Prediction failure and local coherence in a head-final language
Sakshi Bhatia (UMass, Amherst) & Samar Husain (IIT, Delhi)
sakshibhatia@umass.edu, samar.husain@hss.iitd.ac.in

Previous work has shown local coherence effects in Hindi [1] pointing to the fallibility of prediction in a head-final language, contra certain assumptions in the literature [2]. However, the previous studies do not investigate the source of this effect. Research on local coherence in a language like English has shown that a set of words become locally coherent if they have a syntactic and semantic interpretation [3]. In this paper we report two experiments evaluating the contribution of syntactic and semantic factors – agreement and plausibility respectively – in allowing local coherence to arise.

**Experiment 1: Agreement** is a key device to encode syntactic relations between words in Hindi, it could be that the presence of agreement constitutes a 'necessary' condition for local coherence to arise. We investigate this possibility using a 2x2 design crossing RC Word order (Canonical/Non-canonical) and post-RC Agreement type (Overt/Default). Canonical order is SOV (1a,1c); Non-canonical order is SVO (1b,1d). Importantly, all four experimental conditions are ungrammatical – the post-RC verb is thematically incompatible with the RC head noun NP_{Masc}. In the Overt Agreement conditions (1a,1b), the post-RC verb incorrectly bears Feminine agreement matching the RC internal object NP_{Fem} rather than the NP_{Masc}. In the Default Agreement conditions (1c,1d), the post-RC verb bears default agreement as the RC-internal object bears accusative case-marking which blocks agreement in the language. Items were presented using the centered self-paced reading paradigm, followed by an acceptability rating task. N_{Subj}=55 Hindi native speakers, N_{item} = 24. Conditions (1a) and (1b) are a replication of [1]. In [1] the local coherence effect had the following signature during reading: Hindi speakers were faster in reading the Non-canonical sentences at the critical verb and the post-critical region. This pointed to the forgetting of predictions regarding the matrix verb which would have been generated by the head noun.

If overt agreement features drive the local coherence effect observed in [1], then an interaction is predicted such that the effect of word order is magnified for the overt feature condition. Results are in Table 1. In the Ratings, we found a main effect of word order (t=2.56) - canonical order was preferred (replicating [1]) - and a main effect of features (t=4.43) - sentences with overt agreement were preferred. However, no interaction was observed. In Reading Times at the critical region, we observed a main effect of word order (t=2.00) – sentences with non-canonical order were read faster (replicating [1]) – but no other effects were significant. Even though overt agreement with the RC internal NP improves acceptability, in the absence of an interaction in the Ratings or RT data, we cannot conclude that agreement is a necessary component of the local coherence effect.

**Experiment 2: Plausibility** is a key device to encode semantic relations between words, and it could be the case that local coherence arises only when the RC-internal object is a plausible argument of the post-RC verb. In this experiment we cross Plausibility (Plausible/Implausible) with Word order (Canonical/Non-canonical). Conditions (2a), (2b) are identical to Exp 1 – the RC-internal NP_{Fem} is compatible with the post RC verb in terms of plausibility. In conditions (2c), (2d) – a different NP_{Fem} is used which is incompatible with the post RC verb. N_{Subj}=63, N_{item} = 24.

If plausibility drives the local coherence effect observed in [1], then an interaction is predicted such that the effect of Word Order is magnified for the Plausible condition. Results are provided in Table 2. In the Ratings, we found a main effect of word order (t=3.74) where canonical order was rated higher and a main effect of plausibility (t=15.8) – implausible RC-internal NPs correspond to lower ratings. However, no interaction was observed. In Reading Times at the critical region, we observed a main effect of word order (t=3.37) – sentences with non-canonical order were read faster – and a marginal interaction (t=1.94). No significant effects were observed at the post-critical region. Even though the semantic plausibility match between the RC-internal NP and the post-RC verb improves acceptability, the marginal interaction provides only weak support for plausibility being a necessary component of the local coherence effect. This calls for the replication of the findings of this experiment as well as further investigation into online measures of the local coherence effect.

These experiments replicate previous work demonstrating the nature of prediction in the head-final language Hindi. The consistent reading time advantage for non-canonical structures points to the building of locally coherent parses, contra expectations about prediction maintenance in head-final languages. While there is no evidence for the impact of agreement on the online parsing of local coherence structures there is weak evidence suggesting plausibility impacts the construction of these parses. This may have implications for representations assumed in self-organizing parsing accounts [4].
Experiment 1: Agreement

1. Sample Item (bold=critical region, underline=locally coherent parse, ‘\’=region breaks)
   (a) Canonical, Overt Agreement
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ kitaab\ paRhi\ thi\ mujhe\ bechnii\ paRii
   That boyMasc who\ yesterday lots interest with\ bookFem read\ had\ Fem IDat sell\ had-to\ Fem
   (b) Non-canonical, Overt Agreement
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ paRhi\ thi\ kitaab\ mujhe\ bechnii\ paRii
   That boyMasc who\ yesterday lots interest with\ bookFem read\ had\ Fem IDat sell\ had-to\ Fem
   (c) Canonical, Default Agreement
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ kitaab koi\ paRhaa\ tha\ mujhe\ bechnaa\ paRaa
   That boyMasc who\ yesterday lots interest with\ book Acc read\ had\ Def IDat sell\ had-to\ Def
   (d) Non-canonical, Default Agreement
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ paRhaa\ tha\ kitaab koi\ mujhe\ bechnaa\ paRaa
   That boyMasc who\ yesterday lots interest with\ read\ had\ Def book Acc IDat sell\ had-to\ Def

   Locally coherent parse: 1b, 1d. I had to sell the book

Spillover: Post-critical region = sentence level conjunct like and, but, however etc.

Table 1: Results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Ratings</th>
<th>Reading Times in ms (Critical region)</th>
<th>RT (post-crit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Canonical, Overt</td>
<td>4.01</td>
<td>1015 (38)</td>
<td>712 (25)</td>
</tr>
<tr>
<td>b. Non-Canonical, Overt</td>
<td>3.72</td>
<td>962 (38)</td>
<td>692 (25)</td>
</tr>
<tr>
<td>c. Canonical, Default</td>
<td>3.58</td>
<td>1016 (39)</td>
<td>731 (29)</td>
</tr>
<tr>
<td>d. Non-Canonical, Default</td>
<td>3.38</td>
<td>924 (33)</td>
<td>690 (29)</td>
</tr>
</tbody>
</table>

Experiment 2: Plausibility

2. Sample Item (bold=critical region, underline=locally coherent parse, ‘\’=region breaks)
   (a) Canonical, Plausible = 1a
   (b) Non-canonical, Plausible = 1b
   (c) Canonical, Implausible
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ aankh\ dekhi\ thi\ mujhe\ bechnii\ paRii
   That boyMasc who\ yesterday lots interest with\ eyeFem seen\ had\ Fem IDat sell\ had-to\ Fem
   (b) Non-canonical, Implausible
   Vah laRkaa\ jisne\ ka\ bahut dilchaspii se\ dekhi\ thi\ aankh\ mujhe\ bechnii\ paRii
   That boyMasc who\ yesterday lots interest with\ seen\ had\ Fem eye\ Fem IDat sell\ had-to\ Fem

Locally coherent parse: 1b. I had to sell the book. 1d. I had to sell the eye.

Spillover: Post-critical region = sentence level conjunct like and, but, however etc.

Table 2: Results

<table>
<thead>
<tr>
<th>Condition</th>
<th>Ratings</th>
<th>Reading Times, ms (Critical region)</th>
<th>RT (post-crit)</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Canonical, Plausible</td>
<td>4.30</td>
<td>1123 (62)</td>
<td>655 (22)</td>
</tr>
<tr>
<td>b. Non-Canonical, Plausible</td>
<td>3.97</td>
<td>882 (37)</td>
<td>655 (30)</td>
</tr>
<tr>
<td>c. Canonical, Implausible</td>
<td>2.92</td>
<td>1063 (62)</td>
<td>723 (32)</td>
</tr>
<tr>
<td>d. Non-Canonical, Implausible</td>
<td>2.58</td>
<td>921 (34)</td>
<td>687 (21)</td>
</tr>
</tbody>
</table>

References: