Compositionality in Bangla Compound Verbs and their Processing in the Mental Lexicon

Tirthankar Dasgupta¹
TCS Innovation Lab,
New Delhi
iamirthankar@gmail.com

Manjira Sinha¹
Xerox Research Center India, Bangalore
Manjira87@gmail.com

Anupam Basu
Department of Computer Science and Engineering,
IIT Kharagpur
anupambas@gmail.com

Abstract

We conduct a cross-modal priming experiment to determine the mental representation and access strategies for compound verbs (CV) in Bangla. Analysis of reaction time indicates that compositionality among CVs triggers priming effects for both the constituent verbs. On the other hand non-compositional CVs exhibit priming only for the polar verb. Thus, compositional CVs are decomposed into their constituent verbs during processing. On the other hand, non-compositional verb phrases are represented and accessed as a whole in the minds of a Bangla speaker. The reaction time data thus collected are used to evaluate our vector space model for compositionality judgment.

1 Introduction

The term compositionality refers to the idea that the meaning of a complex expression is derived from (i) its structure and (ii) the meanings of its constituents (refer to (Fodor and Pylyshyn, 1988) for details). Compound verbs (henceforth CV) in Bangla are known to exhibit continuum of compositionality. Thus, some of them are compositional (khete bashA “sit to eat”), some are non-compositional (jAliye mArA “to irritate”) and some in between (ghure berAno “loitering”).

There is a considerable linguistic debate on whether CVs are considered as two distinct words connected by some syntagmatic rule or whether they form a single lexical unit (Paul, 2010; Chakrabarti et al., 2008; Butt, 1995) and whether semantic compositionality plays any role deciding the above. These linguistically and computationally challenging issues play an important role in developing lexical resources like WordNet (Fellbaum, 2010).

Since, none of the linguistic arguments and computational approaches has so far led to any consensus, we here for the first time performed a series of psycholinguistic experiments on Bangla compound verbs to study and model how the mental lexicon (ML), organize and process such complex expressions. The term mental lexicon refers to the storage of words in the human mind (Aitchison, 2005). A clear understanding of the structure and processing mechanism of CVs in the ML may provide us insight about how to represent and process CVs in computational lexicons. Further, the reaction time of the subjects for recognizing various lexical items under appropriate conditioning may lead us to develop more robust computational models of automatic identification of CVs.

A plethora of works have been done to explore the representation and processing of words in the mental lexicon (refer to Seidenberg, 2012 to get a detail account of it). Attempts have also been made to study the processing of Bangla derivationally suffixed morphologically complex words (Dasgupta et al., 2010; Dasgupta et al., 2012, 2014) and Bangla compound words (Dasgupta et al., 2015). However, to the best of our knowledge

¹ Work done during the tenure at IIT Kharagpur.
no such prior attempts have been made to psycho-
linguistically analyze the representation and pro-
cessing of the Bangla compound verbs.

Based on the above discussion, the objective of
this paper is to explore the semantic composition-
ality in Bangla CVs and its role in the possible
organization and processing of CVs in the ML.
Accordingly, we first use empirical techniques to
collect user judgment of compositionality for
Bangla CVs. Next, we perform the cross modal
priming experiment to understand the processing
of CVs ML. We found that Bangla CV exhibit
continuum of compositionality. Highly composi-
tional verb phrases $V_1V_2$ trigger significant prim-
ing effect for both its constituent verbs $V_1$ and $V_2$.
On the other hand, non-compositional verb se-
quences exhibit priming effect only for $V_1$. These
observations together imply that the mental lex-
icon decomposes compositional verb sequences
into their constituent verbs and recognize sepa-
rateley during processing. On the other hand non-
compositional verb sequences are organized and
accessed as a whole. Based on the dataset col-
lected from the above experiments, we further de-
velop a vector space model to compute semantic
compositionality in a verb sequence. The pre-
dicted values are compared with the human judg-
ment scores and the priming experiment results
where a significant correlation is found.

2 Compositionality Judgments

We could not find any Bangla CV dataset publicly
available and annotated in terms of composi-
tionality. Thus, we choose 300 verb sequences and
presented them to 36 native Bangla speakers.
Similar to the work discussed in (Reddy et al.,
2011), we request each subject to give composi-
tionality score of the verb sequences by asking a)
how literal the verb phrase is in a given context and b)
how literal the constituent verbs are. Sub-
jects were asked to rate the compositionality un-
der 1-10 point scale, 10 being highly composi-
tional and 1 for non-compositional. In order to
validate the user annotated data, we compute the
inter-annotator agreement using the Fleiss Kappa
measure (Fleiss et al., 1981). We found an agree-
ment of $\kappa = 0.68$.

3 The Priming Experiment

Priming effects are observed because of the way
our brain is supposed to function. Presentation of
a stimulus (say a word $P$) activates a group of neu-
rons (often called a functional web) in the brain.
When another stimulus (say word $T$) is then pre-
sented to the individual within a short duration,
the recognition of $T$ is faster if the functional web
of $T$ shares a lot of neurons with that of $P$. This
fast reaction time to recognize a stimulus due to
the prior exposure to a functionally related stimu-
lus is known as priming (Tulving et al., 1982).
Thus, through priming experiments, we can iden-
tify word pairs whose functional webs have a
stronger overlap, which in turn reveals how brain
organizes the words in the mental lexicon (See
Pulvermüller, 2003; Spivey et al., 2012) for a de-
tailed account of such phenomena and different
types of priming techniques.

Experiment design: We conduct the cross-
modal repetition priming experiments as de-
scribed in (Marslen-Wilson et al., 1994). Here,
a subject gets an audio stimulus (called prime ($P$))
of a verb sequence $V_1V_2$ and immediately at the
offset of the spoken prime, gets a visual represen-
tation of the target ($T$) word $V_1$. The same audio
stimulus is presented to the subject after a random
number of trials with the target $V_2$. Based on the
auditory prime and the visual target probe, the
subjects are asked to decide whether the visually
presented target $V_1$ or $V_2$ is a valid word in the
language. The above experiment is also repeated
with the same target words but with different au-
ditory inputs called the controls($C$). The control
verb sequences do not have any relatedness with
the targets. For example, “khaye felo” (complete
eating) and “khAOyA” (to eat) or “khaye felo” and
“felo” (to drop) are $P$-$T$ pairs, for which the
the corresponding $C$-$T$ pair could be “niye jAO” (take
away) and “khAOyA” (to eat). The time taken by
a subject to complete the lexical decision task af-
after the visual presentation of the target is defined
as the response time (RT). For both the experi-
ments, RTs of all the targets were recorded and
analyzed to observe priming effects.

Materials: The 300 verb pairs as discussed in
section 2, are grouped into two classes highly
compositional (C1) and non-compositional (C2).
We randomly choose 30 $V_1V_2$ pairs from each
class as primes. For each prime, we have two tar-
gets $V_1$ and $V_2$, which makes 120 $P$-$T$ pairs. Con-
sequently, 120 C-$T$ and 240 fillers (or non-words)
have been chosen to restrict the subjects to apply
any strategic guess during the experiment. Over-
all, we have collected RTs for 480 word pairs in
one experiment. A set of 10 $P$-$T$ and $C$-$T$ pairs
were also collected for a practice session before
the true experiments. However, the RTs of these
practice pairs are not included in any analysis.
The experiment was conducted using the DMDX software tool that plays the auditory stimulus and then showed the visual probe for 200ms. Corresponding to each visual probe, subjects had 3000ms to perform the lexical decision after which the system presents the next audio stimulus. The subject performs the decision task by pressing either the “K” key for a VALID word and “S” for INVALID word. The system automatically records the RT, which in this case is the time between the onset of the visual probe and clicking of one of the buttons by the subject. The experiments were conducted on 36 native Bangla speakers; 29 of them have a graduate degree and 7 hold a post graduate degree. The age of the subjects varies between 25 to 35 years.

Results: The RTs with extreme values (>2000ms) and those for incorrect lexical decisions (about 3.2%) were excluded from the data. The raw RTs for all correct responses were inversely transformed (Ratcliff, 1993) and entered into a mixed-design analysis of variance with two factors: priming (primed and unprimed), and condition (C1, C2). In the subject’s analysis (F1), condition and priming were treated as repeated measures. In the items analysis (F2), priming is treated as repeated factor and condition as independent factor. Table 1 summarizes the mean RTs for the prime (P) and control (C) sets of the V1 and V2 for the two classes.

<table>
<thead>
<tr>
<th>Class</th>
<th>Example</th>
<th>Avg Comp</th>
<th>Avg RT V1(P,C)</th>
<th>Avg RT V2(P,C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>khete basA</td>
<td>8.8</td>
<td>666,688</td>
<td>661, 696</td>
</tr>
<tr>
<td></td>
<td>(sit to eat)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>khaye felA</td>
<td>2.3</td>
<td>657,679</td>
<td>687, 659</td>
</tr>
<tr>
<td></td>
<td>(to eat)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Average compositionality and RTs for the different word classes

There was a strong main effect of priming with faster RTs to primed (667ms for V1 and 687ms for V2) than unprimed (679ms and 659ms) targets, $F(1, 36)=47.60$, $p<0.01$; $F(1, 60)=26.00$, $p<0.01$. There was a main effect of condition $F(1, 30)=11.69$, $p<0.01$, $F(1, 60)=3.51$, $p<0.01$, and a significant condition by priming interaction, indicating that priming effects varied across conditions. Thus, we have observed that when V1 is presented as target, significant priming occurred for words belonging to both C1 and C2. These results are statistically significant according to the 2-way ANOVA test. On the other hand, when V2 is presented as target with the same prime stimulus V1V2, priming is observed only in C2. This result indicates that, compositional verb sequences exhibits priming for both the constituent verbs V1 and V2 whereas, non-compositional verb sequences (C1) exhibit priming only with V1. This may be accounted for by the fact that non-compositional verb sequences derive their meaning mainly from the meaning of its first constituent verb V1, thus, priming occurs only with V1. On the other hand, as semantic compositionality between verb sequences increases, both the constituent verbs (V1 and V2) plays role in determining the meaning of the whole expression. Thus, both the constituent verbs exhibit priming behavior when preceded by the prime V1+V2. The above observations together imply that:

**Compositional verb sequences are in general represented in terms of their constituent verbs in the mental lexicon; lexical access and comprehension is facilitated through decomposition of the surface forms into the corresponding constituent verbs. On the other hand non compositional verb sequences are represented and accessed as a whole.**

4 Computational Model

Based on the collected data, we now try to develop a vector space model of semantic compositionality to predict the organization and processing of verb phrases in the mental lexicon. We evaluate our model with the data collected from our human judgment compositionality score and the cross-modal priming experiments.

In our vector space model we represent a word’s meaning in an n-dimensional space. Here, meaning of individual words is represented in terms of its co-occurrences observed in a large corpus. These co-occurrences are stored in a vector format acquired from a corpus following the different procedures as suggested in the literature (refer to (Mitchell and Lapata, 2008) for a detailed overview).

We have considered the lemmatized context words around the target word in a window of size 200 as the co-occurrences. For the purpose of lemmatization, we have used a Bangla morphological analyzer having an accuracy of around
95%. A Bangla corpus of 33 million words is already available. The corpus contains documents from different domains like literature, tourism, news and personal blogs. The top 10000 frequent content words from the corpus are used for the feature co-occurrences. To measure similarity between two vectors, cosine similarity (sim) is used. We have used the raw co-occurrence frequency as the values of the constructed vector.

Given a CV $W_3$ with the constituents $W_1$ and $W_2$, the compositionality score $S_3$ of $W_3$ is computed as $S_3 = f(S_1, S_2)$. Where, $S_1$ and $S_2$ are the literality scores of $W_1$ and $W_2$ respectively and $f$ is the semantic compositionality function defined in Column 1 of Table 1. $S_1$ and $S_2$ are computed as

$$S_1 = \text{sim}(v_1, v_3)$$
$$S_2 = \text{sim}(v_2, v_3).$$

Where, $v_1$, $v_2$ and $v_3$ be the co-occurrence vectors of $W_1$, $W_2$ and $W_3$ respectively and $\text{sim}$ is the cosine similarity between two vectors computed as:

$$\text{sim}(v_1, v_2) = \frac{v_1 \cdot v_2}{||v_1|| ||v_2||}$$

The primary idea behind the constituent based compositionality model is the fact that if a constituent word is used literally in a given verb sequence it is possible that the verb sequence and its constituent share common co-occurrences.

We compare the compositionality score $S_3$ of all the five composition functions namely, $f1$, $f2$, $f3$, $f4$, and $f5$ with the human annotated compositionality score. The constant parameters $(\alpha, \beta, \gamma)$ for all the five models have been computed using list square linear regression technique. We perform a 6 folded cross validation over the test data. The performance of the individual models is reported in Table 2 below. We can observe that the composition functions $f1$ and $f5$ are better predictor of phrase based compositionality than the other models.

<table>
<thead>
<tr>
<th>#</th>
<th>$F$</th>
<th>$\rho$, $R^2$</th>
<th>$(\alpha, \beta, \gamma)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\alpha v_1 + \beta v_2 = v_3$</td>
<td>0.73, 0.80</td>
<td>0.02, 0.40</td>
</tr>
<tr>
<td>2</td>
<td>$\gamma v_1 v_2 = v_3$</td>
<td>0.40, 0.71</td>
<td>0.32</td>
</tr>
<tr>
<td>3</td>
<td>$\alpha v_1 + \beta v_2 + \gamma (v_1, v_2) = v_3$</td>
<td>0.43, 0.77</td>
<td>0.01, 0.41, 0.33</td>
</tr>
<tr>
<td>4</td>
<td>$\alpha v_1 = v_3$</td>
<td>0.30, 0.55</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Table 2: Correlation between the human judgment compositionality and the predicted $S3$ by each of the composition function along with the goodness of fit.

In the final phase of our work, we have compared the phrase level compositionality score $S_3$ and the literality scores $S_1$ and $S_2$ with the priming behavior of targets $V_1$ and $V_2$. The correlation results are reported in Table 3. We observe $S_1$ has got a high correlation with the priming behavior of the second constituent verb $V_2$. On the other hand, weak correlation exists between $S_1$ and $V_1$. The observations are quite expected, as priming behavior in $V_1$ is observed irrespective of the semantic compositionality of the phrase $V_1V_2$.

![Figure 1](image1.png)  
Figure 1: Change in degree of priming of $V_1$ with respect to the semantic distance of $V_1$ and $V_1+V_2$

![Figure 2](image2.png)  
Figure 2: Change in degree of priming of $V_2$ with respect to the semantic distance of $V_2$ and $V_1+V_2$

---

2 www.cel.iitkgp.ernet.in
whereas, priming behavior in \( V_2 \) changes as compositionality in \( V_1V_2 \) changes. This is also reflected in Figure 1 and Figure 2 respectively.

Similar observations are found for \( S_1 \) and \( S_2 \). Overall, the results in Table 3 are in agreement with that of Table 1: compositional verb phrases exhibit priming for both its constituents whereas non-compositional verb phrases shows priming only to its first constituent verb (\( V_1 \)).

<table>
<thead>
<tr>
<th></th>
<th>( V_1 )</th>
<th>( V_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( S_3 )</td>
<td>0.04</td>
<td>0.78</td>
</tr>
<tr>
<td>( S_1 )</td>
<td>0.73</td>
<td>0.01</td>
</tr>
<tr>
<td>( S_2 )</td>
<td>0.12</td>
<td>0.74</td>
</tr>
</tbody>
</table>

Table 3: Correlation between priming in \( V_1 \) and \( V_2 \) and the computed cosine similarity scores (\( S_1, S_2 \)).

5 General Discussion and Conclusion

In this paper we have presented some preliminary psycholinguistic experiments to identify the role of compositionality in the representation and processing of Bangla compound verbs in the mental lexicon. In order to do so, we first computed the user judgment of compositionality score. We found that there is a continuum of compositionality between Bangla compound verbs. Next, we employ the cross-modal repetition priming experiment over a set of Bangla compound verbs. Our initial results show that non-compositional verb sequences exhibit priming only with \( V_2 \). On the other hand compositional verb sequences exhibit priming with both the constituent verbs \( V_1 \) and \( V_2 \). These observations together lead us to believe that mental representation and access of compound verbs in Bangla typically depends on the compositionality of the constituent verbs. Compositional verb phrases are accessed via decomposition of the phrase into its constituent verbs whereas non-compositional phrases are accessed as a whole.

In the next phase of our work, we try to develop a vector space model of semantic compositionality to predict the organization and processing of verb phrases in the mental lexicon. We evaluate our model with the data collected from our human judgment compositionality score and the cross-modal priming experiments. Comparing the models output with that of the empirically collected data, we claim that the proposed vector space model correctly predicts the semantic compositionality of Bangla verb sequences and their possible organization and processing in the mental lexicon. However, it would be premature to conclude anything concrete based only on the current experimental results. We also observe that several other factors including usage frequency and verb ordering affect the overall word recognition time and access mechanisms. Each of these factor calls for rigorous experimentation for understanding the exact nature of their interdependencies that we intend to perform in future.

Acknowledgement

We are thankful to all the participants who actively helped us to conduct the psycholinguistic experiments. We are also thankful to TCS Innovation Lab, Delhi for funding us to travel and attend the conference.

Reference


