Abstract

This study examines the morphosyntactic development, specifically verb morphology, of typically-developing Bangla-speaking children between the ages of two and four. Three verb forms were studied: the Present Simple, the Present Progressive, and the Past Progressive. The study was motivated by the observations that reliable language-specific developmental information is not available in Bangla (which affects research in the areas of speech and language therapy and computational linguistics) and that properties of these verb forms render them suitable for exploring how language typology contributes to the learnability of verb morphology in emerging child language. Children’s performance on these forms was assessed through form-specific language elicitation tasks and spontaneous language samples. Three stages of development of verb morphology were identified by consideration of accuracy of production and error types.
Keywords

Language acquisition, language typology, verb morphology, Bengali, Bangla

Introduction

Typological characteristics of languages have often been found to regulate the patterns of language development in children. Crosslinguistic investigations have revealed that, depending on the morphological character of the language investigated, young typically-developing (TD) children’s productions of verbal inflections is characterised by omission of obligatory tense markers on verbs, failure to mark agreement correctly on verbs, or substitution of incorrect verb forms for target verb forms (e.g. Leonard, Caselli, & Devescovi, 2002; Wexler, 1994). Little research so far has focused on the developmental trends of Indo-Aryan languages, resulting in an under representation of languages with certain morphological properties in the acquisition literature, and thereby making it difficult to develop language-appropriate assessments and interventions for children with language impairment. This paper describes Bangla-speaking children’s development of verb morphology vis-à-vis three forms of verbal inflections.

Typological influences on the development of verb morphology

Fundamental typological differences among languages contribute to variance in how morphological markers are expressed (Leonard, 2014b). Languages differ in flexibility of word order, morphological typology, optionality of morphological markers, and presence/absence of subjects and objects, and, for synthetic languages, whether inflectional systems are predominantly fusional (one morph representing more than one grammatical function) or...
agglutinative (morphs showing a one-to-one relationship with their grammatical functions). These differences impact on the order of emergence and the nature of developmental errors. For example, in German, a fusional non-null-subject language, children’s errors with German tense markers are largely omissions, resulting in non-finite forms (Bittner, 2003). Errors on the tense markers are also abundant in French child data (Kilani-Schoch, 2003). On the other hand, in Italian (a fusional null-subject language), children’s errors are typically ‘near-misses’ that deviate from the target forms by one feature (Pizzuto & Caselli, 1992). Such substitutions are often unidirectional where the unmarked forms replace the more marked forms (e.g. plural forms replaced by the singular forms) (Leonard, Caselli, & Devescovi, 2002). Therefore, unlike the use of non-finite forms found in English or German, in Italian, a third person singular form is a more likely substitute for a third person plural or a first person singular, than for a first person plural.

Turkish, an agglutinative language with exceptionally heavy inflectional morphology, contributes significantly to explaining the influences of typological issues, and provides a platform for research on Bangla. The two languages share the properties of the same word order, flexibility within a given order, agglutinative features, transparency of the inflections and null-subjecthood. Turkish data display evidence for a hierarchy in verb inflections regulated by the cognitive complexity of the concepts the inflections signal (Aksu, 1978; Aksu-Koc & Slobin, 1985). A remarkable finding of the Turkish data is, unlike the early utterances of children in many languages studied, Turkish-speaking children’s utterances are not telegraphic and they contain grammatical markers that are correctly produced (Aksu-Koc & Slobin, 1985). Possible reasons for the relative absence of errors in young Turkish
children’s morphological productions include: i) inflections are postposed, syllabic and stressed, increasing the salience of the markers; ii) inflections are obligatory, creating consistent positive evidence for children; and iii) inflections are transparent and almost all of them have distinct phonetic forms. These reasons for exceptionally high accuracy rates are also found in Bangla to an extent, which makes the Turkish research a significant reference for child language research in Bangla.

In order to position the present study relative to the typological argument, a description of Bangla verb morphology is presented below. As the following description indicates, its inflectional structure resembles Turkish, of the languages in which morphological development has so far been studied, rather than fusional or isolating languages.

Description of Bangla/ Bengali

Bangla or Bengali, an Indo-Aryan language, has branched out from Sanskrit (600-650 BC) through an intermediate language- Prakrit. Bangla is widely spoken in Bangladesh, in the Indian states of PashchimBanga, Assam and Tripura, and among the Bengali migrant population in countries such as the USA, UK, Australia and Canada, and many middle-eastern countries. There are over 107 million first language speakers of Bangla in Bangladesh (Bangladesh Bureau of Statistics, 2011) and 83 million speakers in different states of India (Encyclopaedia Britannica, 2016). With approximately 250 million first and second language speakers, Bangla is now considered the 7th most widely spoken language of the world (Ethnologue, 2016).
Bangla is a null-subject language, and Bangla verbs exhibit agglutinative features in the morphological paradigm, i.e. verbal inflections are suffixed to the verbs with each having a distinct morphological significance (Kar, 2009). Bangla verbs are marked for aspect, tense, and person, with the person marker the only obligatory suffix. In present tense contexts, tense marking is not overtly realised on verbs. Aspect marking on Bangla verbs is realised through two aspect markers: progressive and perfective. There are three distinct person markers indicating first, second, and third person agents. Inflections are attached in a fixed order: verb stem- aspect- tense- person. A present simple verb form contains the verb stem and a person marker, while the past progressive form contains the verb stem, the progressive aspect marker, the past tense marker, and a person marker (Table 1). For example, the inflected verb form pore ‘reads’ contains the verb stem /por-/ and the third person marker /-e/, while the form porchhilo ‘was reading’ contains the verb stem /por-/, the progressive aspect marker /-chh/i-/i/, the past tense marker /-l-/i/ and the third person marker /-o/. The verbal system is linear and transparent to a great extent. Also note that there are phonological variations among the aspect markers that are determined by whether or not the verb stem ends in a vowel. The progressive aspect markers contain geminated forms (/chh/i/- and /chh/i/i/-) when the stems end in vowels (e.g. /kha-/).

Table 1. Examples of verb inflections in Bangla across two verbs where the general form is verb stem - aspect marker - tense marker - person marker*

<table>
<thead>
<tr>
<th>Verb form</th>
<th>Verb stem 1: /por-/</th>
<th>Verb stem 2: /kha-/</th>
</tr>
</thead>
<tbody>
<tr>
<td>Present Simple</td>
<td>She boi pore</td>
<td>She khae</td>
</tr>
<tr>
<td>Verb root - Ø- Ø- e</td>
<td>(S)he book(s) reads</td>
<td>(S)he eats</td>
</tr>
</tbody>
</table>
Verb stem /por-/ (to read) ends in a consonant, verb stem /kʰa-/ (to eat) ends in a vowel, Ø- indicates absence of a grammatical marker.

*Only 3rd person forms are displayed in the table.

**Previous findings from Bangla**

Chakraborty and Leonard (2012) reported high accuracy rates for 19 Bangla-speaking preschool children in Kolkata. They provided responses to elicitation probes for first, second and third person present progressive and past progressive forms encoding continuous aspect (Table 2). The high accuracy rates were attributed to the relatively transparent nature of the composition of Bangla verb morphology and the fact that Bangla verbs agreed with the subject only in person. No difference was observed between children’s accuracy in the present and the past progressive forms. Of tense-aspect-person errors in the present progressive tasks, only 2% were incorrect aspect markers and only 4% were incorrect tense markers; the rest were incorrect person markers. Of the tense-aspect-person errors in the past progressive tasks, tense and aspect inaccuracies constituted only 10% and 5% of the errors respectively indicating a large 76% of errors being inappropriate person markers. They concluded that person markers in Bangla were more difficult for children to learn than the tense and aspect markers.
Table 2. Accuracy (%) in the Bengali verb inflections across the age range (2;0- 4;0)

<table>
<thead>
<tr>
<th></th>
<th>Present Progressive</th>
<th>Past Progressive</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st</td>
<td>2nd</td>
</tr>
<tr>
<td>Mean</td>
<td>93.26</td>
<td>71.16</td>
</tr>
<tr>
<td>SD</td>
<td>5.97</td>
<td>23.43</td>
</tr>
</tbody>
</table>


Note that this study of Bangla included a total of 19 children aged 1;6 – 4;0 and only children who could respond to the majority of test items were included in the analysis. Although the mean accuracy of response was high, there were individual differences, with some children scoring as low as 54% on the 3rd person past progressive form. Given the small sample size, the inclusion criteria, and absence of any analysis of errors forms, as well as possible dialect differences between Kolkata, India and Dhaka, Bangladesh, where our study is based, a more extensive treatment is called for.

Aim

The aim of this study was to map the developmental pathway of a set of verb inflections in a relatively large sample of Bangla-speaking pre-school children across a wider age range than studied previously. Children’s performance in producing the tense and aspect marker combinations used in the present simple, the present progressive and the past progressive forms were examined. The research questions were:

1. What is the developmental route for acquiring the Present Simple, the Present Progressive, and the Past Progressive verb inflections in Bangla for children aged two to four years?
2. What error patterns occur for target forms?

3. How do error patterns inform the process of morphological development in Bangla?

Methods

Participants

There were 70 typically-developing children between 23 months (1;11 years) and 51 months (4;3 years), mean age = 36.96 (SD = 7.6 months), 33 boys and 37 girls. At 23-35 months, 36 – 47 months, and >48 months there were 31, 32 and 7 children respectively. These children were recruited from six daycare centres and pre-schools in Dhaka, Bangladesh, with consent from the daycare centres/ schools. Parents completed a consent form and a demographic questionnaire. The inclusion criteria included no reported speech and/ or language delay or disorder; no family history of speech and/ or language delay or disorder; and exposure to one language only, Bangla. Although SES was not directly measured, four of the six daycare centres operated as part of four companies and they cared for children of employees from all levels of the company, resulting in a diversity of socio-economic status in the sample.

The study was conducted with approval from the Human Ethics Committee of the University of Canterbury, New Zealand (Ref.no. HEC 2012/166).

Tasks

Three language production tasks were designed, piloted, modified, and administered to elicit three different verb conjugations in an experimental format, and a 20-minute language sample was collected for subsequent analysis. All of the language tasks were
Structured conversation (Present Simple). The first author engaged in a conversation with each participant about the daily activities of family members. Some of the stimuli used are presented below:

1. Ammu rannaghore ki kore?
   Mummy kitchen. Loc what do PresSimp 3p
   
   *What does mummy do in the kitchen?*

2. Dadu shokale ki kore?
   Granny morning. Time what do PresSimp 3p
   
   *What does granny do in the morning?*

Since it was a conversation, the number of responses varied (7 to 37 utterances). Samples that contained at least six responses (utterance tokens) were entered into the analysis. Based on this criterion, four children were excluded from the analysis.

Picture book Task (Present Progressive). A picture book of the cartoon character Barney titled *Let’s Go to School* (Bernthal & Full, 2006) was used in this task. The book contained the story of Baby Bop going to school where she and her friends were shown to be engaging in activities (Appendix). Ten pictures were chosen from the book and children were asked set questions, for example ‘*What is Bop doing here?*’ ‘*What are her friends doing?*’ The pictures were chosen based on three criteria. Firstly, the picture had to illustrate an action depicting an early developing verb. As there are no published data on Bangla lexical development,
confirmation was obtained from English and other languages as published on the website of Cross Linguistic Lexical Norm study (http://www.cdi-clex.org). Secondly, a native speaker (the first author) judged whether or not the Bangla counterparts of the identified verbs were common in Bangla. Finally, verbs resulting from this process that could be unambiguously identified from the pictures were chosen for the task.

**Paired picture task (Past Progressive).** There were 10 picture pairs to elicit 10 verbs in their past progressive forms. The first picture of each pair depicted a person or an animal sleeping and the other picture showed the same person or animal doing an action such as eating, running or cooking. Showing the first picture the first author told the child in Bangla, *’Look, the boy is sleeping here. He is sleeping because he is very tired. Do you know why he is tired? Because he was doing something a while ago. Can you say what he was doing?’* Looking at the second picture, children were expected to say what the actor had been doing (Appendix). The same selection criteria as the Present Progressive task were employed for the verb stems. To demonstrate the requirements, two practice items were administered for each task.

**Language sample.** Twenty-minute language samples were collected during a play session between each child and the experimenter (Klee, Stokes, Wong, Fletcher, & Gavin, 2004). Although some researchers have found short samples to be reliable (Heilmann, Nockerts, & Miller, 2010), Gavin and Giles (1996) stated that temporal reliability of language sample measures required longer samples ($\geq 175$ complete and intelligible utterances or 20 minutes). Set toys were a doll’s house, a cooking set, a toy car and some puppets.

**Procedure**
Testing took place in a separate, bright and reasonably quiet room in the children’s daycare centres/ preschools. Interested parents completed and returned the consent forms and questionnaires, and the first author contacted them by telephone to ask a further set of informal questions about family members, to set the scene for the present progressive task. All the tasks and the free play for each child were completed within an hour in one or two sessions within a three-day period. All sessions were audio- and video-recorded using a lapel microphone and a Canon Powershot S5IS camera.

**Scoring**

The tasks and the language samples were transcribed using the conventions of Systematic Analysis of Language transcripts (SALT-NZ) (Miller, Gillon, & Westerveld, 2012) adapted for Bangla (Sultana, 2016), by the first author, who also coded and scored them. The Present Simple task required children to answer questions about family activities. No minimum or maximum possible scores were identified as the forms were generated in conversation. Null responses, use of other markers, utterances with non-finite verb forms, and utterances without verbs (where verbs were expected) were considered errors. There were 10 picture stimuli each for the Present and Past Progressive tasks. Scores were calculated as percent correct. Null responses and imitated responses were considered incorrect. In addition, since the task aimed to test children’s knowledge of inflections and not the verbs themselves, correct inflections with non-target stems were given credit. For calculating MLU from children’s language samples routinised expressions were considered as one unit and written as a word, i.e. *thikase* (It’s alright), *eije* (Here/ There you go). Also, compound words (or semi-compound words) were transcribed as one word.

**Reliability**
Conventional calculations of reliability were used (Fletcher, Leonard, Stokes & Wong, 2005). A second transcriber, a linguistics graduate and a first language speaker of Bangla who was not involved in the study, independently transcribed 20% of the samples. She read a transcription manual (Klee, 2010) as a preparation for the transcription and was trained for the task by the first author.

Two measures were calculated from the transcripts: point-to-point morpheme agreement and point-to-point agreement in utterance boundaries. To calculate the agreement in morphemes, each morpheme was compared in the two transcripts of each participant. Percent agreement was calculated by dividing the number of agreements by the total number of morphemes and multiplying the score by 100. Morphemes coded as unintelligible were excluded from comparison. In order to calculate the agreement in utterance boundary, the number of agreements was divided by the number of utterances and then multiplied by 100. There was 98.9% and 99.7% agreement in morphemes in children’s and examiner’s utterances respectively, and 97.4% and 95.4% agreement in utterance boundaries for children and examiner respectively.

Analysis

All statistical analyses were conducted using IBM SPSS Statistics 20.0 (IBM Corporation, 2011). The major analyses include descriptive measures of items correct, error counts, cluster analysis, and one- and two-way ANOVAs. Descriptive measures of children’s accuracy were calculated in percentages since the number of responses varied across children and tasks. Children’s error patterns were examined to identify their use of non-target forms in relation to accuracy rates. The error patterns for the present simple and present progressive were straightforward. However, there were two dominant error patterns for the past
progressive (see Error Analysis below). This led us to explore whether there were groups of children who had distinctly different error forms. This led in turn to a two-step cluster analysis that was used to group children based on the accuracy of production of the target forms, and the error types. Cluster differences were confirmed by running a series of one-way ANOVAs. The clusters represented three developmental profiles that differed on characteristics of verb form accuracy and error types (see Error Analysis below). Finally, a repeated measures ANOVA was conducted to compare the degree of accuracy among the profiles on the three tests.

Results

Of the 70 children, the responses of 64 children were entered for analysis. Incomplete datasets or inaudible samples were not included (six children, aged 2;0, 2;3, 2;11, 3;0, 3;1, and 4;2 years).

Language sample measures

Mean Length of Utterance (MLU) in morphemes, Total Number of Utterances (TNU), Total Number of Words (TNW) and Number of Different Word Types (NDW) were generated from the language samples. A new measure, Bound Morpheme Type (BMT), was created to explore the possibility that high-performing children might use a greater variety of bound morpheme types. To minimize the effects of conversational and contextual factors on MLU scores (Johnston, 2001; Klee, 1992), an adjusted MLU (MLUadj) was derived by excluding one-word and imitated utterances (both self-imitation and imitation of the interlocutor). This measure was used in the analyses. Table 3 displays the descriptive statistics of the language sample measures.
Table 3. Measures derived from children’s language samples (N= 64)

<table>
<thead>
<tr>
<th>Measure</th>
<th>Range (Min- Max)</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (month)</td>
<td>28 (23- 51)</td>
<td>37.16</td>
<td>7.49</td>
</tr>
<tr>
<td>MLU (morpheme)</td>
<td>2.97 (2.12- 5.09)</td>
<td>3.73</td>
<td>.75</td>
</tr>
<tr>
<td>MLUadj (morpheme)</td>
<td>3.37 (3.53- 6.90)</td>
<td>5.03</td>
<td>.78</td>
</tr>
<tr>
<td>TNU</td>
<td>265 (61- 326)</td>
<td>189.47</td>
<td>60.19</td>
</tr>
<tr>
<td>TNW</td>
<td>806 (110- 916)</td>
<td>437.66</td>
<td>170.56</td>
</tr>
<tr>
<td>Number of word types</td>
<td>162 (33- 195)</td>
<td>110.34</td>
<td>35.88</td>
</tr>
<tr>
<td>TBM</td>
<td>591 (66- 657)</td>
<td>279.80</td>
<td>121.34</td>
</tr>
<tr>
<td>BMT</td>
<td>22 (12- 34)</td>
<td>24.81</td>
<td>5.59</td>
</tr>
</tbody>
</table>

MLU: Mean length of utterance; MLUadj: Adjusted mean length of utterance; TNU: Total number of utterances; TNW: Total number of words; TBM: Total bound morphemes; BMT: Bound morpheme types.

Test scores

The Present Simple form was used with the highest accuracy (mean = 88%, range = 6.25–100; SD = 19.96), followed by the Present Progressive (mean = 67%, range = 0–100, SD = 31.75) and the Past Progressive forms (mean = 44%, range = 0-100, SD = 41.23).

Children’s performance across the three tasks was statistically different (one-way repeated measures ANOVA, F (1.75, 110.38) = 53.14, p < .001, partial η² = .46).

Error analysis

Figures 1, 2, and 3 show children’s accuracy and their errors types on the Present Simple, the Present Progressive and the Past Progressive tasks respectively. Children are
arranged on the horizontal axis by their accuracy in the target form, and identified by their age in months on the X axis. The Y axis shows percentages of children’s responses that were correct, and specific types of error responses made, or ‘other responses’. Recall that the Present Simple is structurally the simplest form with no tense and aspect markers (see Table 2). From Figure 1 it is clear that the accuracy rate is remarkably high for this form. Forty-five percent of children (29 of 64) had no error in the task. Most errors consisted of verbless responses, unclear utterances, and no-responses (coded as ‘other’) and a few substitutions by the present perfect form occurred. Within these three morphological markers, the Present Simple developed first, and thus might be described as the earliest-developing form.

Figure 1. Present Simple accuracy and error patterns
Figure 2 reveals that the Present Progressive in Bangla was the second most accurate form; 56% of children (36 of 64) had 80% or greater success in this form. The dominant
substitute for this form was its Present Simple counterpart, and there were some ‘other’ types. Note that the gradually diminishing ‘other’ errors are not shown in figures 2 and 3 for the purpose of clarity. The Past Progressive responses suggest that this form was challenging for the children and often the target form was replaced by non-target forms. Only 34% children (22 of 64) achieved 80% accuracy or above. There were two dominant error patterns for the Past Progressive, used by two different groups of children. Children with little knowledge of the Past Progressive and the Present Progressive favoured the Present Simple or an ‘other response’ as the error pattern (Figure 3). As children began to master the Present Progressive form it became the preferred substitution for the Past Progressive form, rather than the Present Simple form. These two types of responses warranted further investigation (see below). Finally, when the Past Progressive form emerged and became frequent, there was an expected decline in the use of the Present Progressive form as a substitute.

Stages of acquisition

Error patterns were analysed further to identify potential stages of acquisition. Twenty-three children scored 0% accuracy for the Past Progressive form. A correlation analysis was conducted between children’s use of the Present Progressive form and their use of this form as a substitute for the Past Progressive form. There was a strong association between the two variables ($r_{(23)} = .79$, $p < .001$), i.e. once children mastered the Present Progressive, they tended to use it as a substitute for the Past Progressive instead of a simpler form (the Present Simple). This pattern of correct and incorrect use of morphological markers prompted the use of a cluster analysis to determine if there were reliable identifiable groups of children as we suspected. Further motivation for running a cluster analysis based on accuracy and error in children’s performance instead of their age came from the finding that
accuracy rates and age were only moderately correlated: age and Present Simple: r = .58; age and Present Progressive: r = .51; age with Past Progressive: r = .51. While these are all significant (p < .001), they only explain around a quarter of the variance.

Cluster analysis

The cluster analysis was run using SPSS (IBM Corp., 2011) with a set of six scores obtained from the three tasks: percentage accuracy in the Present Simple, Present Progressive and Past Progressive tasks, and percentage substitution of the Present Simple for Present Progressive, the Present Simple for Past Progressive, and the Present Progressive for Past Progressive. These scores were entered as grouping factors for determining clusters, which we will call Profile 1, Profile 2 and Profile 3 to indicate that the three Profiles differed in the characteristics of verb form accuracy and error types. Since age was not strongly correlated with children’s performance, it was not used in the cluster analysis. However, age is reported, along with a set of language sample measures, for each of the Profiles identified (Table 4). The output of the cluster analysis yields a qualitative judgement of goodness of fit of ‘poor’ to ‘good’ (IBM Corp., 2011), and the resulting clusters were described as ‘good’ in quality.

The analysis produced three clusters among which children were reasonably distributed; Profile 1, Profile 2 and Profile 3 consisted of 18 (28.13%), 19 (29.69%) and 27 (42.19%) children respectively. The mean percentage accuracy for target and error forms appear in Table 4. Profile 3 scored highest on the target forms, followed by Profile 2 and Profile 1. In Profile 3 the frequency of occurrence of error forms was low. In Profile 2, the error pattern for the Past Progressive form was commonly the Present Progressive form (74.21%). In Profile 1, the Present Simple was substituted for the Present Progressive and Past Progressive forms (47.78% and 20%). Children in Profile 1 were younger than those in
Profile 2 who were almost as old as those in Profile 3 ($F(2, 61) = 17.42$, $p < .001$, partial $\eta^2 = .37$; Profile 1: mean= 30.11 months, SD= 4.58; Profile 2: mean= 38.58, SD= 6.15; Profile 3: mean= 40.85, SD= 6.79).

Table 4. Overall performance profile of three profile types

<table>
<thead>
<tr>
<th>Language domains/ measures</th>
<th>Profile 1 (n = 18)</th>
<th>Profile 2 (n = 19)</th>
<th>Profile 3 (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Past Progressive (%)</td>
<td>6.11 (15.01)</td>
<td>15.79 (17.42)</td>
<td>88.15 (12.72)</td>
</tr>
<tr>
<td>Present Progressive (%)</td>
<td>28.33 (25.26)</td>
<td>84.74 (13.07)</td>
<td>80.74 (21.11)</td>
</tr>
<tr>
<td>Present Simple (%)</td>
<td>69.60 (28.31)</td>
<td>92.59 (10.67)</td>
<td>97.83 (3.57)</td>
</tr>
<tr>
<td>Present Simple for Present Simple (%)</td>
<td>51.67 (25.50)</td>
<td>8.95 (10.49)</td>
<td>12.59 (20.11)</td>
</tr>
<tr>
<td>Present Simple for Present Progressive (%)</td>
<td>47.78 (25.10)</td>
<td>4.74 (8.41)</td>
<td>.74 (3.85)</td>
</tr>
<tr>
<td>Present Progressive for Past Progressive (%)</td>
<td>20 (17.82)</td>
<td>74.21 (17.74)</td>
<td>5.19 (7.53)</td>
</tr>
<tr>
<td>Past Progressive (%)</td>
<td>4.33 (.52)</td>
<td>5.08 (.47)</td>
<td>5.46 (.78)</td>
</tr>
<tr>
<td>Bound Morpheme Type*</td>
<td>19 (4.50)</td>
<td>26.05 (3.52)</td>
<td>27.81 (4.48)</td>
</tr>
<tr>
<td>Word Type*</td>
<td>77.44 (25.09)</td>
<td>120.95 (24.93)</td>
<td>124.81 (35.06)</td>
</tr>
<tr>
<td>Age (month)*</td>
<td>30.11 (4.58)</td>
<td>38.58 (6.15)</td>
<td>40.85 (6.79)</td>
</tr>
</tbody>
</table>
*These measures are presented for the purpose of describing the clusters; they were not determinants of them.

Profile differences. A set of one-way ANOVAs with the test scores, the substitution scores, Age and three language sample measures [MLUadj, Bound Morpheme Type (BMT), and Number of Different Words (Word Type)] entered as dependent variables and the Profile number from the cluster analysis entered as the independent variable were run. To identify which Profiles differed from one another on the measures, a post-hoc test appropriate for groups of different sizes (Field, 2009), Hochberg’s GT2, was run. Kruskal-Wallis test values were reported when the assumption of homogeneity was violated. The results appear in Table 5.
Table 5. Post-hoc results of profile differences

<table>
<thead>
<tr>
<th>Variables</th>
<th>Profile differences</th>
<th>Profiles</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>PS</td>
<td>$1 &lt; 2 = 3^a$</td>
<td>$1 &amp; 2$</td>
<td>$r = .53^{**}$</td>
<td>$r = .68^{***}$</td>
<td>$r = .33$</td>
</tr>
<tr>
<td>PresProg</td>
<td>$1 &lt; 2 = 3^a$</td>
<td>$1 &amp; 3$</td>
<td>$r = .81^{***}$</td>
<td>$r = .74^{***}$</td>
<td>$r = .03$</td>
</tr>
<tr>
<td>PastProg</td>
<td>$1 = 2 &lt; 3^a$</td>
<td>$2 &amp; 3$</td>
<td>$r = .37$</td>
<td>$r = .87^{***}$</td>
<td>$r = .86^{***}$</td>
</tr>
<tr>
<td>PS for PresProg</td>
<td>$1 &gt; 2 = 3^a$</td>
<td>$1 &amp; 3$</td>
<td>$r = -.73^{***}$</td>
<td>$r = -.64^{***}$</td>
<td>$r = -.04$</td>
</tr>
<tr>
<td>PS for PastProg</td>
<td>$1 &gt; 2 = 3^a$</td>
<td>$1 &amp; 3$</td>
<td>$r = -.78^{*}$</td>
<td>$r = -.89^{*}$</td>
<td>$r = -.37$</td>
</tr>
<tr>
<td>PresProg for PastProg</td>
<td>$1 = 3 &lt; 2^a$</td>
<td>$2 &amp; 3$</td>
<td>$r = .83^{***}$</td>
<td>$r = -.48$</td>
<td>$r = -.87^{***}$</td>
</tr>
</tbody>
</table>

PastProg

<table>
<thead>
<tr>
<th>Variables</th>
<th>Profile differences</th>
<th>Profiles</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MLUadj</td>
<td>$1 &lt; 2 = 3^b$</td>
<td>$1 &amp; 2$</td>
<td>$d = -1.51^{**}$</td>
<td>$d = -1.70^{***}$</td>
<td>$d = -.59$</td>
</tr>
<tr>
<td>Bound</td>
<td>$1 &lt; 2 = 3^b$</td>
<td>$1 &amp; 2$</td>
<td>$d = -1.75^{***}$</td>
<td>$d = -1.96^{***}$</td>
<td>$d = -.44$</td>
</tr>
</tbody>
</table>

Morpheme Type

<table>
<thead>
<tr>
<th>Variables</th>
<th>Profile differences</th>
<th>Profiles</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
<th>Effect sizes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Type</td>
<td>$1 &lt; 2 = 3^b$</td>
<td>$1 &amp; 2$</td>
<td>$d = -1.74^{***}$</td>
<td>$d = -1.55^{***}$</td>
<td>$d = -.13$</td>
</tr>
<tr>
<td>Age (month)</td>
<td>$1 &lt; 2 = 3^b$</td>
<td>$1 &amp; 2$</td>
<td>$d = -1.56^{***}$</td>
<td>$d = -1.85^{***}$</td>
<td>$d = -.35$</td>
</tr>
</tbody>
</table>

* $p < .05$, ** $p < .01$, *** $p < .001$.

$^a$ analysis using Kruskal-Wallis test.

$^b$ analysis using ANOVA (Hochberg’s GT2).

Profile 2 and Profile 3 were similar in their performance on the Present Simple and Present Progressive forms and they were both significantly better than Profile 1 ($H_{(2)} = 24.12$, $p < .001$; $H_{(2)} = 31.61$, $p < .001$ for Present Simple and Present Progressive respectively). On
the Past Progressive, children with Profile 1 and Profile 2 performed similarly but Profile 3 performed significantly higher than both ($F_{(2,61)} = 211.44, p < .001$). The Kruskal-Wallis test results of the errors in the three elicitation tasks showed that the use of the Present Simple form for the Present Progressive was significantly higher for Profile 1; Profile 2 and Profile 3 were similar in their use ($H_{(2)} = 25.31, p < .001$). Similarly, Profile 1 used the Present Simple form significantly more often in place of the Past Progressive form than Profile 2 and Profile 3 who were found to be similar ($H_{(2)} = 44.45, p < .001$). In the use of Present Progressive for Past Progressive, Profile 1 and Profile 3 had similar scores, but Profile 2 used this form significantly more often than both ($H_{(2)} = 44.91, p < .001$).

There was no significant difference in age for Profile 2 and 3, whereas Profile 1 was significantly younger than both ($F_{(2,61)} = 17.74, p < .001$). Similarly, for the language sample variables, Profile 2 and Profile 3 had significantly higher scores than Profile 1 but the difference between Profile 2 and Profile 3 was not significant (MLUadj: $F_{(2,61)} = 17.01, p < .001$; BMT: $F_{(2,61)} = 24.67, p < .001$; Word Type: $F_{(2,61)} = 15.45, p < .001$).

**Repeated measures ANOVA**

A 3X3 repeated measures ANOVA was run with three Profiles and verb form scores in order to examine any difference in the levels of mastery achieved in the forms within the Profiles. In case of violation of sphericity in the data, Greenhouse-Geisser’s correction was used and Hochberg’s GT2 was used for post-hoc analyses. There was a significant main effect of verb forms, $F_{(2,122)} = 153.31, p < .001$, partial $\eta^2 = .72$, and the interaction effect between verb forms and Profiles was also significant ($F_{(4,122)} = 44.41, p < .001$, partial $\eta^2 = .59$). Within each Profile the accuracy rates in the three forms were significantly different.
For Profile 1, children’s Present Simple score was significantly higher than both the Present Progressive (mean difference = 41.23, p < .001, confidence interval for the difference = 22.56 - 59.98) and the Past Progressive (mean difference = 63.49, p < .001, confidence interval for the difference = 49.59 - 77.39) scores and the Present Progressive was significantly higher than the Past Progressive (mean difference = 22.22, p = .003, confidence interval for the difference = 35.82 – 8.62).

For Profile 2, the Present Simple scores were significantly higher than their Present Progressive (mean difference = 7.85, p = .02, confidence interval for the difference = 1.64 – 14.06) and their Past Progressive (mean difference = 76.80, p < .001, confidence interval for the difference = 68.55 – 85.05) scores. There was also a significant difference between their Present Progressive and their Past Progressive scores (mean difference = 68.95, p < .001, confidence interval for the difference = 61.96- 75.93).

Profile 3 performed significantly higher in Present Simple than in Present Progressive (mean difference = 17.09, p < .001, confidence interval for the difference = 9.10 - 25.07) and in Past Progressive (mean difference = 9.68, p = .001, confidence interval for the difference = 4.39 - 14.06). However, their accuracy scores in the Present Progressive and in the Past Progressive forms were not significantly different (mean difference= 7.41, p = .13), confidence interval for the difference = -2.26- 17.07).

There was an overall difference in the success rates of the forms across the Profiles (Present Simple > Present Progressive > Past progressive). However, with increasing proficiency children arrived at a stage where they performed similarly on the forms: Profile 3 had similar success rates in the Present Progressive and the Past Progressive forms. However, they had not yet mastered these forms to the accuracy level of the Present Simple form.

Discussion
The intention was to discuss the findings in the order of the research questions and analysis, but as it transpired, the relationship between the accuracy of the verb forms, and the types of errors made, required them to be treated together. Therefore, after discussing the complexity/difficulty of the verb forms we turn to a discussion of three Profiles that emerged in consideration of this relationship.

Accuracy of the forms

The performance of the typically-developing children on the three tasks showed that the Present Simple form was acquired with the highest accuracy (88%) followed by the Present Progressive (67%) and the Past Progressive scores (44%) respectively. A repeated measures ANOVA revealed a significant difference in the degree of accuracy among the markers across children. There could be more than one explanation for this order of accuracy.

First, the Present Simple form is the simplest in terms of grammatical complexity, while the Present Progressive and the Past Progressive forms carry more elaborate marking. Therefore, there is likely to be some merit in an interpretation based on structural complexity. Both tense and aspect markers have been discussed in a range of studies as having posed challenges for children (e.g. Leonard, Eyer, Bedore, & Grela, 1997). The absence of tense or aspect markers in the Present Simple form, and the absence of the tense marker in the Present Progressive form lead to a tenable interpretation of the order of acquisition of the verb forms. Second, children’s performance on language tasks or their production of language forms in spontaneous language samples are commonly found to be regulated by the cognitive complexities of the forms (Aksu, 1978). The task to elicit the Past Progressive forms had certain cognitive prerequisites such as conceptualising an ongoing event in past time, comprehending the brief story preceding each stimulus, and making inferences. This is possibly one of the reasons why children often replaced this form with substitutes that were
less taxing on them. Third, children’s use of the markers is likely to be a reflection of how they are used in the language environment, given the grammatical and contextual scope of the markers (Hadley, Rispoli, Fitzgerald, & Bahnsen, 2011; Xanthos et al., 2011). The Present Simple form is pervasive in Bangla due to its frequent use in other linguistic contexts; it is permissible in Bangla to use the Present Simple form in the Present Progressive contexts. However, it is not clear at this stage whether the early mastery of the Present Simple form is due to its structural simplicity or the rich exposure to this form. Systematic data on the children’s language environment is needed in order to tease apart the contributions of the two factors.

Although the accuracy scores revealed a clear difference in success of use across children’s performance on the three tasks, it is possible that the structural and contextual properties of the forms cease to influence children’s performance once a certain level of mastery is achieved. This suggestion is motivated by the finding that no significant difference between scores was found for the most proficient profile, i.e. Profile 3, in the Present Progressive and the Past Progressive tasks. However, even in Profile 3, the Present Simple accuracy rate was still significantly higher than the other two forms. This is likely to be caused by the use opportunities of the Present Simple form, as described above. It is possible that with older children the same tasks would not reveal any difference in performance; all children would have very high accuracy rates across the forms.

On the errors

Errors were near-misses. Investigations of non-target forms across languages revealed diverse error patterns which were primarily regulated by the typological properties of the languages. Results from Bangla, a null-subject language with fairly rich morphological properties, finds converging evidence by identifying that the Present Progressive and the Past
Progressive forms were often substituted by the Present Simple and the Present Progressive forms respectively. Both substitutes deviated from the target forms by one dimension only (near-misses); the former left out the aspect marker, whereas the latter included the aspect marker but omitted the tense marker.

However, a considerable number of children (mainly in Profile 1) supplied Present Simple forms in Past Progressive contexts. Since the Present Simple form differs from the Past Progressive form on two dimensions: aspect and tense, these deviations cannot be termed ‘near-misses’. It is possible that the idea of the ‘near-misses’ is, in fact, an issue of degree. This proposition is consistent with the three Profiles identified. Profile 1 with the ‘least developed’ children used the Present Simple form, whereas Profile 2 used the Present Progressive form for the Past Progressive form. It seems justified to expect that very young children may not have the cognitive facilities to produce complex forms like the Present or the Past Progressive forms. Therefore, a stage where children make compromises in more than one feature appears tenable. However, this pattern of gradually approximating the target forms has not been reported in other rich null-subject languages. This could be attributed to the methodological differences across the studies; probably the appropriate age group has not been explored with the appropriate verb forms; for example, had this study not involved the Past Progressive forms, most errors would have appeared as near-misses. Confirmation regarding whether the gradual approximation property is unique to Bangla data or whether this can also be found in similar languages requires crosslinguistic examination of the child data using the same methodological specifications.

Also, substitutions of the Past Progressive form by the Present Simple may not always be errors on two dimensions. Note that Bangla allows for using the Present Simple form for the Present Progressive, at least in the context of the present study. Therefore, it is likely that
there was a sub-group of children whose substitution of the Present Progressive by the Present Simple form was not necessarily due to their inability to produce the target form. This leads to the possibility that an amount of the Present Simple use is functionally equivalent to the Present Progressive and hence, for these children, substitutions of the Past Progressive by the Present Simple forms still qualify as near-misses. Another view to interpreting this particular substitution pattern, i.e. use of the Present Simple for the Past Progressive form, emerges from the suggestion that the retrieval of the form produced is primarily driven by the similarity in features, and only high frequency forms may change the near-miss patterns due to their strength in the paradigm (morphological richness hypothesis: e.g. Lukács, Leonard, Kas, & Pléh, 2009). As reported previously, the Present Simple form in Bangla seems to have an exceptionally wide scope which makes it likely to also be highly frequent in the language. This interpretation of the substitution patterns is still consistent with the claims associated with near-misses.

*Are errors substitutions or omissions?* The morphological analysis of Bangla verb forms indicates that the verb conjugation system is incremental. Because of the incremental system, the omission of a tense or aspect marker results in another legal verb form in Bangla. Therefore, these deviations can be interpreted both as omissions of the marker, or substitutions by a structurally simpler form. However, the errors appeared more like substitutions than omissions, since the markers were never omitted to the extent of bare forms. Every erroneous form was a legal, inflected form that had attached fewer than the appropriate number of inflections to the stem. Crosslinguistic evidence of substitution by simpler counterparts of the target forms are common (e.g. Pizzuto & Caselli, 1993).

Irrespective of whether Bangla non-target forms found in the present study resulted from omission or substitution, it can be concluded that the linguistic boundaries were maintained
in errors. Bangla-speaking children examined in this study never omitted the markers to the extent of producing an ‘illegal’ form, while the use of bare forms among young children is ubiquitous in languages that permit such use (e.g. Poeppel & Wexler, 1993, German; Wexler, 1994, English).

Hierarchy among markers

In some languages a direct assessment of tense and aspect markers is difficult because in those languages often both markers are encoded in one form, e.g. past –ed in English. Bangla verbs are marked independently for tense and aspect which makes it ideal for testing these markers. In Bangla, any combination from no tense, and no aspect markers to the presence of both is legal; for example, absence of both is typically observed in the Present Simple form, whereas both markers appear in the Past Progressive form. In addition, points between these two ends can be found in the Present Progressive or Present Perfect form (aspect without tense marker), and in the Past Simple form (tense without aspect marker). Some crosslinguistic studies of typical and atypical development have revealed an emergence of grammatical aspect markers prior to tense markers and identified tense to be the challenging area (Leonard, Sabbadini, Leonard, & Volterra, 1987; Rice, Noll, & Grimm, 1997), whereas some reported difficulties in aspect marking itself (Fletcher, Leonard, Stokes, & Wong, 2005; Leonard, Lukács, & Kas, 2012).

Findings from the present study showed that in Bangla, children’s errors contained deviation in both tense and aspect marking. While producing the Past Progressive forms that required both tense and aspect markers (also person markers), the youngest and the least developed group (Profile 1) often failed to mark both which resulted in substitution by the Present Simple form. However, children’s preferences were evident from the substitution of the Past Progressive form by the Present Progressive form. Children’s production of the
correct aspect marker along with the absence of the tense marker appears consistent with the ‘grammatical aspect before tense’ proposal suggested previously. This seems tenable in Bangla also because aspect markers are more prevalent than tense markers in Bangla. The past tense marker is realized in the Past Simple, Past Progressive and Past Perfective forms, where, except for the Past Simple, the other two past forms also take aspect markers. Moreover, the scope of the Past Simple form is restricted and it is often replaced by a more prevalent form in Bangla, i.e. the Present Perfect. The Present Progressive and the Present Perfect forms, interestingly, take aspect markers without a tense marker. Therefore, the distributional patterns of tense and aspect markers in Bangla support the possible emergence of aspect markers before the tense marker in Bangla. A similar interpretation was forwarded for Hebrew-speaking children’s use of the tense marker prior to the person marker (Armon-Lotem & Berman, 2003). Another reason why aspect markers in Bangla are likely to appear more important to children is that Bangla aspect markers are often syllabic and therefore, salient in the language. On the other hand, tense markers in Bangla, i.e. -l- (past) and -b- (future, not discussed here), are non-syllabic. Their lower perceptual salience is likely to contribute towards later emergence of the markers (Leonard, 2014a).

Another perspective for examining the status of these two markers is to assume that both aspect and tense markers have equal weights and they are at the same difficulty level for Bangla-speaking children, and therefore, what children eventually produce in their errors is simply a mismatch vis-à-vis one or more features. However, even within this view, there is room for arguing that aspect markers in Bangla are earlier in the developmental sequence than tense markers. If there were no priorities among the markers, then children, at least sometimes, would produce the Past Simple form (that contains a tense marker without an aspect marker) as a substitute for the Past Progressive form. But this was not the case; when
the target form deviated by one marker, the error was typically on the tense marking. Therefore, based on Bangla-speaking children’s clear preference for aspect markers, it seems tenable to posit that the scope of a morphological marker, as regulated by the language, significantly dictates the emergence and status of that marker.

*Three profile/stages*

Based on the three Profiles identified, the following order of acquisition for the three forms is proposed: Present Simple -> Present Progressive -> Past Progressive. This order is convincing given the linguistic and contextual properties associated with the forms. Also, as discussed earlier, the Present Simple form is hypothesised to have an enhanced representation in children’s cognition because of its exceptionally wide linguistic scope. On the other hand, due to the cognitive demands and the restricted scope of the Past Progressive form, it naturally has limited linguistic opportunities to occur.

The error patterns across the Profiles present interesting opportunities for insights into language development. When children were least developed (Profile 1), the use of the Present Simple form was extensive as both target and non-target forms. Evidence of children’s channeling known forms towards other contexts, is commonly found in the literature (e.g. Lakshmanan, 2006). The groups also showed gradual shifts across the error patterns. For instance, after the mastery of the Present Progressive form, which presumably is available at a higher developmental stage, they began to use the newly mastered form as a substitute for the Past Progressive. The shift in the error pattern suggested that children’s errors became more sophisticated with advancement in mastery of the forms.

However, the Profiles were not necessarily distinct when the distribution of the language sample measures was examined. Profile 1 was significantly younger, and less advanced in terms of MLUadj, Word Type and Bound Morpheme Type, while these scores
were not significantly different between Profile 2 and Profile 3. Therefore, it appears that the only indices that differentiated these two Profiles were children’s accuracy of the Past Progressive form and an associated index, i.e. their substitution of the Past Progressive by the Present Progressive form. Note that the measures that did not bring out the differences between Profile 2 and 3 tend to plateau after a certain point in development (Fletcher, Leung, Stokes, & Weizman, 2000). One possibility could be that growth slows down after stage 2 (Profile 2) with regard to these indices. However, note that the two groups are also similar in age. If the previous explanation was justified, then Profile 3 would contain children who were considerably older than Profile 2, indicating an ongoing chronological progress. This points towards a possibility that within the same age range, some children are able to accomplish certain linguistic milestones while others cannot. Children of the same age with similar scores in a set of other linguistic indices may differ in their mastery of the Past progressive form, which is arguably at an exceptionally high level of difficulty. Therefore, it is likely that Profile 3 consists of children who have advanced language ability relative to their peers.

Conclusion

The developmental trajectory for verb inflection in Bangla accords with previous research on languages with agglutinating characteristics. A consideration of accuracy rates and errors made reveals an orderly accretion of verb forms of increasing complexity, accompanied by systematic substitution strategies based on the child’s current competence. The study lays the groundwork for further language acquisition research in Bangla-speaking contexts for both typically and atypically developing children. Although further research is required, the results could be used to determine possible intervention goals that are developmentally sound. Future investigations should explore the development of Bangla person markers in a range of
linguistic contexts as well as the contribution of external factors such as frequency in parental input on determining the developmental sequence of Bangla verb forms.
References


models of language development. *Journal of Child Language, 19*(3), 491-557. doi:10.1017/S0305000900011557


Appendix

Sample picture used to elicit the Present Progressive form
A picture pair used in the Past Progressive task