Anticipatory Processing with an Impoverished Cue:
An Experimental Study of Wh-Scope Marking Constructions in Hindi

Introduction: We investigated the significance of the informativity of a predictor in anticipatory processes in language comprehension with wh-scope marking constructions in Hindi as the targeted structure. It is important to realize that it is not only the presence of a cue/ predictor that affects the parser establishing a dependency or predicting an expectation. The nature of the predictor in terms of what information it encapsulates about the dependent elements also matters as the parser registers such information and uses it to qualify the dependency formed.

We present experimental evidence that even an impoverished weak predictor results in expectations; giving further evidence for anticipatory processing being a pervasive process during sentence comprehension. A predictor could be impoverished in the sense that it is not a cue to any syntactic or other dependency and is informatively weak in the sense that it does not very specific as to what it leads to an occurrence of, during incremental sentence processing.

1. abhi-ne   **kya**   kaha   tha   ki diya-ne   **kisko**   daanta tha
   abhi-ERG what say be.PST that diya-ERG   whom   scold   be.PST
   'What did Abhi say, whom did Diya scold?'

In 1, *kya* in the matrix clause is a weak predictor of the upcoming wh-word in the embedded clause as there is no syntactic dependency between the two elements assuming Dayal (1994) and Lahiri's (2002) Indirect Dependency Approach. Furthermore, it is not always the case that *kya* in 1 is followed by a wh-word in the embedded clause making it a probable prediction which could be violated. In cases where the wh-word does follow, *kya* does not reveal any information as to the grammatical role or the position of the upcoming wh-word like fronted wh-words or scope markers in certain other languages like English or German do. These features make *kya* a relatively impoverished cue.

Experiment: In order to figure out whether the impoverished cue *kya* leads to an expectation, in addition to structures like 1 (*kya*_wh condition) we had a condition wherein a *kya* was present but there was no following wh-word (*kya*_dec) to test if there was any processing cost when the expectation was violated. Our third condition had an expletive *yeh* 'this' instead of *kya* and a wh-word followed in the embedded clause (*yeh*_wh) to test whether presence of *kya* eased the processing time of the following wh-word. The forth target condition was a baseline condition had an *yeh* followed by a declarative construction (*yeh*_dec). We conducted a self-paced reading experiment (N=32) with a 2X2 paradigm where we manipulated the cue type (*kya* or *yeh*) and the embedded clause type (wh or declarative).

2.   [*yeh*-wh][1]
abhi-ne yeh   kaha tha   ki kal   diya-ne   kisko daanta tha us harkat ke   liye
abhi-ERG this say be.PST   that yesterday diya-ERG   whom   scold   be.PST that doing(N) for
 'Abhi had said whom Diya had scolded.'

3.   [*kya*_dec]
abhi-ne kya   kaha tha   ki kal   diya-ne   radhika-ko daanta   tha us   harkat ke liye
abhi-ERG what say   be.PST   that yesterday diya-ERG   radhikaACC scold   be.PST that doing(N) for
'Did Abhi say that Diya had scolded Radhika?'

4.   [*yeh*_dec]
abhi-ne yeh   kaha tha   ki kal   diya-ne   radhika-ko daanta   tha us harkat ke liye
abhi-ERG this say be.PST   that yesterday diya-ERG   radhika   scold   be.PST that doing(N) for
'Abhi had said that Diya had scolded Radhika.'

Results: The data was analyzed with mixed-effects regression using R and the lme4 package. Reading times were analyzed for the critical region i.e. the embedded object position. We found that the *kya*-wh condition was the fastest to read and the *kya*_dec condition wherein there was a name DP rather

---

[1] The *kya*_wh condition in 1 also had spillovers *kal* 'yesterday' and *us harkat ke liye* 'for that mischief' same as that in 2-4.

than a wh, was the slowest to read. We found a main effect of the embedded clause type (t = -6.61) (at the embedded object and the embedded verb region), the declarative conditions were overall slower that the wh conditions. There was also a significant interaction (t = -3.34) stating that this difference between the wh conditions and declarative conditions were higher with *kya* cue. These results were in accordance with our hypothesis, that *kya* does generate an expectation of a following wh-word.

However, unexpectedly at the embedded object position, the *yeh_dec* condition was also very slow. Most of the post-study feedback comments indicated that the participants had found a lot of names (3 in each target conditions and more in fillers) and had employed the strategy of remembering the names as a considerable number of comprehension questions asked about the names. The concern was that the slowdown at the eo position in the *kya_dec* condition was due to participants remembering names rather than due to the violation of an expectation of a wh-word as hinted by the unexpected slowdown in *yeh_dec* condition as well (both the dec condition had names in the eo position). To mitigate this name-memory effect we analyzed the data only from the first eight targets as we argue that this overt strategy only could have developed over time. We did find a significant main effect of the presentation order of targets with t= -3.29.

READING TIMES AT EACH POSITION - TARGETS 1 - 8
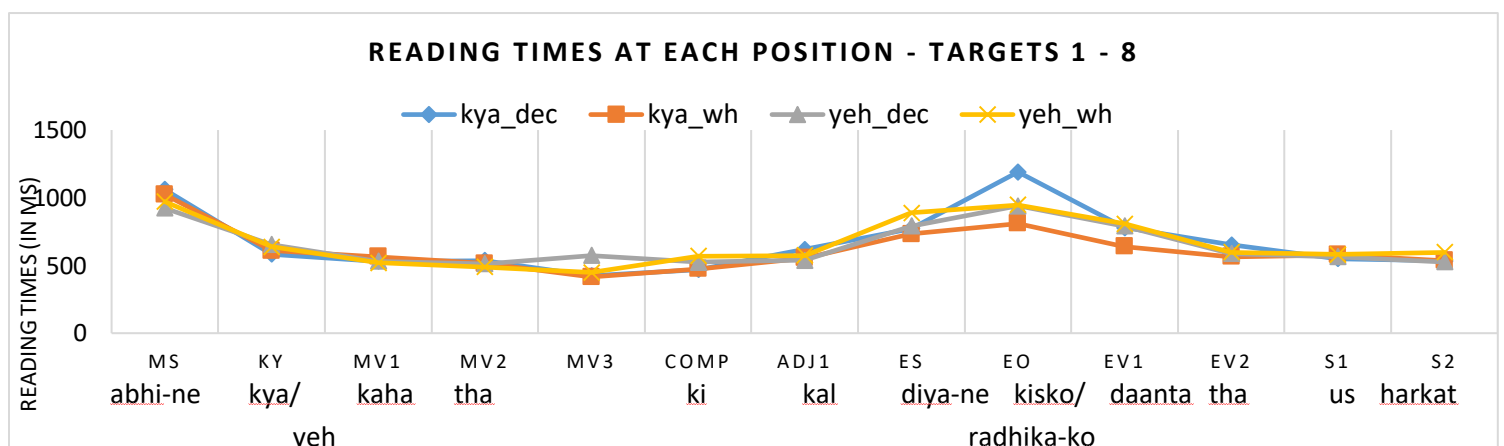
Crucially, as we look at the first eight targets (as shown in the figure above), we indeed find that the *kya_dec* condition is the slowest at read at the eo position. And this slowdown is argued to be a result of the violation of the expectation rather than the participants' using a name-memory strategy as is proved by no such slowdown in the *yeh-dec* condition. In fact the *yeh_dec* condition and the *yeh_wh* condition had similar reading times at the eo position. As per our expectations, the *kya_wh* condition was the fastest to read. Similar to our results for the entire target set, we found a significant main effect of the embedded clause type (t = -3.13) and also a significant interaction (t = -3.83).

Further evidence that the *yeh_dec* condition was slower to read at the embedded object positon due to the name-memory effect is that the names took longer to read even at the embedded subject position for the later targets than for the first eight targets. Presentation order of the targets was calculated to have a marginally significant effect with t= -1.91 at the es position.

Conclusion: Our results indicate that even a weakly informative cue – namely the scope-marker *kya* in Hindi – results in the generation of an expectation. This is evidence that predictive processing is pervasive and does not seem to be constrained by the strength of the cue even though it does take into account the information encapsulated within the cue to qualify what comes ahead.

References: Dayal, V. (1993). Scope marking as indirect wh-dependency. Natural language semantics, 2(2), 137-170. Lahiri, U. (2002). On the proper treatment of "expletive wh" in Hindi. Lingua, 112(7), 501-540.