Reviewing incremental interpretation in second language sentence processing

Shazia Akbar Ghilzai

Abstract

Sentence processing research suggests that the readers interpret sentences in an incremental fashion and attach each incoming word immediately to the developing sentence structure. For instance Just & Carpenter (1980) “immediacy hypothesis" suggests that during comprehension, the reader or listener attempts to interpret the input in as incremental a fashion as possible. Kintsch (1988) proposed that while listening or reading readers update their situation model continually as each word is processed. While processing a sentence and making the attachment decisions if the readers encounter a disambiguating region that doesn’t fit to the evolving sentence the reader revises their initial analysis. This reanalysis is known as garden path effect. While processing the filler gap dependency if the initial analysis is implausible it reduces the garden path effect. It is assumed that it is harder to reanalyse a structure that was initially thought to be plausible than one that was implausible because readers commit more strongly to plausible analysis (Pickering and Traxler, 1998, 2003). Williams et al. (2006, 2001) investigated the degree by which the native and non-native speakers interpreted the sentences incrementally. Performance was evaluated in relation to parsing strategies and sensitivity to plausibility constraints. The present research, however, is reviewing their work on ‘incremental processing in second language sentence processing’.

Keywords: filler gap dependency, plausible, implausible, sentence processing

1 Corresponding Author: Shazia Akbar Ghilzai, faculty member at Department of Linguistics, QAU. Email: sghilzai@qau.edu.pk. MA. Psycholinguistics and Neurolinguistics (The University of Essex, UK). A Doctoral student at USPC Paris13. Research Lab TTN Paris 13
Williams et al. (2001) study of processing long distance dependencies in L1 and L2 provided evidence for immediate computation of plausibility in non-natives. Following sentences (a) and (b) were compared.

(a) **Plausible-at-v**:

```
<table>
<thead>
<tr>
<th>Initial gap filler assignment</th>
<th>filled gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which <em>machine</em> did the mechanic fix—the motorbike with two weeks ago?</td>
<td></td>
</tr>
<tr>
<td>filler</td>
<td>false Gap</td>
</tr>
<tr>
<td></td>
<td>gap</td>
</tr>
</tbody>
</table>
```

In sentence (a) there is a potential gap after the verb *fix*. While reading the sentence the readers initially fill this gap by assuming the machine as the direct object of verb fix. On further reading the sentence when they found that the gap is filled by an unexpected noun phrase *the motorbike*, their reading process slows down in the post-verbal region. The reason is that the filler *machine* that they had assigned initially confuses them. The research question under discussion is whether the filled gap effect is affected by the plausibility of the initial filler gap assignment.

(b) **Implausible-at-V**:

```
<table>
<thead>
<tr>
<th>Initial gap filler assignment</th>
<th>filled gap</th>
</tr>
</thead>
<tbody>
<tr>
<td>Which <em>customer</em> did the mechanic fix—the motorbike with two weeks ago?</td>
<td></td>
</tr>
<tr>
<td>filler</td>
<td>false Gap</td>
</tr>
<tr>
<td></td>
<td>gap</td>
</tr>
</tbody>
</table>
```

In sentence (b) the initial filler is implausible at verb position. (More stop making sense decisions were found at the implausible-at-V condition, but there was an opposite pattern following the postverbal noun.)

Williams et al. (2001) found that the difference between natives and non-natives in manipulating the plausibility information. In reading times the natives showed a plausibility effect. At the postverbal determiner, the implausible at v condition was more slower than the plausible at v condition.

The non-native speakers showed the plausibility in the post verbal noun region where the plausible at v condition was slower than the implausible at v condition. It could be the result of argument
competition, that is the process of substituting the post verbal noun for the filler as the theme in the verb argument structure. This substitution process is more difficult when the initial assignment is plausible. But in this case, the difficulty could be the result of thematic /situation model level process rather than processes of syntactic revision.

Experiment 1

Williams, J. N. (2006) further explored the plausibility effects in the post verbal region by increasing the number of words prior to the noun e.g.

(c) **Plausible - at - V**

false gap

Which *machine* did the mechanic fix --- *the very noisy motorbike* for --- two weeks ago?

<table>
<thead>
<tr>
<th>filler</th>
<th>false gap</th>
<th>gap</th>
</tr>
</thead>
</table>

(d) **Implausible - at - V**

Which *friend* did the mechanic fix --- *the very noisy motorbike* for --- two weeks ago?

| filler   | false gap | Det-int-adj (critical region) |

Hypothesis

It was hypothesized that if for non-native readers plausibility effects reflect argument competition, then reading times in the plausible at V condition will become slower than the implausible condition at and following the post verbal noun. If there is merely a delay in the utilization of plausibility then plausibility effects might still be obtained prior to the noun, but later than for natives.

Procedure

Williams, J. N. (2006) used the self-paced reading technique for the Experiment 1. Participants read the sentences word by word and changed the each word into the next word by clicking the mouse. They were required to pressed the space bar as soon as they thought
that the sentence had stopped making sense. This technique yields reading time per word. The recorded RT allows researchers to see how long participants spend reading a critical word or fragment within a sentence. If RT for individual words is recorded, researchers can combine reading times across several words to obtain RT per region. Differences in RT for critical words, fragments, or regions are then compared for various conditions in the experiment. This technique assumes and exploits the incremental nature of reading; thus, if at a given point in the stimulus, the reader encounters a word or region that does not fit their structural or semantic expectations or that leads to an increase in complexity, an increase in RT is expected on that critical word or soon after (spill over effects).

Results

Stop making sense decisions
**Figure 1:** stop making sense decisions made by the native participants at each position in the two experimental conditions.

All language groups (Native, Chinese and Romance participants) showed the same pattern of stop making sense decisions at the same critical positions in both experimental conditions. The critical region under observation is the immediate post verbal determiner-intensifier-adjective region (figure 1). There was a high rate of stop making sense decisions at the verb “fix” in the implausible-at-V condition, that shows the filler was being assigned to the first gap position resulting in an immediate stop making sense decision. Statistical results showed that there was a significant effect of plausibility at the verb position ($p< 0.05$). At the determiner “the” there were more responses in the implausible-at-V condition than the Plausible-at-V condition. But the pattern was reversed over the intensifier “very” and adjective “noisy”. Williams, j (2006) suggested that the more stop making sense decisions in the implausible-at-V condition on the determiner is a spill over from the effect that is present on the verb.

**Reading times**
Figure 2: Mean Reading times in ms

Figure 2 shows that the reading times pattern for the three language groups was similar. The reading times were slower in the plausible-at-V condition in the det-int-adj region (the very noisy). The reading times were higher at the intensifier position than determiner and adjective. Statistical analysis showed that there was a significant effect of plausibility. The interaction between plausibility and position was also significant for all language groups. Reading time at verb was only slower for the Romance group in the implausible at V-condition.

Discussion

All groups of participants initially interpreted the filler as the direct object of the verb in both the plausible and implausible conditions. There was a higher rate of stop making sense decisions at and immediately following the verb in the implausible condition. In contrast in the plausible condition there were more stop making sense decisions and slower reading times prior to the post verbal noun.

At the implausible condition there was more stop making sense decisions at the verb position than the plausible condition because the filler was implausible as the direct object. But there was no difference between the two conditions in terms of reading times. It means that the participants were able to understand the implausibility at the verb condition. Williams et al. suggested that the experiment one provides the evidence that the non-native readers compute the plausibility of the potential filler gap relation at the verb and this influences the difficulty of reanalysis, it does not tell that these processed operate in other reading situations that do not compel incremental interpretation at the level of situation model. They designed another experiment to analyse the above issue by examining the same critical material but there was no requirement to interpret the sentences in an incremental fashion.
Experiment 2

Procedure
The sentences were presented using the same word-by-word presentation method as in Experiment 1. Memory probes were presented as whole sentences. Participants read two sentences followed by the memory probe for the first sentence, followed by the memory probe for the second sentence.

Which bucket did the lady wash the very large shirt in early this morning?
Which baby did the boy drop the very small toy on just after lunch?
The lady washed a ___
The boy dropped the toy on the___

Responses to the memory probes were given orally.

Results
Figure 3: Experiment 2. Mean reaction times in ms

Figure 3 shows the mean reaction times. At the critical region (det-int-adj) reading times were faster for the natives than the non-natives. There was no main effect of plausibility. However, there was some divergence in the native data between the two conditions at and following the post verbal determines and a small plausibility effect at the critical det-int-adj region. While for the non-natives there was no plausibility effect in the critical det-int-adj region. In fact the plausibility emerged after preposition and there was a significant plausibility over the last three words of the sentence where the same words were used in both conditions. Overall, the plausibility effect in the critical post-verbal region the very noisy was not significant.
In order to investigate the relationship between plausibility effects and memory tasks performance in the native group, the native participants were divided into two groups (i.e. high memory group and low memory group) according to their memory scores. Figure 4 shows that the low memory group showed no effect of plausibility in the det-int-adj region; whereas for high memory group reading time was slower in the plausible condition than the implausible condition, especially at the determiner and intensifier. High memory group showed a larger plausibility effect over the last three words position. It reflects a sustained argument competition as the readers in the plausible-at-v condition find it difficult to maintain a stable representation of the sentence when there are two arguments that can plausibly fulfil the same role. The Low memory group showed the plausibility effect only at the post verbal noun.

**Figure 4**: reading time for high versus low memory natives.
Figure 5: Reading times for high memory non-natives versus low memory non-natives.

In order to see the relevance of memory task performance, the participants were roughly into two groups according to their overall performance on the memory task.

Figure 5 shows that the plausibility effects the det-int-adj appears to be different for the two groups. Only the high memory native group showed greater processing difficulty in the
immediate post verbal region in the plausible-at- v condition. Only the high memory group showed any plausibility effects, but this effect was delayed until the preposition. This might be reflecting a delayed effect of argument competition because when the post verbal noun was encountered its plausibility was not computed immediately for argument competition but delayed until the following word. The low memory group showed a slight effect of argument competition at the noun, but this effect did not persist till the end of the sentence.

The lack of plausibility effect in the critical region shows that although a gap was postulated, its plausibility was either not evaluated or the plausibility information did not immediately affect the subsequent processing. The low memory group did not show the plausibility affect, it means that they did not evaluate the gap. But the high memory group evaluated the gap because they showed the clear differences between the plausible-at-v and implausible-at-v conditions at and following the preposition. But the plausibility was not immediately affecting processing in the high memory non-natives. Williams et al. proposed two possible reasons for this delayed effect of plausibility. First, the plausibility of the filler as the direct object was not computed as much quickly as to influence the reader’s commitment to initial filler gap assignment before structural cues triggered reanalysis. Second reason is that there was no actual delay in the computation of plausibility. The non-native group only updated their interpretation of the filler when they encountered a relevant gap in the input. At the verb the filler would be immediately interpreted as the direct object and theme, regardless of plausibility. This analysis was only modified when the next was encountered at the preposition, where it became clear that the filler performed an adjunct role. Reanalysis would have been harder for the readers when the initial analysis was plausible.
Figure 6 shows the reading profile from the verb to the head of the post verbal noun. The non-native reading profiles were compared over the post verbal region in experiment 1 and 2. The comparison was made to examine the two issues:

- Was there a filled gap effect that was not sensitive to plausibility?
- Was there simply no filled gap effect at all?

It was hypothesized that:

- *If first were true* then it could be claimed that participants were interpreting the sentences at a structural level because they had postulated a gap after the verb (but not at the level of the situation model).

- *If the latter were true*, then it could not be claimed that the participants were even interpreting the sentences at the structural level.

The plausibility effect was larger for experiment 1 than experiment 2. In experiment 1 the reading times in the plausible-at-v and implausible-at-v conditions diverge at the intensifier and noun, where as in experiment 2 the reading times were equivalent at each position.
Over the critical region reading times of experiment 2 seems to pattern with the plausible at-v condition of experiment 1. That shows that the readers postulated at gap at the verb. All three conditions showing a lower rise than in the plausible-at-v condition of experiment 1. It suggests that there was a filled gap effect in this region in both conditions of experiment 2 and the participants were performing a structural analysis of the filler gap relation. If there was no filled gap effect in experiment 2 then one would expect reading times in the critical region to be relatively fast and more like the implausible condition of experiment 1. It is clear that the readers were encountering a processing difficulty in the critical region of in both conditions of experiment 2. It suggests that a gap has been assumed at the verb and it created difficulty for the readers to process the subsequent det-int-adj region. Absence of plausibility in the critical region suggests that although a gap was postulated, but its plausibility was either not evaluated or the plausibility information did not immediately affect the subsequent processing. The increase in the reading time from the adjective to the noun in the plausible-at-v conditions was greater in experiment 1 than in the experiment 2.

**Discussion**

Williams, J. N. (2006) explored the plausibility effects in native and non-native sentence processing. One issue major was whether such effects are delayed in non-native readers, or whether there is any difference in the way in which plausibility affects performance, through the argument competition rather than through a direct interaction with syntactic processing. Experiment-1 showed the same cross over pattern predicted by Pickering and Traxler (1998, 2000). Experiment-1 also provided strong evidence against the argument competition because the non-native group processed the plausibility effects in the disambiguating region even before the post verbal noun. Both natives and non-natives processed the plausibility of the of the potential filler gap relationship at the first verb irrespective of nature of their L1. The results suggest that both native and non-native readers were sensitive to plausibility and both
used the AFS (Active filler strategy). AFS assumes that the parser can identify fillers like *wh*-elements (displaced constituents after movement) and it mandates that these fillers be assigned to the first empty slot available in a syntactic tree a left-to-right parse. It is, therefore, filler-driven (as opposed to gap-driven) parsing.

In experiment 2, when the native and non-native groups were combined reading times were faster for natives than non-natives. The slower reading times for non-natives may be related to automaticity. Language learners may process the target language less rapidly than adult native speakers, possibly reflecting a lack of automaticity (Segalowitz, 2003).

There was no main effect of plausibility as well. Although, numerically, there was a small plausibility effect in the native group, this was only significant in the participants’ analysis and there was no interaction between participant group and plausibility.

However, when the natives and non-natives were divided into high memory and low memory groups, there was variability in the size and location of plausibility effect among the natives and non-natives. For *Native speakers*, the high memory natives showed an effect of plausibility at the determiner (*the*) and intensifier (*very*), while the low memory natives showed effects at the noun. When they did the same analysis on the *non-native speakers*, a different pattern emerged. The *high memory non-natives* showed plausibility effects at the preposition following the noun, while the *low memory non-natives* showed no plausibility effects at all. Williams et al. proposed a cline of incrementality for these systematic variations in degrees of incrementality for the native and non-natives. The cline of incrementality can also be views as differences in the degree of automaticity.

While the *low memory non-natives* showed no plausibility effects at all. Some non-native readers are not able to apply plausibility information during online processing of filler-gap dependencies. One possible reason could be that these non-native readers do not access full semantic representations of fillers, and it may be difficult for them to access plausibility
information online, which could be due to a variety of reasons e.g. working memory or motivation. Another possibility could be that, some L2 speakers rely on a structural heuristic leading them to ‘dump’ the filler at the first verb encountered, regardless of conflicting lexical information. This idea that late L2 learners rely on heuristics during sentence processing was suggested by Clahsen and Felser (2006).

A closer inspection of results show that Experiment 1 results were clearly different from experiment 2: the results of experiment-1 indicated that both native and non-native speakers fill gaps immediately and use the plausibility information to reanalyse filler-gap dependencies. On the other hand in experiment 2 when the native and non-native groups were compared, plausibility effects in the critical post verbal region were not significant. Even, when the subjects were divided into high and low memory groups. Comparing the high memory groups only high memory non-natives showed significant plausibility effect in the critical det-int-adj region. Whereas the best memory non-natives the plausibility effects were not significant in the critical region. Low memory natives and low memory non-natives also didn’t show plausibility effects at the critical region. In short only high memory natives showed plausibility effects in the critical det-int-adj region. William et al. tried to justify their results by arguing that the differences in results are due to different types of tasks they used (stop making sense decision, memory comprehension tasks). When the task focuses attention on plausibility, plausibility information is used, but when the task does not focus attention on plausibility, then this information is not used by all of the readers.

References


