.'  
 b. ??Kare<sub>i</sub>-no tomodachi-ga daigakusei-ni okane-o kasita.  
 he-GEN friend college student-DAT money-ACC lent  
 'His friend lent the college student some money.'
- (10) a. Dono hokengaisya<sub>i</sub>-mo soko<sub>i</sub>-no kogaisya-ni  
 every insurance company-MO soko-GEN subsidiary-ACC  
 nenji hookokusyo-o okutta.  
 annual report-ACC sent  
 'Every insurance company sent an annual report to its subsidiary.'  
 b. \*Soko<sub>i</sub>-no kogaisya-ga dono-hokengaisya<sub>i</sub>-ni-mo nenji  
 soko-GEN subsidiary-NOM every-insurance company-MO annual  
 hookokusyo-o okutta.  
 report-ACC sent  
 'Its subsidiary sent an annual report to every insurance company.'

However, there is an important exception to the restriction on pronouns that precede their clausemate antecedents. When a pronoun precedes its antecedent as a consequence of scrambling, coreference and binding are more acceptable. For example, the pronoun *kare* is more readily associated with a following antecedent in the noncanonical word order in (11) than in the canonical word order in (9b). A similar contrast obtains for the bound-variable interpretation of *soko* when comparing the scrambled word order in (12) with the canonical word order in (10b). These contrasts in acceptability are confirmed in rating studies described below. A standard account of this increase in acceptability is

that scrambled phrases may be interpreted in their canonical position, with the consequence that the pronouns in (11) and (12) are then both preceded and c-commanded by their antecedents, just as in (9a) and (10a) (Miyara 1982; Saito 1985). However, nothing in the studies that follow depends crucially on this account of the contrast. The contrast may also be captured in accounts of constraints on anaphora that focus on argument prominence hierarchies (e.g., Bach & Partee 1980, Pollard & Sag 1994, Steedman 1997), and may be related to the widespread finding in studies of forward anaphora that pronouns are preferentially associated with antecedents that are more thematically prominent (e.g., Chafe 1976; Crawley & Stevenson 1990; Kaiser & Trueswell, in press).

- (11) Kare<sub>i</sub>-no tomodachi-ni daigakusei<sub>i</sub>-ga okane-o kasita.  
 he-GEN friend-DAT college student-NOM money-ACC lent  
 ‘The college student lent some money to his friend.’
- (12) Soko<sub>i</sub>-no kogaisya-ni dono hokengaisya<sub>i</sub>-mo nenji  
 soko-GEN subsidiary-DAT every insurance company-MO annual  
 hookokusyo-o okutta.  
 report-ACC sent  
 ‘Every insurance company sent an annual report to its subsidiary.’

The contrast in acceptability of backwards anaphora between the canonical NOM-DAT word order in (9b) and (10b) and the noncanonical DAT-NOM word order in (11) and (12) forms the basis of the experiments that follow. The experiments test whether Japanese speakers incrementally construct coreference relations in the configurations where such relations are possible (see (11) and (12)), and not in the configurations where coreference is degraded or disallowed (see (9b) and (10b)). If this contrast obtains at regions prior to the verb, it indicates that Japanese speakers form compositional relations among NPs before reaching a verb, and suggests that the structures that are built incrementally have sufficient detail to implement constraints and biases in anaphora and their interaction with scrambling.

We investigate the real-time construction of anaphoric relations using a semantic mismatch paradigm that has been used in a number of studies of English-type languages. For example, van Gompel and Liversedge (2003) compared eye-tracking times for sentences like (13a,b), in which the subject of the main clause either matches or mismatches in gender with a pronoun in a preceding clause. They interpret the slowdown in reading times at mismatching subject NPs as evidence of an attempt to create a coreference link between the pronoun and the main clause subject before it becomes clear that the two NPs are semantically incompatible. We discuss below the possibility that this may reflect construction of a coreference relation even before the subject appears in the input. This *gender mismatch effect* is an analog of “active”

dependency formation effects that have been widely reported in studies of the processing of *wh*-dependencies (e.g., Crain & Fodor 1985, Stowe 1986, Traxler & Pickering 1996; for a review, see Phillips & Wagers 2007).

- (13) a. When {she/he} was fed up, the girl visited the boy very often.  
 b. When {he/she} was fed up, the boy visited the girl very often.

If the Japanese parser actively constructs anaphoric relations among preverbal NPs, then we should expect to find semantic mismatch effects for NPs in potential antecedent positions. Further, if the Japanese parser restricts the on-line construction of anaphoric relations to configurations where binding or coreference are acceptable, then we should expect semantic mismatch effects to be absent for NPs in illicit or dispreferred antecedent positions. This parallels the use of the gender mismatch paradigm in other studies as a tool for investigating the on-line processing of binding constraints (e.g., Clifton, Kennison & Albrecht 1997, Badecker & Straub 2002, Kennison 2003, Sturt 2003, Kazanina et al. 2007; for a review, see Kazanina 2005).

In what follows we describe four off-line and two on-line studies of anaphoric relations in Japanese that follow the logic outlined here. Section 2 presents one off-line study and one on-line study on coreference relations involving the personal pronouns *kare* and *kanojo*. Section 3 presents further evidence for the availability of bound-variable interpretations for the demonstrative *soko*, including two off-line studies, and section 4 presents one on-line and one off-line study of the processing of operator-variable relations involving *soko*. Section 5 discusses the implications of the experimental results, and section 6 concludes.

## 2. Experiment 1: Processing Coreference Relations with *Kare/Kanojo*

Because of the possibility of scrambling, a pronoun inside a sentence-initial dative NP may take a following nominative NP as its antecedent (14a). In contrast, in a sentence that is identical except that the case-markers on the noun phrases are exchanged, yielding a canonical NOM-DAT word order, coreference is less acceptable (14b).<sup>1</sup> Experiment 1 verifies this acceptability

<sup>1</sup> There is good evidence that the genitive pronoun in (14a,b) is embedded inside the sentence-initial argument and is not raised out of that NP. First, Japanese has possessor raising, but it is marked by a case change from genitive to nominative. Second, genitive possessors and their hosts cannot be separated by clefting or scrambling. Third, genitive possessors cannot be antecedents of main clause reflexives, whereas raised nominative possessors are at least marginally acceptable for the speakers that we consulted (i). The reflexive in (i) can also take the subject NP *tomodachi* 'friend' as an antecedent.

- (i) {\**kare*<sub>1</sub>-no/ ?*kare*<sub>1</sub>-ga} dono-tomodachi-ga zibun<sub>1</sub>-o semeta-no?  
 he-GEN he-NOM which-friend-NOM self-ACC blamed-Q  
 'Which of his<sub>1</sub> friends blamed himself<sub>1</sub>?'

contrast and employs it in a gender mismatch paradigm to test for incremental construction of coreference relations in Japanese. If the Japanese parser builds structure incrementally and actively seeks antecedents for pronouns only in configurations where coreference is acceptable, then we should expect to observe gender mismatch effects only in potential antecedent positions; that is, the position of the nominative NP in (14a), but not the position of the dative NP in (14b).

- (14) a. Kare<sub>i</sub>-no tomodati-ni sono otoko<sub>i</sub>-ga hotyoo-de osoikakata.  
 he-GEN friend-DAT the man-NOM knife-with attacked  
 ‘The man attacked his friend with a knife.’
- b. ?\*Kare<sub>i</sub>-no tomodati-ga sono otoko<sub>i</sub>-ni hotyoo-de osoikakata.  
 he-GEN friend-NOM the man-DAT knife-with attacked  
 ‘His friend attacked the man with a knife.’

### 2.1 *Participants*

The experiment consisted of two tasks. An on-line self-paced reading task (experiment 1B) was followed by an off-line acceptability judgment task (experiment 1A). Forty-four native speakers of Japanese participated in both tasks. All of them were students at the University of Maryland, USA, or at Shizuoka University or Shizuoka Sangyo University, Japan. They were paid \$10.00 or its yen equivalent for their participation in the experiment, which lasted about 50 minutes.

The acceptability judgment task was administered after the self-paced reading task, so that on-line reading times would not be affected by the presence of related sentences in the judgment task. However, we first describe the judgment task, since it establishes a key premise for the on-line task.

### 2.2 *Experiment 1A: Acceptability Judgment Task*

The aim of the acceptability judgment task was to test whether Japanese native speakers show the acceptability contrast in (14), where backwards anaphora is more acceptable in noncanonical (scrambled) word order than in canonical word order. The judgment task followed a  $2 \times 2$  factorial design that manipulated the word order of the noun phrases (scrambled vs. canonical) and the type of the sentence-initial noun phrase (non-*wh*-phrase vs. *wh*-phrase). The conditions with sentence-initial *wh*-phrases were included because of their similarity to the materials used in the on-line experiment, which used *wh*-phrases because they have been shown to trigger early “undoing” of scrambled word orders (e.g., Aoshima, Phillips & Weinberg 2004). The conditions with sentence-initial non-*wh*-phrases were included to test whether the acceptability ratings generalize beyond *wh*-phrases. Each target sentence consisted of a two-word NP containing a genitive-marked pronoun, followed

by an adverbial phrase, an NP that was rated as a potential antecedent for the pronoun, and a sentence-final verb. Thirty-two sets of the four conditions were distributed among four lists in a Latin Square design. Each participant saw one of the lists intermixed with 32 unrelated items in a pseudorandom order. Among 32 target items, 16 were derived from items used in the self-paced reading task, and the other 16 were different from those used in the on-line task. A sample set of experimental conditions is shown in (15). The target and filler items were counterbalanced with respect to the gender of the pronoun. Participants were asked to judge whether the underlined pronoun and the underlined noun phrase in each sentence could be understood as coreferential, indicating their rating on a scale from 1 to 5, where 1 means that coreference is impossible, and 5 means that it is fully acceptable. Whereas all target items contained backward anaphoric relations, the filler items also included instances of forward anaphoric relations between pronouns and potential antecedent NPs, covering both unambiguous and ambiguous cases, as in English examples like *John* was watching TV in his room and John was watching TV with Bill in his room.

- (15) a. Non-*wh*-phrase, canonical  
Kare-no gakusei-ga syokudoo-de sensei-ni atta.  
 he-GEN student-NOM cafeteria-at teacher-DAT met  
 'His student met the teacher at the cafeteria.'
- b. Non-*wh*-phrase, scrambled  
Kare-no gakusei-ni syokudoo-de sensei-ga atta.  
 he-GEN student-DAT cafeteria-at teacher-NOM met  
 'His student, the teacher met at the cafeteria.'
- c. *Wh*-phrase, canonical  
Kare-no dono-gakusei-ga syokudoo-de sensei-ni atta-no?  
 he-GEN which- student-NOM cafeteria-at teacher-DAT met-Q  
 'Which of his students met the teacher at the cafeteria?'
- d. *Wh*-phrase, scrambled  
Kare-no dono-gakusei-ni syokudoo-de sensei-ga atta-no?  
 he-GEN which- student-DAT cafeteria-at teacher-NOM met-Q  
 'Which of his students did the teacher meet at the cafeteria?'

Mean coreference acceptability ratings in all conditions were entered into a repeated-measures ANOVA, with noun phrase type (non-*wh*-phrase, *wh*-phrase) and word order (scrambled, canonical) as within-subjects factors. Scores from four participants whose data were excluded from the analysis of the on-line task were also excluded here. Results are shown in Table 1.

There was a significant main effect of word order, due to higher coreference acceptability ratings for scrambled conditions than for canonical conditions,  $F_1(1,39) = 13.91$ ,  $MSE = 1.39$ ,  $p < .001$ ;  $F_2(1,23) = 5.92$ ,  $MSE = 2.8$ ,  $p < .05$ . There was no significant main effect of noun phrase type,  $F_1(1,39) = 1.5$ ,  $MSE = 1.67$ ,  $p = .23$ ;  $F_2(1,23) = 2.23$ ,  $MSE = 1.51$ ,  $p = .15$ , and no

**Table 1. Experiment 1A: mean acceptability ratings for coreference judgments ( $n = 44$ )**

Condition	Mean	SD
Non- <i>wh</i> -phrase, canonical	2.92	1.41
Non- <i>wh</i> -phrase, scrambled	3.23	1.35
<i>Wh</i> -phrase, canonical	3.05	1.34
<i>Wh</i> -phrase, scrambled	3.53	1.29

interaction of word order type and noun phrase type ( $F_s < 1$ ). Pairwise comparisons within each level of the noun phrase type factor showed that in the *wh*-phrase conditions the scrambled condition was rated significantly higher than the canonical condition,  $F_1(1,39) = 10.98$ ,  $MSE = 1.05$ ,  $p < .01$ ;  $F_2(1,23) = 4.56$ ,  $MSE = 2.5$ ,  $p < .05$ , and in the non-*wh*-phrase conditions the scrambled condition was also rated higher than the canonical condition, although this difference was only marginally significant,  $F_1(1,39) = 3.67$ ,  $MSE = 2.16$ ,  $p = .06$ ;  $F_2(1,23) = 2.89$ ,  $MSE = 20.2$ ,  $p = .09$ . Pairwise comparisons within each level of the word order factor yielded no significant differences ( $p_s > .15$ ). Additional analyses showed that there were no differences in ratings between target items that were derived from sentences used in the on-line study and items that were not included in the on-line study.<sup>2</sup>

### 2.3 Experiment 1B: Materials

The experimental items for the self-paced reading study consisted of four conditions organized in a  $2 \times 2$  factorial design that manipulated the gender match between a pronoun and a subsequent NP (match vs. mismatch) and the relative order of nominative and dative NPs (canonical NOM-DAT order vs. scrambled DAT-NOM order). Apart from these differences, all conditions followed a similar scheme. A sample set of items is shown in (16). A genitive-marked pronoun (*kare-no* 'his' or *kanojo-no* 'her') always appeared in the second region of the sentence, embedded inside a *wh*-phrase. The critical potential antecedent NP appeared two regions after the *wh*-phrase and always denoted a gender specific referent that matched the gender of the pronoun in two conditions and mismatched in two conditions (e.g., *oba* 'aunt' and *oji* 'uncle' in (16)). The primary interest of the study was in the presence or absence of gender mismatch effects when readers encounter this NP. The average length

<sup>2</sup> We report coreference acceptability scores here in terms of shifts in mean scores, for ease of exposition and to allow comparison with other studies. However, individual ratings showed a bimodal distribution in all conditions, with few scores of 3 on the 1–5 scale (9–12% across conditions), and the largest number of scores consistently falling in the 2 and 4 categories (49–54% across conditions). The changes in means shown in Table 1 correspond to changes across conditions in the ratio of low ratings (1–2) to high ratings (4–5). Although the large sample of responses justifies our use of parametric statistical tests, we also conducted non-parametric chi-square analyses on the counts of low and high ratings in the scrambled and canonical word orders. These tests corroborated the finding that there were more high ratings in the scrambled conditions, an effect that was reliable for both levels of the NP-type factor.

in Japanese characters was very similar for the gender-matched NPs (4.27 characters) and the gender-mismatched NPs (4.29 characters), and thus it is unlikely that length differences contributed to any reading-time differences. The critical NP was separated from the clause-final verb by two additional phrases to make it possible to clearly distinguish any processing disruption occurring before the verb from effects that begin at the verb position. In all conditions the clause containing the pronoun and the critical NP was a preposed embedded clause that was followed by all main-clause material. The presence of the additional clause ensured that every target item could contain a gender-compatible antecedent for the pronoun, even if the clausemate NP was gender mismatched or was a dispreferred antecedent. Preposing of embedded clauses is common in Japanese, likely because it minimizes center embedding and hence reduces complexity. However, readers had no indication that the opening regions of the sentence formed part of an embedded clause at the point when they read these regions. The first verb always bore the interrogative suffix *-ka* that is used to mark the scope of a *wh*-phrase in Japanese, consistent with the bias for local scope licensing in the processing of Japanese *wh*-questions (Miyamoto & Takahashi 2002, Aoshima et al. 2004). Across all target items, there were equal numbers of male and female pronouns and NPs. Twenty-four sets of items were distributed among four lists in a Latin Square design. Each participant saw one of the lists intermixed with 56 unrelated items in a pseudorandom order. The materials for this and all other studies in this article can be accessed at the third author's website.

(16) a. Scrambled, mismatch

Daidokoro-de kare-no dono-kodomo-ni tyoosyoku-go oba-ga  
 kitchen-at him-GEN which child-DAT breakfast-afte aunt-NOM  
 isoide obentoo-o watasita-ka titioya-wa oboeteita.  
 in-hurry lunchbox-ACC handed-Q father-TOP remembered  
 'In the kitchen the father remembered to which of his children the  
 aunt handed a lunch box in a hurry after breakfast.'

b. Scrambled, match

Daidokoro-de kare-no dono-kodomo-ni tyoosyoku-go oji-ga  
 kitchen-at him-GEN which child-DAT breakfast-after uncle-NOM  
 isoide obentoo-o watasita-ka oba-wa oboeteita.  
 in-hurry lunchbox-ACC handed-Q aunt-TOP remembered  
 'In the kitchen the aunt remembered to which of his children the  
 uncle handed a lunch box in a hurry after breakfast.'

c. Canonical, mismatch

Daidokoro-de kare-no dono-kodomo-ga tyoosyoku-go oba-ni  
 kitchen-at him-GEN which child-NOM breakfast-after aunt-DAT  
 isoide obentoo-o watasita-ka titioya-wa oboeteita.  
 in-hurry lunchbox-ACC handed-Q father-TOP remembered  
 'In the kitchen the father remembered which of his children  
 handed a lunch box to the aunt in a hurry after breakfast.'

## d. Canonical, match

Daidokoro-de kare-no dono-kodomo-ga tyoosyoku-go oji-ni  
 kitchen-at him-GEN which child-NOM breakfast-after uncle-DAT  
 isoide obentoo-o watasita-ka titioya-wa oboeteita.  
 in-hurry lunchbox-ACC handed-Q father-TOP remembered  
 'In the kitchen the father remembered which of his children handed  
 a lunch box to the uncle in a hurry after breakfast.'

2.4 *Experiment 1B: Procedure*

Participants in this experiment were timed in a phrase-by-phrase self-paced noncumulative moving-window reading task (Just, Carpenter & Woolley 1982). Sentences were presented using Japanese characters. Stimulus segments initially appeared as a row of dashes, and participants pressed the space bar of the keyboard to reveal each subsequent region of the sentences.

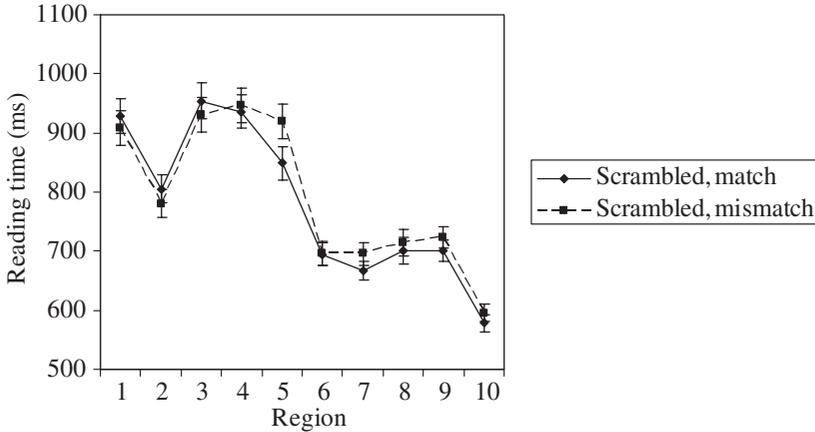
To ensure that participants attended to the stimuli, a subject-verb matching task was presented after each trial. A verb was displayed on the computer screen followed by two agent NPs, corresponding to the nominative-marked NP and the topic-marked NP in the target sentence, and participants had to decide which of the NPs was the subject of the verb in the sentence just read by pressing one of two keys on the keyboard. The two agent NPs were displayed without case marking to avoid case-based question-answering strategies. The task was adopted from Nagata (1993) and was similar to the task used in studies by Miyamoto and Takahashi (2000, 2003). This task was chosen instead of a standard yes/no comprehension question task due to the inclusion of *wh*-interrogative phrases in the target items. Visual feedback indicated whether the answer given was incorrect. The experimental trials were preceded by two screens of instructions and five practice trials.

2.5 *Experiment 1B: Results*

Comprehension accuracy and reading times at each region were entered into a repeated-measures ANOVA, with word order (scrambled, canonical) and gender-match (match, mismatch) as within-subjects factors. All data from participants whose comprehension task accuracy was below 80% for target sentences and below 85% in total were discarded. Four participants failed to meet this criterion (9.1%). Among the 40 participants included in the analysis, average comprehension accuracy was 95.7%. The average response accuracy did not differ significantly across the four conditions (all  $F_s < 1$ ). Additionally, there was no significant difference in the comprehension task accuracy between the participants who live in Shizuoka, Japan (95.0%) and those who live in Maryland, USA (96.3%).

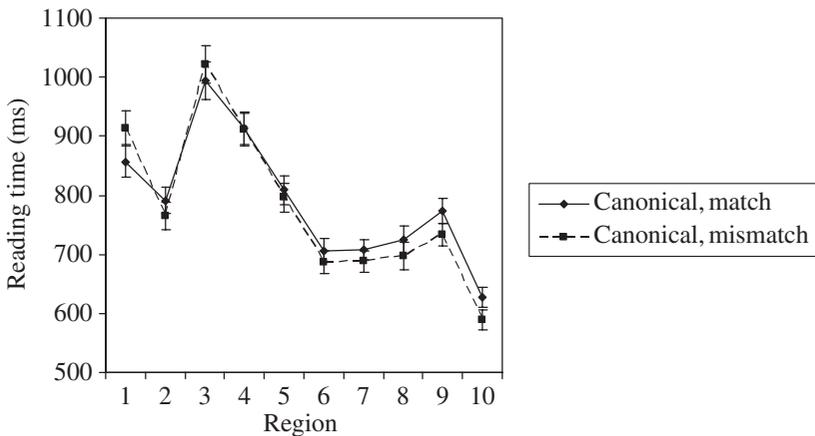
In the analysis of reading time data, reading times longer than 2300 ms were discarded, affecting 3.2% of trials. All trials on which the comprehension

question was answered incorrectly were excluded. The means and analyses we present are based on the remaining trials. Reading times for the conditions with scrambled word order are shown in Figure 1, and for the conditions with canonical word order in Figure 2.



(Adverb<sub>1</sub> His/Her<sub>2</sub> Wh-NP-dat<sub>3</sub> Adverb<sub>4</sub> NP-nom<sub>5</sub> Adverb<sub>6</sub> NP-acc<sub>7</sub> Verb-Q<sub>8</sub> NP-top<sub>9</sub> Verb<sub>10</sub>)

**Figure 1. Experiment 1B: reading times per region for the scrambled conditions.**



(Adverb<sub>1</sub> His/Her<sub>2</sub> Wh-NP-nom<sub>3</sub> Adverb<sub>4</sub> NP-dat<sub>5</sub> Adverb<sub>6</sub> NP-acc<sub>7</sub> Verb-Q<sub>8</sub> NP-top<sub>9</sub> Verb<sub>10</sub>)

**Figure 2. Experiment 1B: reading times per region for the canonical order conditions.**

At all regions preceding the critical NP in region 5 there were no significant differences between reading times as a function of word order or gender match (all  $F_s < 1$ ). At region 5, which contained the critical NP that either matched or mismatched the gender of the pronoun in region 2, there was a significant main effect of gender match, due to slower average reading times in the mismatching conditions,  $F_1(1,39) = 5.32$ ,  $MSE = 109279$ ,  $p < .05$ ;  $F_2(1,23) = 4.7$ ,  $MSE = 164470$ ,  $p < .05$ . There was also a significant main effect of word order due to slower reading times in the scrambled conditions,  $F_1(1,39) = 13.48$ ,  $MSE = 160597$ ,  $p < .001$ ;  $F_2(1,23) = 32.42$ ,  $MSE = 56059$ ,  $p < .001$ . The interaction of gender match and word order was marginally significant,  $F_1(1,39) = 3.89$ ,  $MSE = 85768$ ,  $p = .06$ ;  $F_2(1,23) = 4.4$ ,  $MSE = 97631$ ,  $p = .05$ . However, pairwise comparisons within each level of the word order factor revealed that the gender mismatch effect was restricted to the scrambled conditions, where the mismatching condition was read on average 102ms more slowly than the matching condition,  $F_1(1,39) = 8.65$ ,  $MSE = 98367$ ,  $p < .01$ ;  $F_2(1,23) = 7.43$ ,  $MSE = 153593$ ,  $p < .01$ , and there was no gender mismatch effect in the canonical conditions ( $F_s < 1$ ). This indicates that the two main effects were due to a contrast that was present only in the scrambled conditions.

From region 6 to region 8, there were no significant differences between reading times in the matching and mismatching conditions (all  $F_s < 1$ ). However, at the main-clause subject NP in region 9 there was a main effect of word order that was significant in the participants analysis, but not in the items analysis,  $F_1(1,39) = 4.26$ ,  $MSE = 70222$ ,  $p = .05$ ;  $F_2(1,23) = 1.94$ ,  $MSE = 155706$ ,  $p = .18$ ). There was no significant main effect of gender match ( $F_s < 1$ ), but there was an interaction of gender match and word order that was significant in the participants analysis and only marginally significant in the items analysis,  $F_1(1,39) = 6.94$ ,  $MSE = 63579$ ,  $p < .05$ ;  $F_2(1,23) = 3.3$ ,  $MSE = 120026$ ,  $p = .08$ . Pairwise comparisons within each level of the word order manipulation showed that in the canonical order conditions the matching condition was read significantly more slowly than the mismatching condition,  $F_1(1,39) = 4.61$ ,  $MSE = 100673$ ,  $p < .05$ ;  $F_2(1,23) = 6.74$ ,  $MSE = 68300$ ,  $p < .05$ , and there was no corresponding difference in the scrambled conditions ( $F_s < 1$ ).

A similar trend was revealed at the main verb in the final region (region 10). Although no main effect was observed for either word order,  $F_1(1,39) = 1.56$ ,  $MSE = 60237$ ,  $p = .21$ ;  $F_2(1,23) = 1.94$ ,  $MSE = 43749$ ,  $p = .18$ , or gender match,  $F_1(1,39) = 1.2$ ,  $MSE = 45774$ ,  $p = .28$ ;  $F_2 < 1$ , there was a significant interaction of gender match and word order,  $F_1(1,39) = 4.56$ ,  $MSE = 56240$ ,  $p < .05$ ;  $F_2(1,23) = 4.37$ ,  $MSE = 55900$ ,  $p < .05$ . Pairwise comparisons within each level of the word order factor showed that in the canonical conditions the matching condition was read more slowly than the mismatching condition,  $F_1(1,39) = 4.96$ ,  $MSE = 55137$ ,  $p < .05$ ;  $F_2(1,23) = 4.26$ ,  $MSE = 61016$ ,  $p = .05$ , with no corresponding slowdown in the scrambled conditions ( $F_s < 1$ ). Pairwise comparisons within each level of the gender match factor

yielded no significant differences ( $F_s < 1$ ). The reading time effect at this region likely reflects the continued cost of the disruption observed at the preceding region.

## 2.6 Discussion

The main finding of the off-line test of coreference judgments was confirmation of our prediction that Japanese speakers show higher acceptance of backward anaphora in scrambled word orders. This effect holds irrespective of the type of NP that the pronoun is embedded inside. However, the existence of this contrast is consistent with multiple accounts of the advantage for the scrambled order. Under one account the advantage holds because the scrambled word order is derived from a canonical word order, in which the pronoun would occupy a position that both follows and is c-commanded by the antecedent. Under an alternative account the advantage reflects a preference for pronouns to have antecedents in “prominent” structural or discourse roles, where subject NPs count as particularly prominent arguments. Also, it should be noted that although the advantage for scrambled word order was reliable, it was a far smaller effect than the effect of clear syntactic constraints on backwards anaphora such as condition C (Kazanina et al. 2007). This should not be surprising, given that neither precedence nor c-command are requirements for coreference relations.

The results of the on-line task showed that Japanese speakers were sensitive to the gender match between a pronoun and a subsequent NP in the same scrambled word order configuration, where they showed greater acceptance of coreference in the off-line task. This suggests that the parser actively attempted to construct coreference relations in the scrambled condition, not only before reaching the verb but even before fully processing the gender of the critical NP, paralleling findings about the processing of backwards anaphora in English (van Gompel & Liversedge 2003, Kazanina et al. 2007). In contrast, in conditions with canonical word order there was no gender mismatch effect, which suggests that the parser made no attempt to build a coreference relation between the pronoun and the critical NP in these conditions. This parallels the lower acceptability ratings for backwards anaphora in the canonical order conditions. Taken together, these results indicate that Japanese speakers incrementally build compositional structures involving relations between multiple NPs, and that they do not need to wait for a verb to begin structure building. Furthermore, the results suggest that the representations that are built on-line are sufficiently detailed to capture the contrast in the acceptability of backwards anaphora in scrambled and word orders.

Although there was no effect of gender congruency in the canonical order conditions at the critical NP in region 5, we did observe a gender match penalty in those conditions at the main clause subject NP in region 9 (and continuing to region 10). If we are correct in our conclusion that participants

did not construct an anaphoric dependency at region 5 in these conditions, then region 9 is the first region where an accessible antecedent for the pronoun is encountered. It is unlikely that the gender match penalty merely reflects the cost of long-distance dependency formation, since this region is identical in the matching and mismatching conditions. It is possible that it reflects at least temporary consideration of the earlier NP as a potential antecedent in the gender match condition, leading to a competition-induced slowdown. This may be related to late processing effects of grammatically inaccessible antecedents in Sturt's (2003) study of forwards anaphora processing in English.

The findings from this study go beyond previous studies of incremental parsing in head-final languages in the respect that we demonstrate selective building of compositional relations among NPs in advance of the verb. However, the results are also tempered by concerns about the specific instance of coreference that was tested. First, the acceptability rating task showed the predicted contrast between scrambled and canonical word orders, but ratings for even the most highly rated condition was relatively low (mean = 3.53 on a 1–5 scale). The ratings for the target sentences, all of which probed backwards anaphora interpretations, may have been reduced due to the inclusion in the filler items of cases of forwards anaphora, which yielded a mean rating of 4.14. The inclusion of very easy cases of coreference may have lowered the scores for more difficult backwards anaphora examples. A second, more serious concern involves the use of the personal pronouns *kare* and *kanojo*. The use of these pronouns was well motivated based on the fact that they are the only gender-specific third-person pronouns in Japanese, but it is also the case that they occur relatively infrequently with intrasentential antecedents in naturally occurring Japanese, and are frequently used deictically or with a discourse antecedent. This may make it somewhat unnatural for Japanese speakers to encounter *kare* or *kanojo* in backwards anaphora contexts. Finally, it is also a concern that the study focused on constraints on coreference relations, which may prefer but certainly do not require c-command between the antecedent and the pronoun. Therefore, although the reading-time results in this study were rather clear, the syntactic basis for the study was less clear. In sections 3 and 4 we address this concern by turning to studies of bound-variable anaphora in Japanese, which is subject to a more clear-cut grammatical requirement on c-command relations.

### 3. Binding and Coreference in Japanese

#### 3.1 Referential versus Bound-Variable Interpretations

This section documents the difference between Japanese pronouns that allow bound-variable interpretations and those that do not, and presents results from two acceptability rating studies that test whether scrambling impacts bound-

variable anaphora more strongly than the coreference relations investigated in experiment 1. These studies set the stage for an on-line study of Japanese bound-variable anaphora in section 4.

Although the results of experiment 1 suggest that Japanese speakers are selective in constructing coreference relations online, the specific acceptability contrast tested in experiment 1 is less clear syntactically. It is well known that coreference relations are available in a wider range of configurations than bound-variable relations. The English pronoun in (17a) may corefer with the non c-commanding main clause subject NP, but the corresponding pronoun in (17b) cannot be understood as a bound variable.

- (17) a. While he<sub>i</sub> was in DC, John<sub>i</sub> visited the White House.  
 b. \*While he<sub>i</sub> was in DC, every tourist<sub>i</sub> visited the White House.

However, *kare* and *kanojo* do not readily allow bound-variable interpretations, as has often been pointed out (e.g., Kuroda 1965; Hoji 1985, 1991, 1995; Nakayama 1982; Noguchi 1997; Saito & Hoji 1983). This contrasts with the Japanese anaphor *zibun* 'self', which does allow bound-variable readings. Example (18a) shows that when either *kare* or *zibun* appears in embedded subject position they may corefer with a referential main clause subject, but (18b) shows that in the same configuration *zibun* may be bound by a quantificational antecedent, but *kare* cannot. *Zibun* is, however, unsuitable for our online studies using the gender mismatch paradigm, since it is not semantically selective.

- (18) a. John<sub>i</sub>-ga zibun<sub>i</sub>-ga/ kare<sub>i</sub>-ga Mary-ni kirawareteiru to  
 John-NOM self-NOM he-NOM Mary-by be-disliked COMP  
 omoikondeiru (koto).  
 be-convinced fact  
 'John<sub>i</sub> is convinced that he<sub>i</sub> is disliked by Mary.'
- b. Daremo<sub>i</sub>-ga zibun<sub>i</sub>-ga/ \*kare<sub>i</sub>-ga Mary-ni kirawareteiru  
 everyone-NOM self-NOM he-NOM Mary-by be-disliked  
 to omoikondeiru (koto).  
 COMP be-convinced fact  
 'Everyone<sub>i</sub> is convinced that he<sub>i</sub> is disliked by Mary.'

In light of the difficulty of bound-variable readings for *kare* and *kanojo*, it is unsurprising that these pronouns do not yield binding reconstruction effects, as shown by examples like (19), in which the pronoun inside a scrambled phrase cannot be bound by a following quantificational subject NP (Hoji et al. 2000). This contrasts with the English pronoun *him*, as shown by the translation of (19), and also contrasts with the backwards coreference relations with referential antecedents tested in experiment 1. Additionally, *kare* and *kanojo* allow non-c-commanding antecedents (20).

- (19) \*?[Sensyuu kare<sub>i</sub>-o suisensita sensei-ni] dono  
 last-week he-ACC recommended teacher-DAT every  
 gakusei<sub>i</sub>-mo orei-o okutta.  
 student-mo present-ACC sent  
 'To the teacher who recommended him<sub>i</sub> last week every student<sub>i</sub>  
 sent a present.'
- (20) John<sub>i</sub>-no hahaoya-wa [Mary-ga kare<sub>i</sub>(no-koto)-o suki-da-to]  
 John-GEN mother-TOP Mary-NOM he-ACC like-COMP  
 omotteiru.  
 think  
 'John<sub>i</sub>'s mother thinks that Mary likes him<sub>i</sub>.'

The evidence for on-line, preverbal application of constraints on anaphora in Japanese would be strengthened if it could be shown that structure building is sensitive to clearer structural conditions on binding, such as those that apply to bound-variable anaphora. Fortunately, Japanese has another pronoun, *soko*, that allows bound-variable interpretations when c-commanded by a quantificational antecedent, and that is semantically selective and hence suitable for our experimental paradigm.<sup>3</sup>

Kuroda (1979) observed that Japanese has at least two series of expressions that correspond to English demonstrative expressions such as *that man*. The first series consists of demonstratives that begin with *so-*, as in *so-itu* 'that person' or *so-ko* 'that place'; the second series consists of demonstratives that begin with *a-*, as in *a-itu* 'that person' or *a-soko* 'that place'. Subsequent work by Ueyama, Hoji, and their colleagues has provided a detailed account of these two series of demonstratives and has demonstrated that only demonstratives from the *so*-series allow bound-variable readings (e.g. Hoji 1991, 1995, 2002; Hoji et al. 2000; Takubo & Kinsui 1998; Ueyama 1998, 2000).

Ueyama (1998, 2000) argues that the *so*-words and *a*-words differ both in their referential use and in the availability of bound-variable readings. *So*-words and *a*-words can both be used as referential pronouns, but *a*-words can refer to an individual that has been mentioned in the preceding discourse only if the speaker knows the identity of that referent. In contrast, *so*-words can be used anaphorically to refer to an individual regardless of whether the speaker knows the identity of that individual. For example, the discourse in (21) illustrates a situation where a speaker refers to a person who has been introduced in the discourse, but whose identity the speaker does not know. In this context, the *a*-word *aitu* is unnatural, but the *so*-word *soitu* is entirely natural.

<sup>3</sup> The following discussion is based largely on work by Hajime Hoji and Ayumi Ueyama (Ueyama 1998, 2000; Hoji et al. 2000), to whom we are grateful for extensive discussions of these phenomena.

- (21) [Situation: A wife told her husband on the phone that someone had called for him. He does not know who the person is. The husband then asks his wife:]  
 Soitu<sub>i</sub> /??Aitu<sub>i</sub>-wa nante itteta?  
 that guy-TOP what said  
 ‘What did he say?’

Ueyama also points out that *a*-words may be used at the very beginning of a discourse whereas *so*-words cannot, even if the speaker knows the referent. In (22), for instance, *ano* can refer to a student who visited his professor, without any prior discourse context, whereas *sono* is inappropriate in that context.

- (22) [Situation: A professor asks his secretary about the name of the student who visited him yesterday. He asks the secretary:]  
 Kinoo kita #sono/ano gakusei-no namae nan datta?  
 Yesterday came that student-GEN name what COPULA  
 ‘What is the name of the student who came yesterday?’

Whereas the *so*-word *soko* allows a bound-variable interpretation (23a), the *a*-word *asoko* does not (23b). Furthermore, *soko* allows a bound-variable interpretation only when it is c-commanded by its antecedent, as shown by the examples in (24), where *soko* is embedded inside a subject NP.

- (23) a. Dono kaisya<sub>i</sub>-mo soko<sub>i</sub>-no kogaisya-o suisensita.  
 b. \*Dono kaisya<sub>i</sub>-mo asoko<sub>i</sub>-no kogaisya-o suisensita  
 Every company-mo that place-GEN subsidiary-ACC recommended  
 ‘Every<sub>i</sub> company recommended its<sub>i</sub> subsidiary.’
- (24) a. ?\*Soko<sub>i</sub>-no bengosi-ga [Toyota-to Nissan<sub>i</sub>-to  
 that place-GEN attorney-NOM Toyota and Nissan with  
 torihiki-ga aru kaisya]-o suisensita.  
 business-NOM exist company-ACC recommended  
 Intended but unavailable meaning: ‘Its attorney recommended the company that is doing business with Toyota and Nissan.’  
 b. ?\*Soko<sub>i</sub>-no bengosi-ga [subete<sub>i</sub>-no  
 that place-GEN attorney-NOM all-GEN  
 zidoosya-gaisya-no raibaru gaisya]-o uttaeteiru.  
 automobile-company-GEN rival company-ACC sued  
 Intended but unavailable meaning: ‘Its attorney sued the rival company of each automobile company.’

In contrast, backwards binding of *soko* is possible when it appears inside a fronted (scrambled) object NP (25a), presumably due to the fact that the object NP may be interpreted in its canonical argument position, where it is preceded and c-commanded by the quantificational antecedent. Unsurprisingly, backwards binding is not possible with *asoko* (25b).

- (25) a. [Soko<sub>i</sub> –no kogaisya]-o dono zidoosya gaisya<sub>i</sub>-mo  
 that-place-GEN subsidiary-ACC every auto-company-MO  
 suisensita.  
 recommended
- b. \*[Asoko<sub>i</sub>-no kogaisya]-o dono zidoosya gaisya<sub>i</sub>-mo  
 that place-GEN subsidiary-ACC every auto-company-MO  
 suisensita.  
 recommended  
 ‘Every auto company recommended its subsidiary.’

In sum, *soko* appears to have the appropriate properties for a stronger test of incremental, accurate construction of binding relations in Japanese. We conducted two off-line studies to test the reliability of these properties in native speakers of Japanese, before proceeding to the on-line study reported in section 4.

### 3.2 Experiment 2A: Bound-Variable Interpretation of *Soko* versus *Asoko*

Experiment 2A used an off-line binding acceptability judgment task similar to experiment 1A to test the claim that *soko* allows bound-variable interpretations in scrambled sentences whereas *asoko* does not. The experiment consisted of four conditions organized in a 2 × 2 factorial design that manipulated the order of case markers (‘word order’: canonical NOM-ACC order vs. scrambled ACC-NOM order) and the form of the demonstrative pronoun (*soko* vs. *asoko*). All target items began with a genitive demonstrative pronoun (*a*)*soko* that was embedded inside a larger NP. That NP was followed by a quantificational NP that served as the potential antecedent for the demonstrative pronoun. As in experiment 1A, we predicted that backwards anaphora should be possible only in the scrambled word order. Backwards binding should be impossible in the conditions with canonical word order, since the pronoun is not c-commanded by the quantificational NP. Furthermore, *asoko* is predicted to disallow backwards binding, irrespective of word order. Twelve sets of items were distributed among four lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with 24 filler items in a random order. One full set of experimental condition is shown in (26).

- (26) a. *Soko*-scrambled  
 Shinbunkiji-niyoruto [soko<sub>i</sub>-no itiban yakunitatanai  
 news article-by that place-GEN most useless  
 syain]-o [dono jidoosyagaisya]<sub>i</sub>-mo kubinsita toiu  
 employee-ACC every automobile company-MO fired  
 hookoku-ga dehajimeta-rasii.  
 saying report-NOM appear-began-seem  
 ‘According to a newspaper article, it seems that a report emerged  
 that every automobile company<sub>i</sub> fired its<sub>i</sub> most useless employee.’
- b. *Soko*-canonical  
 Shinbunkiji-niyoruto [soko<sub>i</sub>-no itiban yakunitatanai  
 news article-by that place-GEN most useless  
 syain]-ga [dono jidoosyagaisya]<sub>i</sub>-mo uttaeta  
 employee-ACC every automobile company-MO sued  
 toiu hookoku-ga dehajimeta-rasii.  
 saying report-NOM appear-began-seem.  
 ‘According to a newspaper article, it seems that a report emerged  
 that its<sub>i</sub> most useless employee sued every automobile company<sub>i</sub>.’
- c. *Asoko*-scrambled  
 Shinbunkiji-niyoruto [asoko<sub>i</sub>-no itiban yakunitatanai  
 news article-by that place-GEN most useless  
 syain]-o [dono jidoosyagaisya]<sub>i</sub>-mo kubinsita  
 employee-ACC every automobile company-MO fired  
 toiu hookoku-ga dehajimeta-rasii.  
 saying report-NOM appear-began-seem  
 ‘According to the newspaper article, it seems that a report  
 emerged that every automobile company<sub>i</sub> fired its<sub>i</sub> most  
 useless  
 employee.’
- d. *Asoko*-canonical  
 Shinbunkiji-niyoruto [Asoko<sub>i</sub>-no itiban yakunitatanai  
 news article-by that place-GEN most useless  
 syain]-ga [dono jidoosyagaisya]<sub>i</sub>-mo uttaeta  
 employee-NOM every automobile company-MO sued  
 toiu hookoku-ga dehajimeta-rasii.  
 saying report-NOM appear-began-seem  
 ‘According to the newspaper article, it seems that a report emerged  
 that its<sub>i</sub> most useless employee sued every automobile company<sub>i</sub>.’

Participants were 28 undergraduate students from Meiji Gakuin University, Japan, who were asked to judge whether an underlined pronoun and an underlined quantificational NP in each sentence could be understood as coreferential, indicating their rating on a scale from 1 (coreference impossible) to 5 (coreference fully acceptable). Participants were paid ¥500 for the

**Table 2. Experiment 2A: mean ratings of the acceptability of anaphoric relations ( $n = 28$ )**

Conditions	Anaphoric relation ratings	
	Mean	SD
a. <i>Soko</i> -scrambled	3.21	1.55
b. <i>Soko</i> -canonical	2.49	1.44
c. <i>Asoko</i> -scrambled	1.71	1.07
d. <i>Asoko</i> -canonical	1.95	1.29

experiment, which lasted around 20 minutes. Mean acceptability ratings in each condition were entered into a repeated-measures ANOVA, with pronoun (*soko*, *asoko*) and word order (scrambled, canonical) as within-subjects factors. Results are shown in Table 2.

There was a significant main effect of pronoun type due to higher ratings for the *soko* conditions than for the *asoko* conditions,  $F_1(1,27) = 19.1$ ,  $MSE = 4.55$ ,  $p < .001$ ;  $F_2(1,11) = 93.64$ ,  $MSE = 0.93$ ,  $p < .001$ , but there was no significant main effect of word order,  $ps > .1$ . There was a significant interaction of pronoun type and word order,  $F_1(1,27) = 15.33$ ,  $MSE = 1.27$ ,  $p < .001$ ;  $F_2(1,11) = 21.48$ ,  $MSE = 0.91$ ,  $p < .001$ . Pairwise comparisons within each level of the pronoun type factor showed that in the *soko* conditions the scrambled condition was rated significantly higher than the canonical order condition,  $F_1(1,27) = 9.4$ ,  $MSE = 2.36$ ,  $p < .01$ ;  $F_2(1,11) = 11.81$ ,  $MSE = 1.88$ ,  $p < .01$ , but that there was no corresponding difference in the *asoko* conditions,  $ps > .1$ . Thus, this study confirmed that Japanese speakers more readily allow backward bound-variable interpretations for *soko* in scrambled word orders. Although the overall ratings for backwards anaphora were still relatively low in this study, possibly due to the inclusion of forwards anaphora examples as filler items in the study, the large difference in acceptability ratings between *soko* and *asoko* lends support to the claim by Ueyama and colleagues that *soko* allows bound-variable readings in Japanese whereas *asoko* does not.

### 3.3 Experiment 2B: Referential versus Bound-Variable Readings

Experiment 2B was an acceptability rating study focusing on backwards anaphora in scrambled word orders. The aim was to verify the claim that *soko* allows both bound-variable and coreferential readings in backwards anaphora, whereas *kare/kanajo* allow only coreferential readings. The four conditions manipulated the type of the pronoun (*kare/kanajo* vs. *soko*) and the type of the potential antecedent NP (referential NP vs. quantificational NP). All target sentences started with a genitive pronoun embedded inside a scrambled dative NP. The dative NP was followed by a subject NP that was either referential or quantificational. Twelve pairs of items were constructed for *kare/kanajo* (six for each pronoun) and 12 pairs for *soko*. The test items for *kare/kanajo* and *soko* could not be fully matched due to the differing semantic restrictions on

the two types of pronouns. The 24 pairs of items were distributed among two lists in a Latin Square design. Each participant saw one of the lists intermixed with 48 filler items in a random order. A sample pair of items for each type of pronoun is shown in (27) and (28).

- (27) a. *Kare*-referential NP  
 Kare-no seito-ni sono sensei-wa kyoositu-de  
 he-GEN student-DAT the teacher-TOP classroom-at  
 jisyo-o watasita.  
 dictionary-ACC gave  
 'The teacher gave a dictionary to his student(s) at the classroom.'
- b. *Kare*-quantificational NP  
 Kare-no seito-ni dono sensei-mo kyoositu-de  
 he-GEN student-DAT every teacher-MO classroom-at  
 jisyo-o watasita.  
 dictionary-ACC gave  
 'Every teacher gave a dictionary to his student(s) at the classroom.'
- (28) a. *Soko*-referential NP  
 Soko-no syain-ni sono  
 that place-GEN employee-DAT the  
 hokengaisya-wa cyoorei-de  
 health insurance company-TOP morning meeting-at  
 bunsyo-o watasita.  
 document-ACC gave  
 'The health insurance company gave a document to its employees  
 at the morning meeting.'
- b. *Soko*-quantificational NP  
 Soko-no syain-ni dono  
 that place-GEN employee-DAT every  
 hokengaisya-mo cyoorei-de  
 health insurance company-MO morning meeting-at  
 bunsyo-o watasita.  
 document-ACC gave  
 'Every health insurance company gave a document to its employees  
 at the morning meeting.'

Participants were 44 undergraduate students from Shizuoka University, Japan, who received ¥500 for completing the study, which lasted around 25 minutes. The task was the same as experiments 1A and 2A. Mean acceptability ratings for each condition were entered into a repeated-measures ANOVA, with pronoun (*kare*, *soko*) and antecedent type (referential NP, quantificational NP) as within-subjects factors. Due to lexical differences between the *kare* and the *soko* conditions we report items

**Table 3. Experiment 2B: mean ratings of the acceptability of backwards anaphoric relations ( $n = 44$ )**

Conditions	Anaphoric relation ratings	
	Mean	SD
a. <i>Kare</i> -referential NP	4.08	1.39
b. <i>Kare</i> -quantificational NP	3.45	1.54
c. <i>Soko</i> -referential NP	4.30	1.19
d. <i>Soko</i> -quantificational NP	3.94	1.33

analyses for only the comparisons within each level of the pronoun-type factor. Results are shown in Table 3.

There was a significant main effect of pronoun type due to higher ratings for the *soko* conditions,  $F_1(1,43) = 17.89$ ,  $MSE = 1.85$ ,  $p < .001$ , a significant main effect of antecedent type, due to higher ratings for conditions with referential antecedents,  $F_1(1,43) = 14.52$ ,  $MSE = 4.44$ ,  $p < .001$ , and a marginally significant interaction of pronoun type and antecedent type,  $F_1(1,43) = 3.67$ ,  $MSE = 1.23$ ,  $p = .06$ . Pairwise comparisons within each level of the pronoun type factor showed that the referential condition was rated significantly higher than the quantificational condition, both in the *kare* conditions,  $F_1(1,43) = 14.05$ ,  $MSE = 1.05$ ,  $p < .001$ ;  $F_2(1,23) = 33.0$ ,  $MSE = 1.56$ ,  $p < .001$ , and in the *soko* conditions,  $F_1(1,43) = 8.72$ ,  $MSE = 2.0$ ,  $p < .01$ ;  $F_2(1,23) = 15.5$ ,  $MSE = 1.17$ ,  $p < .001$ . Planned comparisons within each level of the antecedent type factor showed that the *soko* condition received higher ratings in the referential antecedent conditions,  $F_1(1,43) = 4.93$ ,  $MSE = 1.34$ ,  $p < .05$ , and in the quantificational antecedent conditions,  $F_1(1,43) = 17.8$ ,  $MSE = 1.74$ ,  $p < .001$ .

These results conform to our predictions. Native speakers of Japanese in general give higher acceptability ratings for backwards coreference relations than for backwards bound-variable relations. This is consistent with findings from English processing studies that have reported that speakers prefer to interpret an ambiguous anaphoric dependency as a coreference dependency rather than as an operator-variable dependency (Shapiro & Hestvik 1995, Frazier & Clifton 2000, Shapiro et al. 2003). However, the contrast in ratings was much larger for the personal pronouns *kare/kanojo* than for the demonstrative pronoun *soko*. This supports the claim that *soko* allows bound-variable readings, whereas *kare/kanojo* do not (Hoji 1985, 1991; Kitagawa 1981; Nakayama 1982; Noguchi 1997; Saito & Hoji 1983). Note that the overall ratings are higher in experiment 2B than in experiments 1A and 2A, with the lowest rating in 2B being higher than the highest rating in the other studies. We suggest that the rating scores should be used only as comparative measures within a single study. Comparison of scores across these studies is probably inappropriate due to the impact on scores of the distribution of acceptable and unacceptable target and filler items.

Taken together, experiments 2A and 2B indicate that the Japanese demonstrative pronoun *soko* allows backwards binding relations specifically in scrambled word orders, consistent with a standard c-command condition on operator-variable binding relations. This property, together with its semantic selectivity, makes *soko* appropriate for an on-line test of the construction of binding relations.

#### 4. Experiment 3: Real-time Interpretation of the Bound-Variable Pronoun *Soko*

##### 4.1 Experiment 3A: Rationale and Methods

The aim of this experiment was to further test whether Japanese speakers incrementally and accurately construct binding relations among NPs in advance of a verb, taking advantage of the grammatical properties of the demonstrative pronoun *soko*, as described in section 3. *Soko* may be interpreted as a bound variable when it is c-commanded by a quantificational antecedent. As shown in experiment 2A, this includes configurations where *soko* appears inside an NP that has been scrambled to a position that precedes the antecedent of *soko* (29a), and excludes configurations that are superficially similar but do not involve scrambling (29b).

- (29) a. Soko-no           sisyo-ni           dono   tosyokan-mo  
           that place-GEN   librarian-DAT   every   library-MO  
           manyuaru-o   kubatta.  
           manual-ACC   distributed  
           ‘Every library distributed a manual to its librarian(s).’
- b. \*Soko-no           sisyo-ga           dono   tosyokan-ni-mo  
           that place-GEN   librarian-NOM   every   library-dat-MO  
           manuaru-o   seikyuusita.  
           manual-ACC   requested  
           ‘Its librarian(s) requested a manual from every library.’

If the parser actively searches for a structurally appropriate antecedent for a pronoun, it should attempt to construct a binding relation between *soko* and the quantificational subject NP in sentences like (29a), but should not attempt to construct a binding relation between the corresponding NPs in examples like (29b). As a diagnostic of on-line construction of binding relations we use a variant on the gender mismatch paradigm used in experiment 1B. *Soko* may be used to refer to an organization such as a company but not to a person such as a teacher. Therefore, if the potential antecedent NP *tosyokan* ‘library’ in (29) is replaced with an NP that is incompatible with *soko*, such as *riyoosya* ‘user’, it should affect reading times only if the parser is considering a binding relation between that NP and *soko*.

The specific reading-time predictions for this study are less straightforward than in the gender mismatch study with *kare/kanojo* in experiment 1B. In that study we claimed that the parser actively attempts to construct a coreference relation between a pronoun and a potential antecedent NP before it has processed the gender of the antecedent. This leads to situations in which the parser initially constructs a coreference relation that it must subsequently retract due to a gender mismatch, leading to disruption in reading times. This is similar to the interpretation of gender mismatch effects in backwards anaphora presented by van Gompel and Liversedge (2003). If the parser actively constructs coreference relations before it has processed the semantic properties of an NP, then it is also possible that the parser might inappropriately construct a coreference relation in a configuration that turns out to require a variable binding relation. This is consistent with the general preference for coreference relations over variable binding relations reported in English (Shapiro & Hestvik 1995, Frazier & Clifton 2000, Shapiro et al. 2003), and is also consistent with the small but significant advantage in acceptability ratings for coreferential *soko* over bound variable *soko* in experiment 2B. Consequently, even in sentences where *soko* is interpreted as a bound variable, appropriately c-commanded by a quantificational subject NP, there might be processing disruption associated with retroactively changing the interpretation of *soko* from a referential pronoun to a bound-variable pronoun. For this reason, we predict that the semantic compatibility of the subject NP with *soko* in sentences like (29a) could be associated with reading-time disruption, without committing to the direction of the reading time effect. In contrast, the prediction for sentences with canonical word order like (29b) is exactly the same as in experiment 1B: the parser should make no attempt to construct a binding relation in these cases, and hence the semantic congruity between *soko* and the following NP should not affect reading times.

Thirty-six native speakers of Japanese participated in the study. All of them were students at Shizuoka University, Shizuoka Sangyo University, or Meiji Gakuin University, Japan. They were paid ¥500 for their participation in the experiment, which lasted about 30 minutes.

The experiment had four conditions organized in a  $2 \times 2$  factorial design that manipulated the semantic congruity of the potential antecedent NP (congruous vs. incongruous) and word order (scrambled vs. canonical). The NP containing *soko* was marked with dative case in the scrambled conditions and with nominative case in the canonical order conditions. In the congruous conditions the quantificational NP referred to an institution, consistent with the semantic properties of *soko*; in the incongruous conditions the quantificational NP referred to a human. As in experiment 1B the critical regions all appeared in a preposed embedded clause. As a result, it was possible to guarantee that all pronouns had a grammatical sentence-internal antecedent, even in the canonical order and incongruent conditions. Twenty-four sets of four conditions were distributed among four

lists in a Latin Square design. Each participant saw exactly one of the lists intermixed with 72 unrelated items in a random order. One full set of experimental conditions is shown in (30). The self-paced reading procedure was identical to that used in experiment 1B, except that this study used a standard yes/no comprehension question task after each trial. The materials for this and all other studies in this article can be accessed from the third author's website.

(30) a. Scrambled incongruous

Keizaishi-de                    soko-no    syain-ni                    fukeeki-de  
business.magazine-in soko-GEN employee-DAT recession-for  
dono    kacuyoo-mo    ookina fuman-o                    motteiru-toiu  
every    manager-MO    serious complaint-ACC    have-that  
keikoo-ga                    aru-rasii-to    aru    kaisya-ga  
tendency-NOM    be-seem-that    some    company-NOM  
hookokusiteiru.

reports

'In a business magazine a company reports that a tendency has been observed for every manager to bring a serious complaint against its employees due to the recession.'

b. Scrambled congruous

Keizaishi-de                    soko-no    syain-ni                    fukeeki-de  
business.magazine-in soko-GEN employee-DAT recession-for  
dono    kaisya-mo                    ookina fuman-o                    motteiru-toiu  
every    company-MO    serious complaint-ACC    have-that  
keikoo-ga                    aru-rasii-to    aru    keieesya-ga  
tendency-NOM    be-seem-that    some    executive-NOM  
hookokusiteiru.

reports

'In a business magazine an executive reports that a tendency has been observed for every company to bring a serious complaint against its employees due to the recession.'

c. Canonical incongruous

Keizaishi-de                    soko-no    syain-ga                    fukeeki-de  
business.magazine-in soko-GEN employee-NOM recession-for  
dono    kacuyoo-ni-mo                    ookina fuman-o                    motteiru-toiu  
every    manager-DAT-MO    serious complaint-ACC    have-that  
keikoo-ga                    aru-rasii-to    aru    kaisya-ga  
tendency-NOM    be-seem-that    some    company-NOM  
hookokusiteiru.

reports

'In a business magazine a company reports that a tendency has been observed for its employees to bring a serious complaint against every manager due to the recession.'

## d. Canonical Congruous

Keizaishi-de soko-no syain-ga fukeeki-de  
 business.magazine-in soko-GEN employee-NOM recession-for  
 dono kaisya-ni-mo ookina fuman-o motteiru-toiu  
 every company-DAT-MO serious complaint-ACC have-that  
 keikoo-ga aru-rasii-to aru kaisya-ga  
 tendency-NOM be-seem-that some company-NOM  
 hookokusiteiru.  
 reports

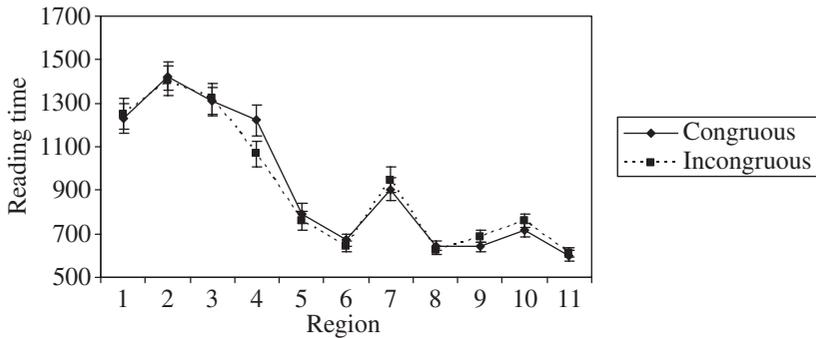
'In a business magazine a company reports that a tendency has been observed for its employees to bring a serious complaint against every company due to the recession.'

## 4.2 Experiment 3A: Results

Comprehension accuracy and reading times at each region were entered into a repeated-measures ANOVA, with word order and congruity as within-subjects factors. All data from participants whose comprehension task accuracy was less than 70% in the target sentences and 80% in total were discarded ( $n = 4$ , 11%). Among the 32 participants included in the analysis, average comprehension accuracy was 83.2%. The average correct response percentage did not differ significantly across the four conditions (all  $F_s < 1$ ). The overall comprehension accuracy in this study may have been lower than in experiment 1B due to the increased complexity of sentences containing quantifiers, and due to the different comprehension task used. Reading times longer than 4000 ms were discarded, affecting 1.6% of trials, and all trials on which the comprehension question was answered incorrectly were excluded. A longer exclusion threshold was used in this study than in experiment 1B, due to the overall increased complexity of the materials and slower reading times.

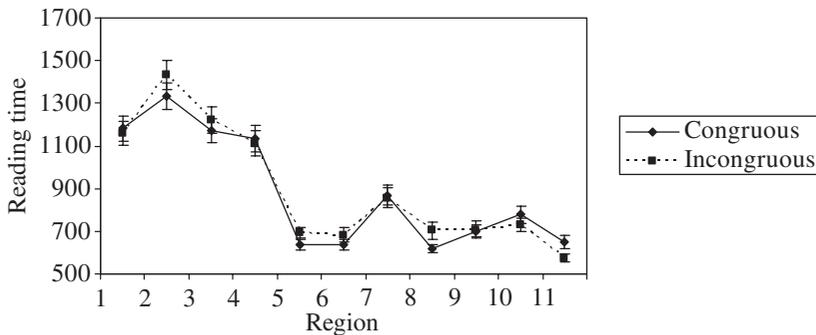
Reading times for the scrambled conditions are shown in Figure 3, and those for the canonical order conditions are shown in Figure 4. At all regions preceding the quantificational NP in region 4 there were no significant differences between reading times as a function of word order or congruency (all  $F_s < 1$ ).

At the quantificational NP in region 4, which varied in its semantic congruity with the pronoun in region 2, there was no significant main effect of word order ( $F_s < 1$ ) or congruity ( $p_s > .15$ ), nor was there an interaction of congruity and word order ( $p_s > .2$ ). However, pairwise comparisons within each level of the word order factor showed that in the scrambled conditions the congruous condition was read 157 ms more slowly than in the incongruous condition. This difference was significant in the participants analysis and marginally significant in the items analysis,  $F_1(1,31) = 4.29$ ,  $MSE = 400514$ ,  $p < .05$ ;  $F_2(1,23) = 3.06$ ,  $MSE = 983830$ ,  $p = .09$ . The relatively weak items effect may reflect greater variance due to the possibility in some items of interpreting *soko* as anaphoric to the sentence-initial adjunct. In contrast, in the



Adverb<sub>1</sub> soko-Gen NP-dat<sub>2</sub> Adverb<sub>3</sub> QNP<sub>4</sub> Adj<sub>5</sub> NP-acc<sub>6</sub> Verb-Comp<sub>7</sub>  
 NP-nom<sub>8</sub> Verb-Comp<sub>9</sub> QNP<sub>10</sub> Verb<sub>11</sub>

**Figure 3. Experiment 3A: reading times per region for the scrambled conditions.**



Adverb<sub>1</sub> soko-Gen NP-nom<sub>2</sub> Adverb<sub>3</sub> QNP<sub>4</sub> Adj<sub>5</sub> NP-acc<sub>6</sub> Verb-Comp<sub>7</sub>  
 NP-nom<sub>8</sub> Verb-Comp<sub>9</sub> QNP<sub>10</sub> Verb<sub>11</sub>

**Figure 4. Experiment 3A: reading times per region for the canonical order conditions.**

canonical order conditions there was no significant difference between congruous and incongruous condition ( $F_s < 1$ ).

There were no significant differences in reading times at regions 5 or 6. At the embedded nominative subject NP in region 8 there was no main effect of word order or congruity, but there was an interaction among these factors that was significant in the participants analysis but only marginally significant in the items analysis,  $F_1(1,31) = 6.45$ ,  $MSE = 69051$ ,  $p < .05$ ;  $F_2(1,23) = 3.29$ ,  $MSE = 109253$ ,  $p = .08$ . Pairwise comparisons within each level of the word order factor showed that in the canonical order conditions the incongruous condition was read more slowly than the congruous condition,  $F_1(1,31) =$

7.89,  $MSE = 89599$ ,  $p < .01$ ;  $F_2(1,23) = 4.16$ ,  $MSE = 123517$ ,  $p = .05$ , and there was no corresponding slowdown in the scrambled conditions. At the sentence-final main verb in region 11 there was a significant interaction of word order and congruency that was significant in the participants analysis but not in the items analysis,  $F_1(1,31) = 4.2$ ,  $MSE = 52765$ ,  $p < .05$ ;  $F_2 < 1$ . Pairwise comparisons within each level of the word order factor revealed a slowdown in reading times for congruous conditions in the canonical order conditions,  $F_1(1,31) = 4.31$ ,  $MSE = 68672$ ,  $p < .05$ ;  $F_2(1,23) = 2.39$ ,  $MSE = 136447$ ,  $p = .1$ , but no corresponding slowdown in the scrambled conditions.

#### 4.3 Experiment 3B: Sentence Completion

Before discussing the results of the on-line experiment we want to address a potential concern about the materials for this study raised by a reviewer. Since the quantificational particle *-mo* is incompatible with the nominative case particle *-ga*, the critical quantificational NP in the scrambled conditions was also compatible with interpretation as an accusative NP, yielding a clause with canonical word order and a null subject. If we were incorrect to assume that participants would parse *dono NP-mo* 'every NP' as nominative, then the logic of the study is undermined. To test this assumption we conducted an additional off-line sentence fragment completion study.

The experimental materials consisted of 20 sets of four sentence fragments derived from the materials for experiment 3A, as illustrated in (31). The four conditions within each set of items manipulated the case of the sentence-initial NP (nominative-initial vs. dative-initial) and the semantic type of the universally quantified second NP (animate vs. organization). The demonstrative pronoun *soko* that appeared in all critical items in experiment 3A was not included in the fragments, since the current study was not directly concerned with backwards anaphora. The 20 sets of items were distributed among four lists in a Latin Square design and combined with 40 filler items of varying degrees of plausibility. Participants were 28 native speakers of Japanese who were students at Hiroshima University, Japan, or residents in the Washington, D.C. area. The experiment was presented in the form of a pencil-and-paper questionnaire in which participants were asked to complete the fragments in a manner that sounded natural to them. The study lasted around 20 minutes.

- (31) a. DAT-initial, animate  
 Hisyo-ni raikanji dono-sinkiriyoosya-mo...  
 librarian-DAT visit-on every new-user...
- b. DAT-initial, organization  
 Hisyo-ni raikanji dono-siritutosyokan-mo...  
 librarian-DAT visit-on every municipal library...
- c. NOM-initial, animate  
 Hisyo-ga raikanji dono-sinkiriyoosya-mo...  
 librarian-NOM visit-on every new-user...

## d. NOM-initial, organization

Hisyo-ga raikanji dono-siritutosyokan-mo...  
 librarian-NOM visit-on every municipal library...

The study yielded a total of 532 codable sentence fragment completions. Information in the sentence completions was used to infer the intended grammatical role of the potentially ambiguous NP. The ambiguous NP could be identified as nominative if a dative-initial fragment was completed with just a verb that selects a dative and a nominative argument, or if the fragment was completed with an accusative NP and a ditransitive verb. We found that 88% of completions in the dative-initial conditions (235/267 trials) followed one of these patterns. In the remaining 12% of completions in the dative-initial conditions (32/267 trials) the ambiguous NP was interpreted as the subject of an embedded clause, as indicated by completions containing two verbs. Fisher Exact tests showed that the proportion of monoclausal versus biclausal completions was unaffected by the semantic type of the NP (animate vs. institution). Results from the nominative-initial conditions showed that even here participants showed a bias to interpret the ambiguous NPs as nominative, as evidenced by a majority of biclausal completions (75.5%, 200/265 trials). Together these results indicate that Japanese speakers show a strong bias to interpret *dono NP-mo* 'every NP' as nominative, and that in dative-initial clauses this leads to a bias to interpret *dono NP-mo* as the subject of a clause with scrambled word order. This confirms the assumption behind experiment 3A.

#### 4.4 Discussion

The main finding of the on-line study was that in the scrambled conditions Japanese readers slowed down upon encountering a quantificational NP that was a semantically congruous potential antecedent for the pronoun *soko*, relative to reading times for semantically incongruous antecedents. No corresponding slowdown was observed in the canonical order conditions. This contrast between scrambled and canonical word orders parallels a similar effect observed for the processing of the pronouns *kare* and *kanojo* in experiment 1B and lends further support to the notion that Japanese speakers are selective in the on-line construction of anaphoric dependencies.

However, the results of this study contrasted with experiment 1B in the respect that the experiment with *kare/kanojo* showed slower reading times for semantically incongruous potential antecedents, whereas the experiment with *soko* showed slower reading times for semantically congruous antecedents. As discussed in section 4.1 the congruity effect in this study may reflect the fact that the Japanese parser initially attempts to construct a coreference relation between *soko* and its antecedent, which subsequently must be reanalyzed as an operator-variable relation once the quantificational NP is fully processed. We

assume that, upon encountering the demonstrative pronoun *soko* in a scrambled phrase, the Japanese parser preferentially interprets the pronoun as referential and then actively posits a coreference dependency with the yet-to-be-encountered subject position. Upon encountering a semantically incongruous subject NP the parser must revise its prediction that the subject NP would provide an antecedent for the pronoun, but it need not revise the interpretation of the pronoun. In contrast, when the parser encounters a semantically congruous subject QP it can confirm its expectation for an antecedent, but it must rescind its prediction of a referential antecedent and it must revise the interpretation of the pronoun *soko* as a bound variable and it must revise the nature of the anaphoric relation. This may account for the surprising semantic congruity effect in this study.

Although there was no effect of semantic congruity at the crucial QP in region 4 in the canonical conditions, there was a slowdown in the incongruous condition at the nominative NP in region 8. Importantly, the NP in this region is identical in both the congruous and incongruous canonical order conditions, and in neither does it provide a suitable antecedent for *soko*. We cannot exclude the possibility that the parser is more disrupted in the incongruous condition at this region because it continues to search for an antecedent in that condition but has considered the earlier NP in region 4 as a possible antecedent in the congruous condition. This would imply that although the parser only actively constructs anaphoric dependencies in the scrambled conditions, it might nevertheless consider dependency formation at a later point in the canonical order conditions. If this occurs, then it could explain the subsequent slowdown at region 11 in the canonical congruous condition, which may reflect recognition of an additional potential antecedent NP in region 10.

Taken together with the findings from experiment 1B, these results indicate that Japanese speakers actively search for the antecedent of a pronoun only in structurally acceptable positions. They are able to construct coreference and operator-variable relations between a pronoun and its antecedent before any verb information is available, consistent with the claims of accounts that predict incremental, accurate structure building in head-final languages.

## **5. Real-time Computation of Syntactic Relations between Noun Phrases**

Across a series of four off-line acceptability rating studies and two on-line reading time studies, we showed that Japanese speakers are sensitive to constraints and biases on the formation of anaphoric dependencies, and that they appear to use these constraints to guide incremental language processing. Importantly, the on-line studies showed that semantic congruity effects occurred immediately upon presentation of a potential antecedent NP, which indicates that dependencies between pronouns and their antecedents are formed before any verb appears in the input. The immediate semantic congruity effects in the processing of backwards anaphora were observed in

sentences with scrambled, noncanonical word order, and were not observed in sentences with canonical subject-first word order. This indicates that the Japanese parser does not blindly search for semantically compatible antecedents for pronouns but rather is sensitive to the case and configurational properties of potential antecedents. Aoshima (2003) presents an explicit account of how this might be achieved, based on an extension of a computational model developed by Schneider (1999).

These results favor accounts of the parsing of head-final languages that emphasize incremental assembly of detailed grammatical structure as each new word is encountered (e.g., Fodor & Inoue 1994, 1998; Inoue & Fodor 1995; Mazuka & Itoh 1995; Kamide & Mitchell 1999; Miyamoto & Takahashi 2002; Mazuka, Itoh & Kondo 2002; Fodor & Hirose 2003; Kamide, Altmann & Haywood 2003; Nakatani & Gibson 2003), rather than models in which structure building is delayed until a clause-final verb is reached (e.g., Pritchett 1991, Mulders 2002). A number of previous studies have presented arguments in favor of incremental structure building in head-final languages (e.g., Bader & Lasser 1994; Mazuka & Itoh 1995; Fanselow, Kliegl & Schlesewsky 1999; Miyamoto & Takahashi 2002), but the current results go beyond the previous arguments in a couple of respects. The crucial processing effects occur well in advance of any verb in the input and are therefore difficult to reinterpret in terms of a verb-driven model of structure building. More importantly, since our results implicate the selective formation of compositional relations (specifically, anaphoric relations) among NPs in advance of the verb, they go beyond previous evidence based on sensitivity to preferred resolution of ambiguous NPs (e.g., Bader & Lasser 1994) or effects of sensitivity to canonical sequencing of NPs (e.g., Miyamoto & Takahashi 2002).

The evidence for preverbal sensitivity to constraints on coreference relations is relevant for recent discussions of the relation between grammatical knowledge and real-time language processes. A number of linguists, working in a diverse set of linguistic paradigms, have attempted to characterize knowledge of grammar as knowledge of how to construct grammatical sentence structures in real time (e.g. Bresnan 1978; Phillips 1996, 2003; Steedman 2000; Kempson, Meyer-Viol & Gabbay 2001; Kempen & Harbusch, 2002), contrasting with the more standard view that a speaker's knowledge of grammar characterizes the well-formedness of sentences but makes no specific commitments about how sentences are constructed in real time. The finding that Japanese speakers build grammatically accurate structures in advance of the verb cannot show that a real-time account of grammatical knowledge is necessary, but it does help to set aside one possible argument against such approaches.

In describing our results we emphasized that the semantic congruency effects observed in experiments 1B and 3A suggest that the Japanese parser actively searches for antecedents for a pronoun, and forms anaphoric relations before it has processed all of the relevant details of a potential antecedent NP. For example, the gender mismatch effect in experiment 1B was explained by

assuming that the parser assigns a nominative subject NP as the antecedent of a fronted pronoun before it has confirmed that the subject NP is compatible with the gender of the pronoun. It should be noted that this account of semantic congruity effects merely provides an upper bound on the time when coreference relations are constructed, and is compatible with a range of more or less “active” accounts of the formation of anaphoric dependencies. Under one scenario the parser waits until it encounters an appropriately case-marked NP before starting to form an anaphoric dependency. Under this account, dependency formation occurs only very shortly before gender and other semantic information is processed (cf. van Gompel & Liversedge 2003). A more aggressively predictive approach is one in which the parser constructs some anaphoric dependencies before it has seen any properties of the antecedent, based only on the properties of the dependent pronoun (Kazanina et al. 2007), implying predictive building of antecedents. This type of highly active dependency formation may be particularly plausible in a situation where the parser encounters a sentence-initial nonsubject NP and can be confident that it will subsequently encounter a subject NP. In such a situation the parser might predict that the subject NP will provide an antecedent for a pronoun that it has seen, before it knows any specific properties of the subject NP. The results of our studies are largely compatible with either the more conservative or the more predictive account of when coreference relations are built. However, it may be easier to capture the contrast between the semantic mismatch penalty for referential dependencies (experiment 1B) and the semantic match penalty for bound-variable anaphora (experiment 3A) in an approach that posits referential dependencies immediately upon encountering a pronoun.

The possibility of predictive building of antecedents for pronouns raises an alternative account of the word order effects in our studies, which deserves to be taken seriously. We argued that the absence of semantic congruity effects in the canonical NOM-DAT word orders in experiments 1A and 3A reflected the Japanese parser’s sensitivity to a bias for preceding and c-commanding referential antecedents (experiment 1B) and to the c-command constraint on bound-variable anaphora (experiment 3A). Under this account, the parser does not immediately consider the dative NP as a potential antecedent for a pronoun because of its configurational properties. However, another description of our results is that active dependency formation effects are found on nominative NPs, which can be reliably predicted to occur in most Japanese clauses, and not on dative NPs, which can be predicted with much lower reliability. This is consistent with the fact that previous demonstrations of gender mismatch effects in backwards anaphora have been restricted to antecedents in subject positions (van Gompel & Liversedge 2003, Kazanina et al. 2007). However, if the parser is able to consider the dative NPs as potential antecedents for *kare/kanojo* (experiment 1B) and *soko* (experiment 3A), and is merely limited in its ability to anticipate the appearance of the dative NP, then we should expect to also find evidence of anaphoric

dependency formation in the canonical conditions in experiment 3A. In that experiment we suggested that reanalysis of *soko* from a referential pronoun to a bound-variable pronoun is responsible for the slowdown in reading times at the semantically congruous antecedent in the scrambled conditions. If the parser is also able to form binding dependencies in the canonical conditions, then we should expect to observe an early semantic congruity effect in those conditions, contrary to fact. (See Kazanina et al. 2007 for studies that distinguish the contributions of predictability and subjecthood in processing English backwards anaphora.)

## 6. Conclusion

Much previous research on the processing of head-final languages like Japanese and German supports speakers' intuitions that they process sentences incrementally, without waiting for crucial information from sentence-final verbs. However, what has been less clear is the quality of the representations that speakers construct in advance of the verb, whether these representations involve compositional interpretation of preverbal constituents, and whether they incorporate the amount of structure that is needed to capture detailed acceptability judgment patterns in off-line tasks. As a step toward answering this question, in this article we presented findings from two self-paced reading studies and four off-line studies on backwards anaphoric dependencies in Japanese. The findings indicate that the parser incrementally constructs anaphoric relations among NPs in advance of the verb, and more importantly that it does so selectively, only in configurations where anaphora is judged to be more acceptable in off-line tasks. These results imply that, at least in the domain of anaphora, head-final word order does not undermine a speaker's ability to deploy his grammatical knowledge immediately in real-time language use.

## References

- Aoshima, S. 2003. The grammar and parsing of *wh*-dependencies. Ph.D. dissertation, University of Maryland, College Park.
- Aoshima, S., C. Phillips & A. Weinberg. 2004. Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51: 23–54.
- Bach, E. & B. Partee. 1980. Anaphora and semantic structure. In *CLS 16: Papers from the Parasession on Pronouns and Anaphora*, ed. J. Kreiman & A. E. Ojeda, 1–28. Chicago: The University of Chicago Press.
- Badecker, W. & K. Straub. 2002. The processing role of structural constraints on the interpretation of pronouns and anaphora. *Journal of Experimental Psychology: Learning Memory and Cognition* 28: 748–769.
- Bader, M. 1994. *Syntactische Verarbeitung im Deutschen* [Syntactic processing in German]. Doctoral dissertation, University of Stuttgart, Germany.
- Bader, M. & L. Lasser. 1994. German verb-final clauses and sentence processing: Evidence for immediate attachment. In *Perspectives on sentence processing*, ed. C. Clifton, Jr., L. Frazier & K. Rayner, 225–242. Hillsdale, NJ: Lawrence Erlbaum.

- Boland, J. E., M. K. Tanenhaus, S. M. Garnsey & G. N. Carlson. 1995. Verb argument structure in parsing and interpretation: Evidence from *wh*-questions. *Journal of Memory and Language* 34: 774–806.
- Bresnan, J. 1978. A realistic transformation grammar. In *Linguistic theory and psychological reality*, ed. J. Bresnan, 1–59. Cambridge, MA: MIT Press.
- Chafe, W. L. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and topic*, ed. C. N. Li, 25–55. New York: Academic Press.
- Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Clifton, C., S. M. Kennison & J. E. Albrecht. 1997. Reading the words *him* and *her*: Implications for parsing principles based on frequency and on structure. *Journal of Memory and Language* 36: 276–292.
- Crain, S. & J. D. Fodor. 1985. How can grammars help parsers? In *Natural language parsing: Psycholinguistic, computational, and theoretical perspectives*, ed. D. Dowty, L. Karttunen & A. M. Zwicky, 94–128. Cambridge: Cambridge University Press.
- Crawley, R. & R. Stevenson. 1990. Reference in single sentences and in texts. *Journal of Psycholinguistic Research* 19: 191–210.
- Crocker, M. W. 1996. *Computational psycholinguistics: an interdisciplinary approach to the study of language*. Dordrecht: Kluwer.
- Fanselow, G., R. Kliegl & M. Schlesewsky. 1999. Processing difficulty and principles of grammar. In *Constraints on language, aging, grammar, and memory*, ed. S. Kemper & R. Kliegl, 171–201. Dordrecht: Kluwer.
- Ferreira, F. 2003. The misinterpretation of non-canonical sentences. *Cognitive Psychology* 47: 164–203.
- Fodor, J. D. & Y. Hirose. 2003. What Japanese parsing tells us about parsing. In *Japanese Korean Linguistics 12*, ed. W. McClure, 193–205. Stanford, CA: CSLI Publications.
- Fodor, J. D. & A. Inoue. 1994. The diagnosis and cure of garden paths. *Journal of Psycholinguistic Research* 23: 407–434.
- Fodor, J. D. & A. Inoue. 1998. Attach anyway. In *Reanalysis in sentence processing*, ed. J. D. Fodor & F. Ferreira, 101–141. Dordrecht: Kluwer.
- Ford, M., J. Bresnan & R. Kaplan. 1982. A competence-based theory of syntactic closure. In *The mental representation of grammatical relations*, ed. J. Bresnan, 727–796. Cambridge, MA: MIT Press.
- Frazier, L. & C. Clifton. 2000. On bound variable interpretations: The LF-Only hypothesis. *Journal of Psycholinguistic Research* 29: 125–139.
- Friederici, A. D., E. Pfeifer & A. Hahne. 1993. Event-related brain potentials during natural speech processing: Effects of semantic, morphological, and syntactic violations. *Cognitive Brain Research* 1: 183–192.
- Garnsey, S. M., M. K. Tanenhaus & R. M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18: 51–60.
- van Gompel, R. P. G. & S. P. Liversedge. 2003. The influence of morphological structure on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29: 128–139.
- Harada, K. I. 1972. Constraints on *wh*-*q* binding. In *Studies in descriptive and applied linguistics 5*, 180–206. Tokyo: International Christian University.
- Harada, S.-I. 1977. Nihongo ni henkei wa hitsuyou da [Transformations are needed in Japanese]. *Gengo* 6:(10)88–95, (11)96–103.
- Hoji, H. 1985. Logical Form constraints and configurational structures in Japanese. Ph.D. dissertation, University of Washington, Seattle.
- Hoji, H. 1991. Kare. In *Interdisciplinary approaches to language: Essays in honors of S.-Y. Kuroda*, ed. C. Georgopoulos & R. Ishihara, 287–304. Dordrecht: Kluwer.

- Hoji, H. 1995. Demonstrative binding and principle B. In *Proceedings of NELS 25*, ed. J. N. Beckman, 255–271. Amherst, MA: GLSA Publications.
- Hoji, H. 2002. Surface and deep anaphora, sloppy identity, and experiments in syntax. In *Anaphora: A reference guide*, ed. A. Barss, 172–236. Oxford: Blackwell.
- Hoji, H., S. Kinsui, Y. Takubo & A. Ueyama. 2000. Demonstratives, bound variables, and reconstruction effects. *Proceedings of the Nanzan GLOW: The Second GLOW Meeting in Asia*, 295–329. Nagoya, Japan: Nanzan University.
- Inoue, A. 1991. A comparative study of parsing in English and Japanese. Ph.D. dissertation, University of Connecticut, Storrs.
- Inoue, A. & J. D. Fodor. 1995. Information-paced parsing of Japanese. In *Japanese sentence processing*, ed. R. Mazuka & N. Nagai, 9–63. Hillsdale, NJ: Lawrence Erlbaum.
- Johnson-Laird, P. N. 1983. *Mental models: Towards a cognitive science of language, inference, and consciousness*. Cambridge, MA: Harvard University Press.
- Just, M. A., P. Carpenter & J. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 11: 228–238.
- Kaiser, E. & J. Trueswell. In press. Investigating the interpretation of pronouns and demonstratives in Finnish: Going beyond salience. In *The processing and acquisition of reference*, ed. E. Gibson & N. Pearlmuter. Cambridge, MA: MIT Press.
- Kamide, Y., G. T. M. Altmann & S. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language* 49: 133–156.
- Kamide, Y. & D. C. Mitchell. 1999. Incremental pre-head attachment in Japanese parsing. *Language and Cognitive Processes* 14: 631–662.
- Kazanina, N. 2005. The acquisition and processing of backwards anaphora. Ph.D. dissertation, University of Maryland, College Park.
- Kazanina, N., E. Lau, M. Lieberman, C. Phillips & M. Yoshida. 2007. The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language* 56: 384–409.
- Kitagawa, C. 1981. Anaphora in Japanese: *Kare* and *zibun*. In *Coyote Papers: Working papers in linguistics from A to Z 2: Proceedings of the Arizona Conference on Japanese Linguistics, the Formal Grammar Sessions*, ed. A. K. Farmer & C. Kitagawa, 61–75. Tucson: University of Arizona.
- Kempen, G. & K. Harbusch. 2002. Performance grammar: a declarative definition. In *Computational linguistics in the Netherlands 2001*, ed. A. Nijholt, M. Theune & H. Hondorp, 148–162. Amsterdam: Rodopi.
- Kempson, R., W. Meyer-Viol & D. Gabbay. 2001. *Dynamic syntax*. Oxford: Blackwell.
- Kennison, S. M. 2003. Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language* 49: 335–352.
- Kuno, S. 1986. Anaphora in Japanese. In *Papers from the first SDF workshop in Japanese syntax*, ed. S.-Y. Kuroda, 11–70. San Diego: University of California.
- Kuroda, S.-Y. 1965. Generative grammatical studies in the Japanese language. Ph.D. dissertation, MIT.
- Kuroda, S.-Y. 1979. (Ko) so, a nitute [On (ko), so, and a]. In *Eigo to Nihongo to [English and Japanese]*, 41–59. Tokyo: Kuroshio Syuppan.
- Kutas, M. & S. A. Hillyard. 1980. Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science* 207: 203–204.
- Marslen-Wilson, W. D. 1973. Linguistic structure and speech shadowing at very short latencies. *Nature* 244: 522–523.
- Marslen-Wilson, W. D. 1975. Sentence perception as an interactive parallel process. *Science* 189: 226–228.
- Mazuka, R. & K. Itoh. 1995. Can Japanese speakers be lead down the garden path? In *Japanese sentence processing*, ed. R. Mazuka & N. Nagai, 295–329. Hillsdale, NJ: Lawrence Erlbaum.

- Mazuka, R., K. Itoh & T. Kondo. 2002. Cost of scrambling in Japanese sentence processing. In *Sentence processing in East Asian languages*, ed. M. Nakayama, 131–166. Stanford, CA: CSLI Publications.
- McElree, B. & T. Griffith. 1995. Syntactic and thematic processing in sentence comprehension: evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 21: 134–157.
- Miyagawa, M. 1989. *Structure and case-marking in Japanese*. New York: Academic Press.
- Miyamoto, E. T. & S. Takahashi. 2000. The processing of *wh*-phrases and interrogative complementizers in Japanese. In *Japanese Korean linguistics 10*, ed. N. Akatsuka & S. Strauss, 62–75. Stanford, CA: CSLI Publications.
- Miyamoto, E. T. & S. Takahashi. 2002. Sources of difficulty in the processing of scrambling in Japanese. In *Sentence processing in East Asian languages*, ed. M. Nakayama, 167–188. Stanford, CA: CSLI Publications.
- Miyamoto, E. T. & S. Takahashi. 2003. Typing mismatch effects in the processing of *wh*-phrases in Japanese. Ms., University of Tsukuba and MIT.
- Miyara, S. 1982. Reordering in Japanese. *Linguistic Analysis* 9: 307–340.
- Mulders, I. C. M. C. 2002. Transparent parsing: head-driven processing of verb-final structures. Doctoral dissertation, Utrecht: LOT.
- Nakatani, K. & E. Gibson. 2003. An on-line study of Japanese nesting complexity. Poster presented at *The 16th annual CUNY Conference on Human Sentence Processing*, MIT.
- Nakayama, S. 1982. On English and Japanese pronouns. Master's thesis, University of Tokyo, Tokyo.
- Nagata, H. 1993. Unimmediate construction of syntactic structure for garden path sentences in Japanese. *Journal of Psycholinguistic Research* 22: 365–381.
- Neville, H. J., J. Nicol, A. Barss, K. Forster & M. Garrett. 1991. Syntactically-based sentence processing classes: Evidence from event-related brain potentials. *Journal of Cognitive Neuroscience* 3: 151–165.
- Nicol, J. L. & D. Swinney. 1989. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18: 5–19.
- Noguchi, T. 1997. Two types of pronouns and variable binding. *Language* 73: 770–797.
- Phillips, C. 1996. Order and structure. Doctoral dissertation, MIT.
- Phillips, C. 2003. Linear order and constituency. *Linguistic Inquiry* 34: 37–90.
- Phillips, C. 2006. The real-time status of island phenomena. *Language* 82: 795–823.
- Phillips, C. & M. Wagers. 2007. Relating structure and time in linguistics and psycholinguistics. In *The Oxford handbook of psycholinguistics*, ed. G. Gaskell, 739–756. Oxford: Oxford University Press.
- Pickering, M. J. & M. J. Traxler. 2001. Strategies for processing unbounded dependencies: Lexical information and verb-argument structure assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27: 1401–1410.
- Pollard, C. & I. Sag. 1994. *Head-driven phrase structure grammar*. Stanford, CA: CSLI Publications.
- Pritchett, B. L. 1988. Garden path phenomena and the grammatical basis of language processing. *Language* 64: 539–576.
- Pritchett, B. L. 1991. Head position and parsing ambiguity. *Journal of Psycholinguistic Research* 20: 251–270.
- Pritchett, B. L. 1992. *Grammatical competence and parsing performance*. Chicago: The University of Chicago Press.
- Reinhart, T. 1983. *Anaphora and semantic interpretation*. Chicago: The University of Chicago Press.

- Resnik, P. 1992. Left-corner parsing and psychological plausibility. In *Proceedings of the 15th International Conference on Computational Linguistics (COLING'92)*, 191–197. Nantes, France.
- Saito, M. 1983. Comments on the papers on generative syntax. In *Studies in generative grammar and language acquisition: A report on recent trends in linguistics*, ed. Y. Otsu, H. van Riemsdijk, K. Inoue, A. Kamio & N. Kawasaki, 79–89. Tokyo: International Christian University.
- Saito, M. 1985. Some asymmetries in Japanese and their theoretical implications. Ph.D. dissertation, MIT.
- Saito, M. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*, ed. M. R. Baltin & A. S. Kroch, 182–200. Chicago: The University of Chicago Press.
- Saito, M. 1992. Long-distance scrambling in Japanese. *Journal of East Asian Linguistics* 1: 69–118.
- Saito, M. & H. Hoji. 1983. Weak crossover and move  $\alpha$  in Japanese. *Natural Language & Linguistic Theory* 1: 245–259.
- Schneider, D. 1999. Parsing and incrementality. Ph.D. dissertation, University of Delaware.
- Shapiro, L. & A. Hestvik. 1995. On-line comprehension of VP-ellipsis: Syntactic reconstruction and semantic influence. *Journal of Psycholinguistic Research* 24: 517–532.
- Shapiro, L., A. Hestvik, L. Lesan & A.R. Garcia. 2003. Charting the time-course of VP-ellipsis sentence comprehension: Evidence for an initial and independent structural analysis. *Journal of Memory and Language* 49: 1–19.
- Stabler, E. P. 1994. The finite connectivity of linguistic structure. In *Perspectives on sentence processing*, ed. C. Clifton, Jr, L. Frazier & K. Rayner, 303–336. Hillsdale, NJ: Lawrence Erlbaum.
- Steedman, M. 1989. Grammar, interpretation, and processing from the lexicon. In *Lexical representation and process*, ed. W. Marslen-Wilson, 463–504. Cambridge, MA: MIT Press.
- Steedman, M. 1997. *Surface structure and interpretation*. Cambridge, MA: MIT Press.
- Steedman, M. 2000. *The syntactic process*. Cambridge, MA: MIT Press.
- Stowe, L. A. 1986. Evidence for on-line gap-location. *Language and Cognitive Processes* 1: 227–245.
- Sturt, P. 1997. Syntactic reanalysis in human language processing, Ph.D. dissertation, University of Edinburgh.
- Sturt, P. & M. W. Crocker. 1996. Monotonic syntactic processing: A cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes* 11: 449–494.
- Sturt, P. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48: 542–562.
- Sussman, R. S. & J. C. Sedivy. 2003. The time-course of processing syntactic dependencies: Evidence from eye-movements during spoken *wh*-questions. *Language and Cognitive Processes* 18: 143–163.
- Takezawa, K. 1987. A configurational approach to case-marking in Japanese. Ph.D. dissertation, University of Washington, Seattle.
- Takubo, Y. & S. Kinsui. 1998. On the relation between the deictic and the non-deictic use of the Japanese demonstratives. Paper presented at *the Symposium on Diachronic and Synchronic Studies on Syntax of East Asian Language*, University of Southern California, Los Angeles.
- Tanenhaus, M. K., M. J. Spivey-Knowlton, K. M. Eberhard & J. C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268: 1632–1634.

- Townsend, D. J. & T. G. Bever. 2001. *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.
- Traxler, M. J. & M. J. Pickering. 1996. Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35: 454–475.
- Trueswell, J. C., M. K. Tanenhaus & C. Kello. 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.
- Ueda, M. 1990. Japanese phrase structure and parameter setting. Ph.D. dissertation, University of Massachusetts, Amherst.
- Ueyama, A. 1998. Two types of dependency. Ph.D. dissertation, University of Southern California, Los Angeles.
- Ueyama, A. 2000. Nihongo kara mieru “bunpoo” no sugata. [Characters of grammar from the view of Japanese.]. *Nihongogaku [Japanese Linguistics]* 19: 169–181.
- Vosse, T. & G. Kempen. 2000. Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition* 75: 105–143.

*Sachiko Aoshima*

*sachiko.aoshima@gmail.com*

*Masaya Yoshida  
Northwestern University  
Department of Linguistics  
2016 Sheridan Road  
Evanston, IL 60208  
USA*

*m-yoshida@northwestern.edu*

*Colin Phillips  
University of Maryland  
Department of Linguistics  
1401 Marie Mount Hall  
College Park, MD 20742  
USA*

*colin@umd.edu*

## **Appendix: Experimental Materials**

Full sets of experimental materials for all experiments in this article can be accessed from the third author’s website at <http://www.ling.umd.edu/colin/>.