Locality and Word Order in Active Dependency Formation in Bangla

Dustin Alfonso Chacón1,4*, Mashrur Imtiaz2, Shirsho Dasgupta3, Sikder Monoare Murshed2, Mina Dan3, Colin Phillips1

1Department of Linguistics, University of Maryland, College Park, MD, USA
2Department of Linguistics, University of Dhaka, Dhaka, Bangladesh
3Department of Linguistics, University of Calcutta, Kolkata, India
4Department of Linguistics, New York University, New York, NY, USA

* Correspondence: Dustin Alfonso Chacón, Department of Linguistics, New York University, New York, NY, USA
dustin.alfonso@nyu.edu

Keywords: filler-gap dependencies, locality, Bangla, sentence processing, islands.

Abstract

Research on filler-gap dependencies has revealed that there are constraints on both possible and grammatical gap sites, and that real-time sentence processing is sensitive to these constraints. Additionally, much of the evidence used to establish the real-time construction of filler-gap dependencies relies on indirect cues, such as measuring sensitivity to a disrupted resolution site. However, neither the mechanisms that select preferred gap sites nor the mechanisms used to detect whether these preferences are met are well understood. In this paper, we report on three experiments in Bangla, a language in which gaps may occur in either a preverbal embedded clause or a postverbal embedded clause. This word order variation allows us to manipulate whether the first gap linearly available is contained in the same clause as the filler, which allows us to dissociate structural locality from linear locality. In Experiment 1, an untimed ambiguity resolution task, we show that there is a global bias to resolve a filler-gap dependency with the first gap linearly available, regardless of structural hierarchy. In Experiments 2 and 3, which use the filled-gap paradigm, we find sensitivity to disruption only when the blocked gap site is both structurally and linearly local, i.e., the filler and the gap site are contained in the same clause. This suggests that the comprehender may not show sensitivity to the disruption of all preferred gap resolutions.

1 Introduction

The formation of linguistic dependencies is subject to a wide variety of constraints. Some constraints are conditions on grammatical well-formedness, whereas others define the interpretations that are preferred in real-time sentence processing. Locality constraints on filler-gap dependencies are one particularly well-studied example of both constraint types. Some locality constraints distinguish acceptable filler-gap dependencies from unacceptable filler-gap dependencies, as long recognized by syntacticians (Ross, 1967; Huang, 1982; Chomsky, 1986; Rizzi, 1990; Rizzi, 2013; Lasnik and Saito, 1992; Manzini, 1992; Szabolsci and den Dikken, 1999; Boeckx, 2008). For instance, the filler-gap dependency in (1a) between who and the position in which it is interpreted (marked as ___ ) is judged acceptable, in contrast with the sentence in (1b). This is because filler-gap dependencies may not cross into clauses (marked S') in the subject position of another clause (the sentential subject
Constraints on acceptable filler-gap dependencies are called island constraints.

(1) a. **Who** did it surprise Dale $[S'\text{ that Sarah saw } \_\_]$?
   b. * **Who** did $[S'\text{ that Sarah saw } \_\_]$ surprise Dale?

Other locality constraints determine which gap sites are preferred when multiple possibilities are available. In on-line tasks, this manifests as a preference for early resolution, a process called active dependency formation (Fodor, 1978; Crain and Fodor, 1985; Stowe, 1986; Frazier, 1987; Frazier and Flores d’Arcais, 1989). For instance, Stowe (1986) observed longer reading times at the direct object *us* in (2a) compared to the control sentence in (2b), which lacks a filler-gap dependency. This increase in reading times, called the filled-gap effect, suggests that readers make an early commitment to resolve *who* as the direct object of *bring* before it is clear whether there is a direct object gap. Encountering the overt direct object pronoun *us* then triggers a reanalysis process, leading to an increase in processing difficulty.

(2) a. My brother wanted to know **who** Ruth would bring *us* home to **_** at Christmas.
   b. My brother wanted to know if Ruth would bring us *home to **_** at Christmas.

The filled-gap effect illustrates a few key points. First, it demonstrates that comprehenders commit to an early resolution site, i.e., there are locality constraints that bias readers to postulate a gap in the object position of *bring* rather than the object position of the preposition *to*. Additionally, it demonstrates that comprehenders can quickly detect when this commitment is inappropriate, and that recovering from the inappropriate commitment is costly. This implies that the filled-gap effect diagnoses not only where gap sites are preferred, but also where gap sites are postulated strongly and early enough to induce a processing cost when the early commitments are not confirmed.

The preference to make an early commitment to a gap is commonly regarded as arising from constraints on resources, e.g., maintaining a filler in memory is costly and therefore it is preferable to resolve filler-gap dependencies earlier (e.g., Wanner and Maratsos, 1978), or longer dependencies in general are more costly to represent and thus dispreferred (e.g., Gibson, 1998).

There has been much interest in determining whether these two types of constraints are the same, following from some independently motivated restrictions on linguistic processes, e.g. restrictions on memory capacity (Deane, 1991; Pritchett, 1992; Klüender and Kutas, 1993; Klüender, 1998; Klüender, 2004; Hofmeister and Sag, 2010; for discussion see Phillips, 2013a). Explaining island phenomena as a consequence of resource limitation has the potential to radically simplify grammatical theories.

If island constraints are indeed reducible to constraints on preferred gap sites, then both sets of constraints should be sensitive to the same properties of the linguistic representation being computed. In other words, the notion of “local” that is relevant should be the same. It is relatively uncontroversial that island constraints are defined in terms of formal linguistic structure, either hierarchical syntactic relations (Ross, 1967; Chomsky, 1981; Huang, 1982; Chomsky, 1986; Lasnik and Saito, 1992; Rizzi, 1990; for review, see Rizzi, 2013), or semantic/pragmatic relations (Erteschik-Shir, 1973; Kuno, 1976; Szabolcsi and Zwarts, 1993; Truswell, 2007; Ambridge and Goldberg, 2008; Abrusán, 2011a; Abrusán, 2011b). However, it is unclear what notion of locality is relevant for determining preferred gap sites. For instance, the direct object position of *bring* may be preferred because fewer nodes separate it from alternative gap sites (i.e., there is an additional PP node separating the filler and prepositional object gap site, illustrated in (3)), or this position may be
preferred because it is the first position that is linearly available. That is, the locality constraints on preferred gap sites may be defined in terms of structural locality or linear locality. If the constraints on preferred gap sites are sensitive to linear locality, then this motivates maintaining a distinction between island constraints and locality constraints on preferred gap sites.

(3) My brother wanted to know

who [s Ruth would [vp bring us home [pp to ___] at Christmas]]

Most research on filler-gap dependency processing cannot decide among these hypotheses, because most studies are conducted on languages like English, where structural and linear locality converge. However, results in Japanese, a language in which complement phrases precede their containing phrase are more illuminating, because they allow us to dissociate these two notions of locality. Findings in Japanese show that filler-gap dependencies are preferentially resolved with the first gap site that is linearly available (Aoshima et al., 2004; Yoshida, 2006; Omaki et al., 2014). This suggests that locality constraints determining preferred gap sites are sensitive to linear order rather than to hierarchical linguistic structure, unlike island constraints.

In this paper, we report on three experiments in Bangla (Bengali) that investigate locality constraints on preferred gap sites. Bangla is a valuable language for this purpose, because embedded clauses may either precede or follow the embedding verb, as shown in (4)–(5). Additionally, like English and Japanese, Bangla allows filler-gap dependencies with question phrases. These filler-gap dependencies may resolve in either the main clause, or an embedded clause on either side of the main verb, as shown in (6)–(7). This allows us to manipulate whether the first gap site is structurally local or distant within the same language, paralleling the cross-language comparison between English and Japanese by Omaki et al. (2014).

(4) raj bollo [š s̪ e ašbe ]
   Raj said he come.FUT
   ‘Raj said that he will come.’

(5) raj [š s̪ e ašbe ] bollo
   Raj he come.FUT said
   ‘Raj said that he will come.’

(6) raj kākhon ___ bollo [š s̪ e ___ ašbe ]
    raj when said he come.FUT
    ‘When did Raj say ___ that he will come ___?’

(7) raj kākhon [š s̪ e ___ ašbe ] ___ bollo
    raj when he come.FUT said
    ‘When did Raj say ___ that he will come ___?’

Experiment 1 was a within-language replication of the cross-language experiment in Omaki et al. (2014). In Experiment 1, we investigated how ambiguous filler-gap dependencies like (6) and (7) are resolved using an off-line ambiguity resolution task. We found that filler-gap dependencies are resolved with the first position linearly available across word orders. In main verb first word orders as in (6), the filler-gap dependency was resolved with the main verb. In embedded verb first word orders like (7), it was resolved with the embedded verb. This provides within-language support for the generalization that locality constraints on preferred gap sites are sensitive to linear locality, and
not structural locality. This motivates separating island constraints from the constraints on preferred gap sites, since they are defined over different linguistic properties.

In Experiment 2, we investigated the preference for linearly local gap sites in an on-line, filled-gap paradigm task. Like Experiment 1, we leveraged the flexible word order of Bangla to manipulate whether the first available gap site was in the same clause as the filler or in an embedded clause. We find a filled-gap effect when resolution with the first gap site is blocked in main verb first word orders like (6), where structural locality and linear locality align, but not in embedded verb first word orders like (7). In other words, there is only sensitivity to disruption of a linearly local gap when this position is structurally local, but not when this position is structurally distant. The comparison between Experiments 1 and 2 suggests a contrast between gap site preferences and sensitivity to disruption.

The apparent mismatch in Experiments 1 and 2 may be due to the on-line/off-line contrast between the two experiments, or to the ambiguity resolution/filled-gap paradigm difference. In Experiment 3, we diagnosed the cause of this mismatch. Experiment 3 was an off-line acceptability judgment task, like Experiment 1, using the filled-gap paradigm, like Experiment 2. We again only find sensitivity to disruption when a filler-gap dependency is blocked from resolving with the first position linearly available in main verb first word orders, as in Experiment 2. This suggests that the contrast between locality preferences and sensitivity to disruption for embedded verb first word orders in Experiments 1 and 2 is not due to the off-line/on-line contrast, but rather the specific mechanisms underlying filled-gap detection.

2 Locality in Filler-Gap Dependencies

There is substantial evidence that shorter filler-gap dependencies are preferred to longer filler-gap dependencies. This is commonly understood to reflect an additional processing cost associated with longer dependencies. For instance, Frazier and Clifton (1989) found that reading times were increased for sentences containing filler-gap dependencies spanning multiple clauses compared to controls (see also Kluender and Kutas, 1993; Kluender, 1998; Dickey, 1996). This bias against longer filler-gap dependencies is also reflected in offline acceptability judgments, where sentences containing filler-gap dependencies spanning multiple clauses are rated lower than sentences with shorter filler-gap dependencies (Phillips et al., 2005; Alexopoulou and Keller, 2007; Sprouse et al., 2012).

Online studies show that the preference for shorter filler-gap dependencies manifests as a preference for early resolution. For instance, the filled-gap effect discussed above shows that blocking early filler-gap dependency resolution triggers a costly reanalysis process (Crain and Fodor, 1985; Stowe, 1986; Lee, 2004), if the comprehender can detect that his or her early commitment is inconsistent with the rest of the sentence. Converging evidence comes from the plausibility mismatch paradigm (Garnsey et al., 1989; Traxler and Pickering, 1996). For instance, in a series of eye-tracking experiments, Traxler and Pickering (1996) observed that gaze times increase on the verb wrote in (8b) compared to (8a).

(8) 
  a. We like the book that the author wrote unceasingly and with great dedication about ___ while waiting for a contract.
  b. We like the city that the author wrote unceasingly and with great dedication about ___ while waiting for a contract.
This suggests that the city is as first interpreted as the object of wrote. Comprehenders can then detect that the early gap commitment yields an implausible interpretation. Then, they reject this commitment, and search for a different gap, yielding a reanalysis cost. Thus, like we argued for the filled-gap effect, the plausibility mismatch effect illustrates not only early commitment to a local gap, but also sensitivity to disruption when this position is unavailable. Other converging evidence for active dependency formation comes from EEG studies (Garnsey et al., 1989; Kaan et al., 2000; Phillips et al., 2005), the “stops making sense” task (Tanenhaus et al., 1985; Boland et al., 1995), cross-modal lexical priming (Nicol and Swinney, 1989; Nicol et al., 1994), and ‘visual world’ eye-tracking (Sussman and Sedivy, 2003).

This bias towards early filler-gap dependency resolution in real-time behavior and towards shorter dependencies in offline judgments is commonly attributed to resource limitations. For instance, unintegrated fillers may require memory resources to be actively maintained (Jackendoff and Culicover, 1971; Wanner and Maratsos, 1977). Alternatively, longer dependencies in general may be more costly, leading to a dispreference for longer filler-gap dependencies (Gibson, 1998; Hawkins, 2005). Other analyses contend that longer filler-gap dependencies may cause increased processing difficulty because the filler must be retrieved from memory at the gap site, which may be costly and error-prone in the case of longer dependencies (McElree, 2006; Wagers, 2012). Lastly, more local gaps may be preferred because comprehenders attempt to resolve as many grammatical requirements as early as possible (Pritchett, 1992; Weinberg, 1992; Altmann and Kamide, 1999; Sedivy et al., 1999; Aoshima et al., 2004; Wagers and Phillips, 2009). These accounts all imply that the comprehender should minimize filler-gap dependency length in order to optimize resource consumption. However, these accounts make no commitment as to whether linear span or structural span are relevant in selecting preferred gap sites.

Island constraints, in contrast, are typically described in structural terms. Island constraints are restrictions on possible filler-gap dependencies, with several illustrated in (9).

(9)  a. Relative Clause Island:
    * Who did Dale comfort [NP the woman that [s saw ___ ?]]
 b. Whether Island:
    * Who did Dale wonder [whether Bob frightened ___ ?]
 c. Wh-Island:
    * Who did Dale say [who saw ___ behind Laura’s bed?]?
 d. Subject Island:
    * Who did [the fact that Sarah saw ___ ] surprise Dale?
 e. Adjunct Island:
    * Who did Dale ruminate [while Harry interrogated ___ ?]
 f. Coordinate Structure Constraint:
    * Who did [Dale suspect ___ and Harry interrogate Leland?]
 g. Factive Island:
    * Why did Dale remember [that Ben was suspicious ___ ?]

Island constraints have long been studied in theoretical linguistics, where they typically are characterized as constraints on well-formed linguistic representations, either as formal syntactic constraints (Ross, 1967; Chomsky, 1973; Chomsky, 1981; Huang, 1982; Chomsky, 1986; Lasnik and Saito, 1992; Rizzi, 1990; Rizzi, 2013), or as constraints on well-formed and felicitous semantic/pragmatic forms (Erteschik-Shir, 1973; Kuno, 1976; Szabolcsi and Zwarts, 1993; Truswell, 2007; Ambridge and Goldberg, 2008; Abrusán, 2011a; Abrusán, 2011b). As such, island constraints
are typically defined over the hierarchical structure of the sentence, or the formal relations between the words and phrases. This can be demonstrated with pairs like (10), repeated from (1), in which the filler-gap dependency that spans fewer words is dispreferred to a filler-gap dependency that spans more words. This contrast follows if there is a constraint against gaps in subject clauses, but not extraposed clauses, for instance (Ross, 1967).

(10) a. **Who** did it surprise Dale [\textit{s} that Sarah saw ___]?
    b. * **Who** did [\textit{s} that Sarah saw ___] surprise Dale?

Island constraints are observed to be robust in both off-line and on-line measures. Off-line acceptability judgments show that speakers give low ratings to sentences with island violations (Sobin, 1987; Alexopoulou and Keller, 2007; Cowart, 1996; Cowart, 2003; Heestand et al., 2011; Sprouse et al., 2012). Additionally, the effects of active dependency formation typically disappear in island constructions. There are no filled-gap effects or plausibility mismatch effects inside island contexts (Stowe, 1986; Bourdages, 1992; Traxler and Pickering, 1996). Similarly, results from EEG studies (Neville et al., 1991; Kutas and Klünder, 1993; McKinnon and Osterhout, 1996) and speed-accuracy tradeoff studies (McElree and Griffith, 1998) suggest that comprehenders immediately detect island boundaries. The rapid application of island constraints can be explained in theories of sentence processing that posit rapid and faithful use of grammatical constraints (e.g., Lewis and Phillips, 2015) or theories that posit that representations with gap sites inside island contexts are too costly to represent (Gibson, 1998; Hawkins, 2005).

Some data suggests the constraints on preferred gaps should be dissociated from island constraints (Phillips, 2006; Wagers and Phillips, 2009; Sprouse et al., 2012; Yoshida et al., 2014). For instance, whereas island constraints are defined in terms of hierarchical structure, constraints on preferred gap sites are defined in terms of linear locality. This can be seen in studies on Japanese, a language that can dissociate structural locality and linear locality. Japanese speakers prefer to resolve filler-gap dependencies in embedded clauses, likely because this is the first position linearly available. For instance, Aoshima et al. (2004) found filled-gap effects for sentences like (11), in which the fronted dative phrase \textit{dono-syain-ni} ‘which employee- DAT’ is blocked from resolving with the embedded clause because of the case-matched noun phrase \textit{kacyoo-ni} ‘assistant manager- DAT’ (see also Yoshida, 2006). Similarly, Omaki et al. (2014) showed that speakers of Japanese interpret an ambiguously fronted \textit{wh}-phrase, as in (12), with the embedded clause in a Question after Story task, a task that provides an untimed measure of how speakers prefer to interpret ambiguous questions (de Villiers et al., 1990). This shows that in off-line measures of gap locality preferences and on-line measures of sensitivity to disruption, Japanese speakers prefer a linearly local resolution.

(11) \textbf{Dono-syain-ni} senmu-wa [syacyoo-ga kaigi-de
    which employee-DAT managing director-TOP president-NOM meeting-at

    kacyoo-ni syookyu-o yakuoku-sita-to] iimasita-ka?
    assistant manager-DAT raise-ACC promised-DECLC told-Q?
    ‘Which employee did the managing director tell ___ that the president promised a raise to
    the assistant manager at the meeting?’

(12) \textbf{Doko-de} Yukiko-chan-wa [choucho-o tsukumaeru-to ] itteta-no?
    where-at Yukiko-DIM-TOP butterfly-ACC catch-DECLC was telling-Q?
    ‘Where did Yukiko say that she will catch butterflies?’
In this paper, we further investigate this generalization in Bangla, a language with variable word order that permits us to manipulate whether the most linearly local potential gap site is within the same clause as the fronted filler (i.e., structurally local), or in an embedded clause (i.e., structurally non-local). In Section 3, we describe the relevant properties of Bangla syntax. In Sections 4–7 we describe the results of three experiments on Bangla filler-gap dependency processing.

3 Grammatical Properties of Bangla

Bangla is a language spoken primarily in Bangladesh and the eastern Indian state of West Bengal, with approximately 180 million speakers worldwide (Lewis et al., 2015). Bangla is in the Eastern Zone of the Indo-Aryan branch of the Indo-European language family. Due to its contact with multiple linguistic areas, Bangla has many properties typical of northern Indo-Aryan, Dravidian, and Southeast Asian languages. For more complete descriptions of the language, see Thompson (2010) and David (2015).

Embedded clauses in Bangla may either precede or follow an embedding verb, shown in (13). Postverbal embedded clauses may be introduced with the complementizer je, shown in (14a). Preverbal embedded clauses may appear with the complementizer bole at the end of the clause, shown in (14b), or with je in a clause-internal position, shown in (14c). Dasgupta (2007) describes the clause-internal je as an “anchor”, which may be a distinct lexical category. Examples are taken from Bayer (1996).

(13) a. še bollo ora ašbe
    he said they come.FUT

   b. še ora ašbe bollo
      he they come.FUT said
      ‘He said that they will come’

(14) a. chele-ṭa bollo [š, je tar baba ašbe ]
     boy-CL said that his father come.FUT

   b. chele-ṭa [š, tar baba ašbe bole ] bollo
      boy-CL his father come.FUT that said

   c. chele-ṭa [š, tar baba je ašbe ] bollo
      boy-CL his father that come.FUT said
      ‘The boy said that his father will come’

These constructions are used in similar contexts, although there are subtle syntactic and semantic differences that we leave aside (for discussion see Bal, 1990 on related constructions in Oriya, and Bayer, 1996; Bayer, 1999; Bayer, 2001; Simpson and Bhattacharya, 2000; Simpson and Bhattacharya, 2003).

Like English and Japanese, Bangla also permits unbounded filler-gap dependencies. Gaps may either occur in preverbal or postverbal embedded clauses. Extraction from a postverbal clause is shown in (15), taken from Simpson and Bhattacharya (2003). In (15a), the noun phrase hæmleṭ ‘Hamlet’ is interpreted as the direct object of the verb poreche ‘read’. In (15b) and (15c), hæmleṭ ‘Hamlet’ appears either one or two clauses away from the embedded clause, but is still interpreted as the direct
object of *poreche* ‘read’.

(15) a. jon bhablo [s’ sery bollo [s’ su hæmlet poreche ]]  
    John thought Mary said Sue Hamlet read

    b. jon bhablo [s’ sery hæmlet bollo [s’ su ____ poreche ]]  
    John thought Mary Hamlet said Sue read

    c. jon hæmlet bhablo [s’ sery bollo [s’ su ____ poreche ]]  
    John Hamlet thought Mary said Sue read

    ‘John thought that Mary said that Sue has read Hamlet’

Extraction from preverbal clauses is shown in (16). In (16a), the noun phrase *tomar beřal-ke* ‘your cat-ACC’ is interpreted as the object of the embedded verb *kamreche* ‘bit’, but it appears in the left edge position of the main clause. Similarly, in (16b), the prepositional phrase *bas theke* ‘bus from’ appears in the left edge position of the main clause, but is interpreted as a modifier of the embedded clause. Compare this with other languages with both preverbal and postverbal clauses, like Basque which disallows gap sites in preverbal clauses (Uriagereka, 1992), and Malayalam which only allows direct object gaps in preverbal clauses, but not adjunct phrases like *bas theke* ‘bus from’ (Srikumar, 2007). In Bangla, the fronted phrase may appear in either the preverbal position, as in (15), or at the left-edge position, as in (16).

(16) a. **tomar beřal-ke** amra ṣobai  
    your cat-ACC we everyone

    [s’ paš-er bari-r kukur ____ kamreche bole ] ṣunechilam  
    neighbor-GEN dog bit that heard

    ‘We had all heard that the neighbor’s dog has bitten your cat’

    b. **bas theke** amar didi  
    bus from my sister

    [s’ ṭogulo duronto bacca laphiy e nambe bole ] bhabe ni  
    so many uncontrollable child jumping descend.FUT that didn’t think

    ‘My sister hasn’t thought that so many children could jump down from a bus.

To summarize, Bangla permits embedded clauses to precede or follow the embedding verb. Additionally, fillers in the main clause may resolve with gap sites in the main clause or in an embedded clause on either side of the embedding verb. This means the schematic representations in (17) are all permissible, making Bangla an excellent language for testing locality biases.

(17) a. Postverbal embedded clause, main clause resolution:  
    … filler … ____ … V … [s’ … ] …

    b. Postverbal embedded clause, embedded clause resolution:  
    … filler … V … [s’ … ____ … ] …

    c. Preverbal embedded clause, embedded clause resolution:  
    … filler … [s’ … ____ … ] … V …
d. Preverbal embedded clause, main clause resolution:

\[ \ldots \text{filler} \ldots [s \ldots ] \ldots \_\_\_ \ldots V \ldots \]

If the locality constraints on preferred gap sites are sensitive to linear order, as suggested by findings in Japanese, then the dependencies schematized in (17a) and (17c) should be preferred to those in (17b) and (17d). However, if locality constraints on preferred gap sites are sensitive to structural locality, then the representations in (17a) and (17d) should be preferred, since the filler and gap site are structurally more local to the filler. We test these predictions in Experiments 1–3.

4 Experiment 1

4.1 Rationale

In Experiment 1, we used the Question after Story task (de Villiers et al., 1990) to determine whether Bangla speakers prefer linearly local gap sites across word orders. We adapted the design used by Omaki et al. (2014), who probed for word order effects on filler-gap dependency resolution using a between language comparison. In their study, participants viewed a series of vignettes in which a character acted out an event in one location and reported on it in another location. Afterwards, participants were asked to respond to a question that contained a fronted \textit{wh}-filler that could resolve in either the embedded clause or main clause. Participants’ responses revealed in which clause they preferred to resolve the filler-gap dependency. In English, which conflates linear and structural locality, the ambiguous filler-gap dependency was most commonly resolved with the main clause in Omaki and colleagues’ studies. Conversely, in Japanese, the filler-gap dependency was preferentially resolved in the embedded clause. They took this as evidence for a universal preference to resolve filler-gap dependencies with the first position linearly available.

Our study took advantage of the flexible word order in Bangla to further test this hypothesis. The study had two main conditions: a main verb first condition, shown in (18a), and an embedded verb first condition, shown in (18b). For both sentences, the fronted \textit{wh}-filler \textit{kothae} ‘where’ can be resolved in the embedded clause, modifying the catching event, or the main clause, modifying the telling event. If gaps are preferentially constructed in the first position linearly available, as suggested by Omaki and colleagues’ cross-language contrast, then we expect \textit{kothae} ‘where’ to be resolved with the main verb in word orders like (18a), and with the embedded verb in word orders like (18b).

(18) a. **Main Verb First Condition:**

\[ \text{şumi kothae ækjôn-ke boleche [s\_je \_še prɔjapoti dhorbe]?} \]

Shumi where someone-ACC told that she butterfly catch.FUT

b. **Embedded Verb First Condition:**

\[ \text{şumi kothae [s\_že prɔjapoti dhorbe bole] ækjôn-ke boleche?} \]

Shumi where she butterfly catch.FUT that someone-ACC told

‘Where did Shumi tell someone that she will catch butterflies?’

4.2 Materials

The materials were adapted from Omaki et al. (2014). The stories and audio were translated by three of the authors to standard colloquial Dhakaiya Bangla. Some lexical material was changed to better suit the different cultural context, including names. The questions were presented in a paper questionnaire, which participants were instructed to fill out when the video was paused at the end of
each story. There was one question for each vignette. Across all questionnaires, we rigidly alternated between a target item and a filler item, in order to reduce priming or perseveration effects. The target items were two-clause sentences with an ambiguous *wh*-dependency, i.e., like (18). The fillers were one-clause sentences with an unambiguous *kæno* ‘why’ question.

Participants were split into two groups – the “between participants” group and the “within participants” group. The “between participants” group was included to make a closer comparison to the existing literature comparing English and Japanese. The division of participants is illustrated in Table 1. Questionnaires were prepared for each group. For the “between participants” questionnaires, the target items all had either main verb first word orders or embedded verb first word orders. The remaining participants received a “within participants” questionnaire, where the target items contained both verb first word order and embedded verb first word orders. These also alternated, such that there were two questions of each word order in each questionnaire.

The stories were animated vignettes made from clipart images. In each vignette, a character goes to four different locations, and performs an action in each. A sample story from Omaki *et al.*’s (2014) English study is given in (20). The videos are included as supplementary material.

(19) **[Introduction]**
It was a beautiful day in spring so Lizzie decided she was going to go catch butterflies in the park.

**[1st Location]**
Her Mom and Dad weren’t home, so Lizzie thought she should tell her brother or sister about going to the park, so that Mom and Dad would know where she was when they got back. She first went to her brother’s room, but he was taking a nap and she couldn’t tell him about catching butterflies.

**[2nd Location]**
Instead, Lizzie looked for her sister. She looked all over the house but didn’t see her sister anywhere! When she was about to give up, Lizzie heard her sister’s voice in the basement! She went to the basement and said to her sister: “I’m gonna catch butterflies in the park!”

**[3rd Location]**
Then, on her way to the park, Lizzie passed by a parking lot and saw a butterfly near it. She walked slowly towards the butterfly, but before Lizzie could get there, another girl came along and caught the butterfly! Lizzie didn’t see any more butterflies there, so she kept walking towards the park.

**[4th Location]**
There were lots and lots of butterflies in the park, and she caught one in a jar and took it home with her. She liked the one that she caught, but she wished she could have caught more butterflies.

Each vignette consisted of six phases. The first phase introduced the protagonist, displayed in the center of the screen. The following four phases depicted him or her at each of the four locations. The protagonist succeeded or failed to perform some intended action as announced in the introductory phase, or succeeded or failed to report on it. The contrast between successes and failures was intended to make the event-location pairings more memorable, and to ensure that the “where” test questions were felicitous. In locations where the protagonist succeeded on performing his or her stated action or reported on it, there was a visual trace left behind (i.e., a butterfly in a bottle, or a word balloon). The first two and last two locations were relevant for either the main clause event
(i.e., the reporting event), or the embedded clause event (i.e., the intended action). In the sixth and final phase, the protagonist returned to the center of the screen, and then the story concluded.

To avoid any potential recency bias, the ordering of the events within each story was counterbalanced, such that the first pair of events pertain to the reporting event in half of the stories, and to the embedded clause event in the other half of the stories. In each case, the story provided motivation for continuing to the next series of events. For instance, in (20), the reporting events are motivated by the character’s need to tell her siblings where she was going. The pairings of quadrant position and event were randomized across stories so that participants could not predict which locations would correspond to which actions.

4.3 Materials

Experiment 1 was an adaptation of Omaki et al.’s (2014) question-after-story task (de Villiers et al. 1990). Participants were instructed in Bangla to watch a sequence of 8 vignettes. At the end of each vignette, the screen displayed “write your answer now” in Bangla. At this point, the experimenter paused the video and instructed the participant to read a question printed on a paper questionnaire. Participants were instructed to write a brief response. This was because in pilot studies, participants attempted to recapitulate large portions of the story in their response. After responding, the experimenter resumed the video, which progressed to the next vignette.

96 participants were recruited for Experiment 1. Forty-eight adult native speakers of Bangla were collected from the student population at The University of Dhaka in Dhaka, Bangladesh, and 48 participants were from the student population at Calcutta University in Kolkata, India. Bangladeshi participants were compensated 500 Bangladeshi Taka (BDT), and Indian participants were compensated 200 Indian Rupees (INR). This session took approximately 15 minutes. Experiment 1 was conducted after participants completed either Experiment 2 or another experiment unrelated to the current study. These populations were each split into two groups, a “within-subjects” and a “between-subjects” group, as discussed in section 4.3. We tested participants in both India and Bangladesh to probe for any potential influence of dialect difference, especially given that Indian Bangla speakers are likely to be competent in Hindi, which uses different wh-scope marking properties (e.g., Dayal, 1996; Manetta, 2012). Additionally, we included a within-subjects and between-subjects manipulation to check for any effect of self-priming in the experiment. This was important for comparing our within-language findings to results from previous between-language comparisons, where participants in each language, e.g., Japanese and English, saw only one of the word orders tested in Bangla.

4.4 Results

We coded each response as either a main clause response or an embedded clause response, depending on which location the participant named. Responses that either failed to answer the question or that provided both possible answers were excluded. The proportions of excluded observations are given in Table 2. After excluding these observations, participants responded with the main verb location in 81% of the main verb first word orders, but only 23% of the embedded verb first word orders.

Using the lmer package in R (Bates et al., 2015), we submitted the results to a logit mixed effects model with a bobyqa optimizer. The predicted variable was main clause response, coded as 1. For fixed effects, we included word order (main verb first or embedded verb first), location (Dhaka or Kolkata), and list type (within participants or between participants), with their interaction terms. For random effects, we included participant and items. We also included random slopes for word order
by participant and by item. The results of the model are given in Table 3. The p-values in Table 3 were generated using the lmerTest package (Kuznetsova et al., 2015). The mean proportion of main verb responses is given in Figure 1.

Only word order had a significant effect on the proportion of main clause responses. The effect was as predicted: for the main verb first word order, participants showed a strong bias to answer with main verb locations. With embedded verb first word orders, there was a strong bias to answer with embedded verb locations. There was no significant effect of city, implying that there were no systematic dialect differences detected in Experiment 1. Additionally, there was no significant effect of list type, i.e., participants typically responded with the event denoted by the first verb linearly available regardless of whether they saw lists with only one word order or lists with mixed word order. For the main verb first word order, participants responded with the location denoted by the main verb in 72% of the trials in the within-participants list, and 81% of the trials in the between-participants list. For the embedded verb first word order, participants responded with the location denoted by the main verb in 28% of the within-participants trials, and 19% of the between-participants trials. Thus, we replicated Omaki and colleagues’ cross-language findings in the between participant group, and showed a robust bias to resolve the filler-gap dependency with the first verb across word orders in the within participant group as well.

4.5 Discussion

In Experiment 1, we showed that Bangla speakers preferentially resolve a filler-gap dependency with the first position linearly available, regardless of whether this position is in the same clause as the filler or in a more deeply embedded clause. This suggests that the locality constraints determining preferred gap sites are primarily sensitive to linear distance, as previously shown in a between-language comparison by Omaki et al. (2014). Importantly, this contrasts with observations about island constraints, which appear to be defined in terms of hierarchical structure.

This within-language demonstration of sensitivity to linear order is also important, because it helps keep constant all other grammatical properties between the word order comparisons. The results found by Omaki and colleagues may be due to some other grammatical distinction between English and Japanese apart from word order. For instance, obligatory long-distance wh-dependencies as observed in English have different properties than the optional wh-dependencies observed in Japanese (“scrambling”, Saito, 1985; Mahajan, 1990), which might indirectly bias the filler-gap dependency resolution preferences in these languages. These concerns are less likely to impact the results of Experiment 1, particularly because the effect is robust in the within participant questionnaires.

An advantage of the Question after Story task in Experiment 1 is that it directly probes participants’ preferred resolution sites instead of measuring sensitivity to disruption, unlike the filled-gap effect. However, the Question after Story task does not reveal the time course of dependency formation. We cannot infer from these data that there is early commitment to the linearly first gap site. For this reason, in Experiment 2, we use a filled-gap paradigm in a self-paced reading task to probe for sensitivity to disruption of filler-gap dependency resolution across word orders.

5 Experiment 2

5.1 Rationale
In Experiment 1, we showed that comprehenders prefer to resolve filler-gap dependencies with the first verb linearly available. The goal of Experiment 2 was to test whether this follows from an early and confident commitment to this gap location. We used the filled-gap paradigm in a self-paced reading task (Crain and Fodor, 1985; Stowe, 1986), which probes for sensitivity to disruption of a preferred gap on-line. If commitment to the first gap site is made early and confidently enough across word orders, then we expected a filled-gap effect when filler-gap dependency resolution with the first verb was blocked, regardless of whether this occurred in the same clause as the filler or in an embedded clause.

5.2 Materials

We crossed the factors word order (main verb first or embedded verb first) and extraction type (argument or adjunct extraction). In all target items there was a long-distance wh-filler gap dependency. The critical conditions contained an argument wh-filler (ka-ke ‘who-ACC’) marked in the accusative case. This argument wh-filler was blocked from resolving with the linearly first gap position by a case-matching noun phrase occupying that position. This was the filled-gap region and the critical region. The adjunct extraction conditions (kókhon ‘when’, kothae ‘where’) were the control conditions, since the accusative-marked noun phrase did not block resolution of an adjunct wh-dependency in that clause.

In the main verb first conditions, the argument wh-filler was blocked from resolving as the indirect object of the verb boleche ‘said/told’. The wh-filler must then resolve as the direct object of the later, embedded verb. Conversely, in the embedded verb first conditions, the argument wh-filler was blocked from resolving as the direct object of the embedded verb. It must therefore resolve as the indirect object of the main verb boleche. In the adjunct extraction conditions, an extra pronoun take ‘him/her-ACC’ was introduced as the object of the embedded verb in main verb first conditions, and the verb boleche ‘said’ in the embedded verb first conditions. This was necessary to ensure that all verbs had all argument roles discharged. In all conditions, the fronted wh-phrase appeared on the left edge of its containing clause to maximize the distance between the wh-phrase and the filled-gap region. The adjunct wh-phrases were counterbalanced between kókhon ‘when’ and kothae ‘where’.

The subject of the main clause containing the wh-filler always denoted a referent of high status, and the pronoun in the most deeply embedded clause and its verb were morphologically marked with politeness agreement. This was done to minimize the complexity induced by any retrieval operations needed in each pronoun and verb region, by maximizing the distinguishability of the referents introduced in the sentence. All target conditions were embedded in an additional clause (raśad jiggaśa koreche... ‘Rashad asked...’). This was to ensure that participants could not predict the word order of the target items on the basis of the first few words. Table 4 gives a sample set of items. There were 32 sets of target items and 48 complexity-matched fillers. The sentences were presented in a Latin Square design, with order randomized for each participant.

5.3 Methods

Sentences were presented on a PC laptop using the Ibex software (http://www.spellout.net/ibexfarm) in a self-paced, word-by-word, moving window paradigm (Just et al., 1982). Ibex is intended for web-based tasks, but the task was run offline by one of the authors. Each trial began with a screen presenting a sentence in which the words were masked by dashes, with spaces intact. Each time the participant pressed the spacebar, a word was revealed and the previous word was again hidden behind a dash. A yes/no comprehension question appeared all at once after the participant completed each sentence. The participant was instructed to use the ‘f’ key to respond ‘yes’, and the ‘j’ key to respond ‘no’, with on-screen reminders of this key-response pairing. On-screen feedback informed
the participant whether the response was correct. Participants were instructed to read carefully at a natural but quick pace, and to answer the questions carefully. The order of presentation of responses was randomized for each participant. All instructions and feedback were given in Bangla.

Participants were 32 adult native speakers of Bangla from the University of Dhaka student community. Due to a technical error, 3 participants’ responses were not recorded, and thus we report on 29 participants. They were compensated 500 BDT for their time. The task took approximately 20–30 minutes to complete.

5.4 Results

Analyses were conducted on comprehension task accuracy and reading times. Trials that received incorrect responses in the comprehension task were removed from analysis. 4 participants whose mean accuracy fell below 70% were removed from analysis. The mean accuracy on the comprehension questions was 80.6% after removing these 4 participants.

Using the lme4 package in R (Bates et al., 2015), we analyzed the reading times for the filled-gap region and the subsequent regions using linear mixed effects models for each word order. We included log-transformed reading times as the predicted variable, and extraction type (argument vs. adjunct) as the predictor factor. We also included random intercepts for participants and items. Models with slopes by extraction type, slopes by participants, and slopes by items failed to converge, and thus these factors were not included in the model. For the main verb first conditions, we found no effect of extraction type in the filled-gap region (rugī-ke ‘patient-ACC’, $\beta = 0.04$, SE = 0.04, t(259) = 1.0, $p = 0.32$). However, in the region immediately following the filled-gap region, there was a main effect of extraction type, due to longer reading times in the argument extraction condition (boleche ‘said’, $\beta = 0.08$, SE = 0.04, t(270) = 2.1, $p = 0.04$). This indicates a filled-gap effect for the main verb first word order, suggesting that readers made an early commitment to a gap for the wh-filler in this position. For the embedded verb first conditions, there was no effect of extraction type at the filled-gap region (rugī-ke ‘patient-ACC’, $\beta = -0.03$, SE = 0.04, t(285) = -0.69, $p = 0.49$) or in the following two regions (cikīśā ‘treatment’, $\beta = 0.05$, SE = 0.04, t(302) = 1.3, $p = 0.20$; korechen ‘did’, $\beta = 0.003$, SE = 0.004, t(300) = .08, $p = 0.94$). Thus, we found no filled-gap effect in the embedded verb first word order. The mean reading times by region are given in Figure 2 and Figure 3.

5.5 Discussion

Experiment 2 was designed to probe for sensitivity to blocking of a preferred gap location across word orders using an on-line measure. If the preference for linearly local gaps found in Experiment 1 reflects an early and confident commitment in both word orders, then we predicted sensitivity to disruption when this resolution was unavailable in the filled-gap paradigm. However, we only found sensitivity to disruption in the main verb first word orders, i.e., when structural locality and linear locality converged. In other words, we did not find evidence of early commitment to this position in Experiment 2. This suggests that the class of gaps that are preferred is not identical to the class of gaps that are committed to early enough to elicit a filled-gap effect.

This difference in measures may be due to a selective sensitivity to structural locality. For instance, the bias to resolve with the first gap linearly available may only manifest as an early commitment when this position is also structurally local. If the biases for preferred gap sites are sensitive to structure in this way, then this undermines one argument for the separation of island constraints from biases on preferred gap sites, i.e., the argument that they should be separated because they refer to different properties of the representation.
However, the differences between the results in Experiments 1 and 2 may reflect differences between the tasks. Experiment 1 probed directly for resolution preferences in an off-line task. The sentences were globally ambiguous, and somewhat simpler than the three-clause sentences in Experiment 2. Conversely, Experiment 2 was an on-line reading task in which participants read sentences word-by-word, somewhat artificially.

Participants also seemed to have some difficulty with this task, since their accuracy on the comprehension questions are somewhat lower than average. Additionally, in pilot versions of Experiment 2, participants read at very different paces. Furthermore, participants reported different levels of familiarity with computers, which may have exacerbated the unnaturalness of the task. Some participants struggled with the instructions, e.g., some participants held the space bar down, failing to release it between words. These additional complications in Experiment 2 may have masked an early commitment to resolve a filler-gap dependency with the first gap linearly available, even in embedded verb first word orders. In other words, participants may have made an early commitment to resolve with a linearly local gap site in embedded verb first word orders, but this was selectively masked in Experiment 2.

In Experiment 3, we address these cross-experimental concerns by using the filled-gap paradigm in an offline acceptability judgment task. Experiment 3 was an off-line acceptability judgment task that used two-clause sentences. This was intended to make it as similar to Experiment 1 as possible. The target sentences all contained a filled-gap in a linearly local position, as in Experiment 2. Thus, Experiment 3 relied on an indirect measure of disruption like Experiment 2. Additionally, Experiment 3 was an untimed pen-and-paper task, like Experiment 1, nullifying the unnaturalness of the self-paced reading paradigm in Experiment 2. If we find evidence of sensitivity to disruption only in main verb first word orders in Experiment 3, i.e. when structural and linear locality converge, then we can conclude that structural locality affects the processes involved in making an early commitment to a gap site. Conversely, if there is a filled-gap effect in Experiment 3 across word orders, then we can infer that the failure to find a filled-gap effect in Experiment 2 was due to the design of that experiment.

6 Experiment 3

6.1 Rationale

In Experiment 3, we again investigated whether Bangla speakers prefer gap sites that are linearly local or structurally local. In Experiment 1, we found evidence for an off-line bias for linearly local gap sites. In Experiment 2, we found sensitivity to disruption with a filled-gap in a linearly local position, but only with main verb first word order. In Experiment 3, we investigated whether this mismatch between the results in Experiments 1 and 2 was due to the on-line/off-line contrast between the studies, or the ambiguity resolution/filled-gap paradigm contrast.

Experiment 3 was an off-line acceptability judgment task. In this task, participants read sentences in an untimed way, as in Experiment 1. However, like in Experiment 2, we used a filled-gap paradigm. Although the filled-gap paradigm is typically used in on-line measures, it can also be used to detect filled-gap effects in off-line measures (Sprouse, 2008). This is because the disruption associated with a filled-gap effect also lowers ratings in acceptability judgment tasks. Thus, we can compare the ratings for sentences in which the preferred gap is unavailable with controls. If we find a decrease in acceptability, then we take this to be a filled-gap effect. If we find a filled-gap effect across both word orders in Experiment 3, then we can infer that the lack of an embedded clause bias in embedded verb first word orders is due to the design of Experiment 2. Conversely, if we find evidence for a
filled-gap effect in the main verb first word order only, then this implies that the difference between Experiments 1 and 2 may be due to the different nature of ambiguity resolution tasks (Experiment 1) and the mechanisms involved in detecting filled-gap effects (Experiments 2-3).

6.2 Materials

The materials in Experiment 3 were constructed in a similar way to the materials from Experiment 2. We crossed three factors – word order (main verb first or embedded verb first), extraction type (argument extraction or adjunct extraction), and filled-gap position (linearly local or linearly distant). This third factor was added to test for any filled-gap effect with the main clause verb in embedded verb first word orders, i.e., to probe for a filled-gap effect in a position that was linearly distant but structurally local. We constructed 8 lists with an equal number of items per condition, and an equal number of items across lists. There were 24 sets of target items, and 36 complexity-matched fillers, 18 of which were ungrammatical. Each participant saw 3 sentences from each condition and all the fillers in a randomized order.

There were a few differences between the target items in Experiments 2 and 3 that are worth noting. First, the target items in Experiment 3 contained two clauses, unlike the three clause sentences in Experiment 2. This is more similar to the materials in Experiment 1. Additionally, the wh-phrase appeared in the preverbal position like in Experiment 1, not the left-edge position as in Experiment 2. This was done because the preverbal position is the more canonical position for wh-fillers (Simpson and Bhattacharya, 2003). Lastly, in the adjunct conditions we did not include the additional object pronoun. In Experiment 2, we included this extra pronoun to ensure that the verb with which the argument wh-filler was interpreted had an overt argument in the adjunct extraction conditions. However, this may have been unnecessary, since Bangla permits null arguments. Thus, we did not include this extra pronoun, to maximize similarity between the argument and adjunct extractions. A sample set of materials is given in Table 5.

6.3 Methods

Experiment 3 was a pen-and-paper acceptability judgment study. Participants were instructed to read the sentences carefully, and then circle a number ranging from 1 to 7, with lower scores indicating unacceptability. They were given example sentences with values already circled to illustrate how to use the scale.

Participants were adult native speakers of Bangla drawn from the University of Dhaka and Calcutta University student populations. There were 32 participants from each group. Participants in Dhaka were compensated 500 BDT for their time, and participants in Kolkata were compensated 200 INR. The experiment lasted approximately 10-20 minutes, and was conducted after either Experiment 2 or another unrelated experiment.

6.4 Results

We submitted the ratings to a linear mixed effects model, using the lme4 package (Bates et al. 2015). We included random effects for participant and item. We included word order (main verb first or embedded verb first), extraction type (argument extraction or adjunct extraction), and filled-gap position (linearly local or linearly distant) as predictors, together with their interaction terms. We also included location (Dhaka or Kolkata) in the model. The estimates of the model are presented in Table 6. The means of the ratings by condition are given in Figure 5. We then performed pairwise
comparisons for extraction type within the two word orders and two filled-gap positions, using the least-squares means estimates with Tukey adjustment. These are shown in Table 7.

There were two main findings in Experiment 3. First, there was a main effect of word order. Ratings were significantly increased in main verb first word order (β = 0.87, SE = 0.20, t(1435) = 4.26, p = 2e-5). This is consistent with the observation that main verb first word orders are the preferred word order for clausal embedding in Bangla. Secondly, there was a three-way interaction between word order, filled gap position, and extraction type (β = -1.39, SE = 0.40, t(1435) = -3.4, p = 0.0006). The only significant pairwise comparison was between argument and adjunct extraction in the main verb first, local filled gap conditions (β = 1.28, SE = 0.20, t(1450) = 6.38, p < 0.0001*). This reflects the lowered ratings with main verb first word order, local filled gap, and argument extraction conditions. In other words, there was a decrease in ratings when an argument wh-filler could not resolve with the main verb in main verb first word orders. This is a replication of the filled-gap effect in Experiment 2 in off-line acceptability judgments. Crucially, this was only observed in the word orders in which linear and structural locality aligned, i.e., we only found a filled gap effect in situations where the first potential gap position was both structurally and linearly local.

6.5 Discussion

The goal of Experiment 3 was to determine whether Bangla speakers are sensitive to disruption of a linearly local filler-gap dependency resolution using an off-line measure. We conducted Experiment 3 to determine whether the lack of a filled-gap effect in embedded verb first word orders in Experiment 2 was due to the design of that experiment, or whether it reflects a difference between sensitivity to disruption and general locality preferences, as explored in Experiment 1.

The results from Experiment 3 show that Bangla speakers are only sensitive to disruption of a linearly local filler-gap dependency resolution in main verb first word orders. There was no filled-gap effect with linearly distant filled-gaps in either word order, and there was no filled-gap effect with linearly local filled-gaps in embedded verb first word orders. In other words, we again found a contrast between word orders with respect to sensitivity to disruption with the linearly local gap position.

Thus, we conclude that there is a general bias to resolve filler-gap dependencies with the first position linearly available, but that this only translates into an early and confident commitment when this is also a structurally local position. This means that the bias to resolve with the first position linearly available is only one component of detecting a disrupted filler-gap dependency, as measured with paradigms like the filled-gap effect.

7 General Discussion

In this paper, we investigated filler-gap dependency formation in Bangla. Bangla features flexible word order that permits us to manipulate whether the first position linearly available is in an embedded clause or in the main clause. This allows us to manipulate whether linear and structural locality converge or diverge. In main verb first word orders, the first gap position linearly available is also structurally local, whereas in embedded verb first word orders the first gap position linearly available is structurally distant. In Experiment 1, we found a strong bias to resolve an ambiguous filler-gap dependency with the first position linearly available, regardless of its structural depth. However, in Experiments 2 and 3, we only found evidence of sensitivity to disruption when linear and structural locality converge. We interpreted these results as showing that there is a general bias
for linearly local gaps, but that this only translates into a strong early commitment to this gap site when the linearly local gap is also structurally local.

If we start with the assumption that preferred gaps are typically detectable in a filled-gap paradigm, then the finding that these measures diverge for embedded verb first word orders in Bangla is surprising. The filled-gap paradigm, the measure of sensitivity to disruption that we used in Experiments 2 and 3, depends on multiple processing mechanisms. The paradigm requires that participants make an early and confident commitment to a gap site, presumably in accordance with their linear locality preferences. Upon encountering the filled gap, the comprehender must quickly detect that the preferred gap is unavailable, and then instigate a costly reanalysis process. The lack of a filled gap effect in embedded verb first word orders might be attributed to any of these processes failing to apply quickly.

Bangla speakers may not have shown a filled-gap effect in embedded verb first conditions in Experiments 2 and 3 because this word order is dispreferred. This was reflected in the lowered ratings for this word order in Experiment 3. This may be in part because preverbal embedded clauses have specific semantic and syntactic restrictions, unlike postverbal embedded clauses (e.g., Bayer, 2001). In our estimation, long preverbal embedded clauses are also likely less frequent in naturalistic speech, and may carry certain pragmatic or discourse functions that also make them atypical. As a consequence, Bangla speakers may find processing preverbal embedded clauses more difficult, and have less facility making fine-grained predictions in preverbal embedded clauses for that reason. This contrasts with Japanese, in which preverbal embedded clauses are canonical (Tanaka, 2001), and filled-gap effects are found in preverbal embedded clauses (Aoshima et al., 2004; Yoshida, 2006).

Alternatively, the difference in the filled-gap effect between word orders may be attributed to differences in information flow. In both word orders in Experiments 2 and 3, the wh-filler in the argument extraction conditions is blocked from resolving with the first position linearly available by another noun phrase with the same accusative case marking. No verb in Bangla can take two accusative arguments. The case-marking is the first cue that the wh-filler cannot resolve in this position. In the main verb first word order conditions, this cue comes several words after the wh-filler. The only intervening linguistic material is an adverbial in Experiment 2 (e.g., khubi ɔbakkhahe ‘very surprisedly’), and the subject of the clause and adverbial material in Experiment 3 (e.g., daktar-ta khubi ɔbakkhahe ‘doctor-CL very surprisedly’). The comprehender is likely only expecting one verb at this point, because there is no evidence for an additional clause. Thus, encountering two accusative-marked noun phrases is sufficient for detecting the local ungrammaticality. The additional intervening linguistic material should not dramatically affect this process. Conversely, in the embedded verb first word order, several other computations must take place before encountering the filled-gap. In this word order, there is an additional nominative noun phrase intervening between the wh-filler and the embedded clause. This means that the comprehender begins a reanalysis process before encountering the filled gap, since the additional nominative is a cue to construct the embedded clause (Miyamoto, 2002). Constructing the additional clause may obscure the relation between the two accusative-marked noun phrases, since two accusative marked noun phrases does not necessarily induce a local ungrammaticality. Alternatively, the additional clause may facilitate processing at the filled gap, because the comprehender now has constructed two predicates that the wh-filler can resolve in – the preferred embedded clause, and the dispreferred main clause. In other words, encountering the filled gap is less costly because the gap site can be “moved” to the already constructed embedded clause cheaply.

8 Conclusion
Much work in theoretical linguistics and psycholinguistics demonstrates that there are robust locality constraints on gaps. Both structural locality and linear locality play important roles in selecting gaps in real-time sentence processing. Structural locality is relevant for determining which gap sites are grammatically well-formed, and linear locality is relevant for determining which gaps are preferred when multiple potential gaps are available (Aoshima et al., 2004; Omaki et al., 2014). Locality biases on filler-gap dependencies can reveal themselves in different ways – as a general preference for certain gap sites, or as an early commitment. Typically, these are taken to reflect the same processes of active dependency formation, but different measures show that they dissociate. We investigated the dissociation of linear locality and structural locality and the dissociation of gap locality preferences and early gap commitments in Bangla. Bangla is a particularly useful language for investigating filler-gap dependencies and locality, because of its flexible word order properties, which permit us to manipulate the structural position of the first gap linearly available, which allows testing the contribution of structural and linear locality.

In Experiment 1, we showed that Bangla speakers have a preference for linearly local gaps, regardless of structural position. This replicated findings from a previous English and Japanese comparison within the same language (Omaki et al., 2014), and thus supports the generalization that filler-gap dependency locality preferences are primarily sensitive to linear locality. In Experiments 2 and 3, we showed that this preference for linearly local gaps only leads to early commitment in main verb first word orders. We highlighted a few reasons why this difference between word orders might hold. Specifically, we suggested that gaps in preverbal embedded clauses may be difficult to maintain, because of the status of preverbal embedded clauses in Bangla. Alternatively, we suggested that the informativity of the filled-gap cue might differ between word orders, which may impact the confidence in the early gap commitment. These facts contrast with Japanese, which has a strong bias for gaps in preverbal embedded clauses (Aoshima et al., 2004; Yoshida, 2006; Omaki et al., 2014). We leave further investigation into the differences between these two languages for future research.

9 Acknowledgments

The authors would like to thank Awalin Sopan, Titir Santra, and Fuad Hasan Sobuj for assistance in preparing the materials and recruiting participants. We previous presented these results at the 26th and 27th CUNY Sentence Processing conferences, and the 3rd and 4th Formal Approaches to South Asian Languages conference. This work was funded by grant NSF-GRFP grant DGE-0750616 awarded to DAC, NSF grant NSF BCS-0848554 awarded to CP, and NSF grant DGE-0801465 awarded to UMD.

10 References


Table 1: Distribution of Participants in Experiment 1. There were 96 participants in Experiment 1, 48 for each city, Dhaka and Kolkata. Each city was split into two groups. One group of 24 in each city saw both conditions in the same questionnaire, the within-participants group. Another group of 24 in each city, was further divided in two groups of 12, one seeing only lists with main verb first word orders and the other seeing only lists with embedded verb first word order.
## Table 2: Proportion of removed responses in Experiment 1.

<table>
<thead>
<tr>
<th></th>
<th>Dhaka Between Participants</th>
<th>Dhaka Within Participants</th>
<th>Kolkata Between Participants</th>
<th>Kolkata Within Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main verb first</strong></td>
<td>25%</td>
<td>29%</td>
<td>31%</td>
<td>21%</td>
</tr>
<tr>
<td><strong>Embedded verb first</strong></td>
<td>15%</td>
<td>33%</td>
<td>8%</td>
<td>23%</td>
</tr>
<tr>
<td>Fixed Effects</td>
<td>Estimate</td>
<td>SE</td>
<td>z</td>
<td>p</td>
</tr>
<tr>
<td>---------------------</td>
<td>----------</td>
<td>-----</td>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>(Intercept)</td>
<td>-2.03</td>
<td>1.44</td>
<td>-1.41</td>
<td>0.16</td>
</tr>
<tr>
<td>Word Order</td>
<td>5.22</td>
<td>2.31</td>
<td>2.26</td>
<td>0.02*</td>
</tr>
<tr>
<td>City</td>
<td>-2.51</td>
<td>1.56</td>
<td>1.61</td>
<td>0.11</td>
</tr>
<tr>
<td>List Type</td>
<td>-0.71</td>
<td>1.27</td>
<td>-0.56</td>
<td>0.57</td>
</tr>
<tr>
<td>Order * City</td>
<td>2.53</td>
<td>2.16</td>
<td>1.17</td>
<td>0.24</td>
</tr>
<tr>
<td>Order * List Type</td>
<td>-0.40</td>
<td>1.84</td>
<td>-0.22</td>
<td>0.83</td>
</tr>
<tr>
<td>City * List Type</td>
<td>3.56</td>
<td>2.10</td>
<td>1.69</td>
<td>0.09</td>
</tr>
<tr>
<td>Order * City * List Type</td>
<td>-1.61</td>
<td>2.84</td>
<td>-0.57</td>
<td>0.57</td>
</tr>
</tbody>
</table>

Table 3: Results of best-fit logistic regression model for Experiment 1.
Rashad asked {who|where} ...  

Main Verb First:  ...  đakter-ṭa khubi ṭbakhhabe ṭugi-ke bolechen je tini purano  

... doctor-CL very surprisingly patient-ACC told that he.POL old  

hospatal-e cikitşa korechen  

hospital-LOC treatment did  

‘Rashad asked {who|where} the doctor told the patient that he treated in the old hospital’

Embedded Verb First:  ...  đakter-ṭa tini purano haspatal-े ṭugi-ke cikitşa korechen bole khubi  

... doctor-CL he.POL old hospital-LOC patient-ACC treatment did that very  

ṭbakhhabe bolechen  

surprisedley told  

‘Rashad asked {who|where} the doctor told/said that he treated the patient in the old hospital’

Table 4: Sample materials from Experiment 2.
Jahid {ka-ke|kothae} ... {Argument|Adjunct}

Jahid {who-ACC|where} …

---

**Main V,** 
Local FG

... *khubi ōbakbhabe* tar *bondhu-ke* *boleche je nipa party-te* dekheche

... very surprisingly *his friend-ACC* said that *nipa party-at* saw

‘Who/where did Jahid very surprisingly tell *his friend* that *Nipa saw* at the party?’

**Main V,** 
Distant FG

... *khubi ōbakbhabe* *boleche je nipa party-te* *tar bondhu-ke* dekheche

… very surprisingly told that *nipa party-at* *his friend-ACC* saw

‘Who/where did Jahid very surprisingly tell that *Nipa saw* his friend at the party?’

---

**Embedded V,** 
Local FG

… *nipa party-te* *tar bondhu-ke* *dekheche* *bole* *khubi ōbakbhabe* boleche

… *Nipa party-at* *his friend-ACC* saw that very surprisingly told

‘Who/where did Jahid tell very surprisingly that *Nipa saw* his friend at the party?’

**Embedded V,** 
Distant FG

… *nipa party-te* *dekheche* *bole* *khubi ōbakbhabe* *tar bondhu-ke* boleche

… *Nipa party-at saw* that very surprisingly *his friend-ACC* told

‘Who/where did Jahid tell *his friend* very surprisingly that *Nipa saw* at the party?’

---

Table 5. Sample materials for Experiment 3.
<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Estimate</th>
<th>SE</th>
<th>t</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Intercept)</td>
<td>4.43</td>
<td>0.22</td>
<td>20.0</td>
<td>152</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Word Order</td>
<td>0.87</td>
<td>0.20</td>
<td>4.26</td>
<td>1435</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Filled-Gap Position</td>
<td>-0.04</td>
<td>0.20</td>
<td>-0.21</td>
<td>1428</td>
<td>0.83</td>
</tr>
<tr>
<td>Extraction Type</td>
<td>-0.25</td>
<td>0.20</td>
<td>-1.25</td>
<td>1432</td>
<td>0.21</td>
</tr>
<tr>
<td>Location</td>
<td>-0.08</td>
<td>0.25</td>
<td>-0.32</td>
<td>62</td>
<td>0.75</td>
</tr>
<tr>
<td>Word Order * FGPosition</td>
<td>0.06</td>
<td>0.29</td>
<td>0.21</td>
<td>1450</td>
<td>0.83</td>
</tr>
<tr>
<td>Word Order * ExtType</td>
<td>0.24</td>
<td>0.29</td>
<td>0.85</td>
<td>1434</td>
<td>0.40</td>
</tr>
<tr>
<td>FGPosition * ExtType</td>
<td>0.11</td>
<td>0.28</td>
<td>0.38</td>
<td>1430</td>
<td>0.70</td>
</tr>
<tr>
<td>WOrder * FGPosition * ExtType</td>
<td>-1.39</td>
<td>0.40</td>
<td>-3.44</td>
<td>1435</td>
<td>0.0006*</td>
</tr>
</tbody>
</table>

Table 6. Results from Experiment 3.
<table>
<thead>
<tr>
<th>Word Order, Filled Gap Position</th>
<th>Estimate</th>
<th>SE</th>
<th>t ratio</th>
<th>df</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main Verb First, Local Filled Gap</td>
<td>1.28</td>
<td>0.20</td>
<td>6.38</td>
<td>1450</td>
<td>&lt;0.0001*</td>
</tr>
<tr>
<td>Main Verb First, Distant Filled Gap</td>
<td>0.009</td>
<td>0.20</td>
<td>0.043</td>
<td>1436</td>
<td>0.97</td>
</tr>
<tr>
<td>Embedded Verb First, Local Filled Gap</td>
<td>0.14</td>
<td>0.20</td>
<td>0.70</td>
<td>1428</td>
<td>0.48</td>
</tr>
<tr>
<td>Embedded Verb First, Distant Filled Gap</td>
<td>0.25</td>
<td>0.20</td>
<td>1.24</td>
<td>1433</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Table 7. Results from pairwise comparisons in Experiment 3. Comparisons were between argument and adjunct extractions within word order and filled-gap position. Comparisons were made using least squares means with Tukey HSD adjustment.
Figure 1: Sample image from Experiment 1 materials. In this vignette, the character Shumi successfully caught butterflies in the park, and reported on it in the first floor. The parking lot and bedroom are distractor locations.
Figure 2: Proportion main clause response by word order in Experiment 1. Error bars correspond to one standard deviation from the mean. Proportions are collapsed across list types and locations.
Figure 3: Reading times by region in Experiment 2, main verb first conditions. Mean reading times by region. Line type corresponds to extraction type. Error bars represent one standard error from the mean.
Figure 4: Reading times by region in Experiment 2, embedded verb first conditions. Mean reading times by region. Line type corresponds to extraction type. Error bars represent one standard error from the mean.
Figure 5: Ratings by condition in Experiment 3. Error bars represent standard errors of the mean.