#### Samar Husain and Sakshi Bhatia (Not) Forgetting Verbs in Hindi Doubly Center-embedded Structures Indian Institute of Technology, Delhi; University of Massachusetts, Amherst **AMLaP 2018**

# Question

- What is the nature of forgetting in the head-final language Hind
- This study:
- Hindi native speakers are not susceptible to verb forgetting effects in center-embedded structures
- Hindi patterns with other head-final languages
- Consistent with the predictions of the language adaptability hypothesis
- However, comprehension data points to the possibility of shallow parsing

# **Verb Forgetting in Center-embedded Structures**

- Working-memory constraints are known to induce 'forgetting' effects in cen embedded constructions in English [2].
- Forgetting the prediction of the upcoming VP is argued to underlie the illusi of grammaticality observed in sentences with a missing verb phrase:
- \*The patient who the nurse who the clinic had hired met Jack. (1)
- The verb forgetting effect has not been observed in head-final languages such as German and Dutch [1, 3], but see [4].
- This asymmetry in processing in head-final languages has been attributed the parser's adaptability to certain language characteristics such as head directionality.
- The parser encounters a large proportion of head-final structures and, therefore becomes very efficient in predicting and maintaining the upcoming verbal heads.

## Motivation

- Prediction processes are fallible in Hindi. [5, 6]
- Ungrammatical sentences with center embedded relative clauses [6]
- Word order manipulation within RC Canonical (2a) & Non-canonical (2
- Speakers are unable to sustain the prediction of the matrix verb that was be integrated with the head noun in the face of the locally coherent pars available in the Non-canonical order (2b)
- Indexed by faster reading times at the matrix verb for non-canonical ord
- a.  $NP_{Masc}$  [Rel-pro<sub>*Erg*</sub> ...  $NP_{Fem}$  RC-V<sub>*Fem*</sub>]  $NP_{Dat}$  **Verb**<sub>*Fem*</sub> **Aux**<sub>*Fem*</sub> (2) NP<sub>Masc</sub> [Rel-pro<sub>Erg</sub> ... RC-V<sub>Fem</sub> NP<sub>Fem</sub> ] NP<sub>Dat</sub> Verb<sub>Fem</sub> Aux<sub>Fem</sub>
- In light of these results, both the 'forgetting hypothesis' and the 'adaptation hypothesis' need to be tested further cross-linguistically.

# **EXPERIMENT**

- Doubly center embedded structures of the type previously employed by [1]
- Experimental manipulation:
- a. Grammatical: All verbs present b. Ungrammatical: Missing V2
- NP1; [Relpro; NP2.object; [Relpro; NP3 NP4.object V3] V2] V1 ... (3) NP1; [Relpro; NP2.object; [Relpro; NP3 NP4.object V3]  $\emptyset$ ] V1 ...

http://web.iitd.ernet.in/ samar/index.html; www.blogs.umass.edu/sakshibhatia

lt	tem, Fillers, Methods	
(4)	Experimental Item ('/' indicates region	h breaks. Critical region
	a. vah dhobi/ jo/ us doctor That washerman who that doctor <b>raha tha</b> -ing was	ko/ jo/ mariiz se/ pa Acc who patient from m
	'That washerman was angry who	was seeing the doctor w
	<ul> <li>b. vah dhobi/ jo/ us doctor</li> <li>That washerman who that doctor</li> </ul>	· ko/ jo/ mariiz se/ pa · Acc who patient from m
(5)	Spillover region ('/' indicates region br	eaks. Post-critical regio
	<b>magar</b> / baad me/ uskaa/ gussaa/ kar but later Loc his anger rec 'but later on his anger died down.'	m ho gaya. luce happen went
Fill	ers:	
•	Simple declarative sentences C	lausal complements
•	Embedded RCs R	ight-extraposed RCs
•	Correlatives Se	entential coordination
Me	thods:	
•	Centered self-paced reading	<ul> <li>24 latir</li> </ul>
	Y/N comprehension questions on 66% of t	trials ► N=48 n
H	Y/N comprehension questions on 66% of to the set of the	trials N=48 n
Fo	Y/N comprehension questions on 66% of to <b>IYPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> > <sub>R</sub>	trials ► N=48 n TIONS TUngrammatical
For	Y/N comprehension questions on 66% of t IYPOTHESES and PREDIC rgetting Hypothesis [2]: Grammatical > <sub>R</sub> Reading times at V1 (and possibly at the p	trials ► N=48 n TIONS T Ungrammatical post-critical region due to
For	Y/N comprehension questions on 66% of to <b>IYPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably	It is N=48 mmonology in the second secon
For	Y/N comprehension questions on 66% of the <b>IYPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the possibly at the possibly the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$	TIONS TIONS <sup>T</sup> Ungrammatical post-critical region due to been forgotten at NP4 in TUngrammatical (b) should be slower than
For	Y/N comprehension questions on 66% of t <b>IYPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the possibly at the possibly at the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b).	TIONS T Ungrammatical post-critical region due to been forgotten at NP4 in T Ungrammatical (b) should be slower than
For	Y/N comprehension questions on 66% of the <b>IVPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b).	It is N=48 means and N=48 means N=48 means N=48 means N=48 means N=48 means N=48 means N=4 m
For Ada •	Y/N comprehension questions on 66% of the <b>IVPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b).	TIONS TIONS TOUS TUNGRAMMENT TUNGRAMENT TUNGRAMMENT TUNGRAMMENT TUNGRAMMENT TUNGRAMMENT TUNGRAMMENT TUNGRAMENT TUNG
For Ada •	Y/N comprehension questions on 66% of the <b>IVPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b).	Frials ► N=48 m TIONS T Ungrammatical post-critical region due to been forgotten at NP4 in T Ungrammatical (b) should be slower than d-effects models.
For Ada •	Y/N comprehension questions on 66% of the <b>IVPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b). <b>RESULTS: Reading Times</b> Log RTs were analyzed using linear mixed Significant difference at the post-critical reading to the post-critical reading times at V1 in (	TIONS TIONS TUngrammatical post-critical region due to been forgotten at NP4 in Tungrammatical (b) should be slower than d-effects models. region (t=2.4): Grammatical
For Ada A	Y/N comprehension questions on 66% of the approximation of the second provide the second provide the second provide the second provides and the second provide the second provides and the second prov	TIONS TUNGRAMMATICAL post-critical region due to been forgotten at NP4 in TUNGRAMMATICAL (b) should be slower than d-effects models. region (t=2.4): Grammatical Figure 1. Reading Times of
For Ada •	Y/N comprehension questions on 66% of the approximation of the second se	TIONS TUNG TUNGRAMMATICAL post-critical region due to been forgotten at NP4 in the slower than d-effects models. region (t=2.4): Grammatical Figure 1. Reading Times of
For Ada •	Y/N comprehension questions on 66% of the <b>IYPOTHESES and PREDIC</b> rgetting Hypothesis [2]: <b>Grammatical</b> $>_R$ Reading times at V1 (and possibly at the powing to the fact that V2 has presumably aptation Hypothesis [1]: <b>Grammatical</b> $<_R$ opposite pattern – reading times at V1 in ( number of verbal heads in (b). <b>RESULTS: Reading Times</b> Log RTs were analyzed using linear mixed Significant difference at the post-critical restricted of the pos	TIONS TUNG TUNG TUNG TUNG TUNG TONS TUNG
For Adda •	Y/N comprehension questions on 66% of the approximate of the second structure	TIONS TUngrammatical post-critical region due to been forgotten at NP4 in the slower than b) should be slower than d-effects models. region (t=2.4): Grammatical Figure 1. Reading Times of
	Y/N comprehension questions on 66% of the second state of the sec	TIONS <sup>7</sup> Ungrammatical post-critical region due to been forgotten at NP4 in a⊤ Ungrammatical (b) should be slower than d-effects models. region (t=2.4): Grammati Figure 1. Reading Times of
	Y/N comprehension questions on 66% of the second state of the sec	TIONS TUNS TUNGrammatical post-critical region due to been forgotten at NP4 in TUNGrammatical (b) should be slower than d-effects models. region (t=2.4): Grammati Figure 1. Reading Times of Tungrammatical
	Y/N comprehension questions on 66% of the average of the second state of the second s	TIONS TUngrammatical post-critical region due to been forgotten at NP4 in TUngrammatical (b) should be slower than d-effects models. region (t=2.4): Grammati Figure 1. Reading Times of The second
For Adda •	Y/N comprehension questions on 66% of the action of the second state of the second st	TIONS TUngrammatical post-critical region due to been forgotten at NP4 in TUngrammatical (b) should be slower than d-effects models. region (t=2.4): Grammati Figure 1. Reading Times of The second

#### n=V1 bolded)

rahaa tha/ dekh raha tha/ **gussa kar** aise/ le noney take -ing was see -ing was anger do

ho was taking the money from the patient. aise/ le rahaa tha/  $\emptyset$  / **gussa kar raha tha**... anger do -ing was noney take -ing was on 1 bolded)

- Declarative sentences with complex NPs
- n-squared items, 48 fillers
- native speakers of Hindi at IIT, Delhi

to spillover) in (a) should be slower than (b), (b).

(a) because of not encountering the required

# ical <<sub>RT</sub> Ungrammatical. (ms)

Condition Grammatical --- Ungrammatical

Position

# **Results: Comprehension Accuracy**

- Experimental items: only 33% of the total participants exceed 70% in their comprehension accuracy for the items.
- Filler sentences: 100% of the participants exceed 70% comprehension accuracy in the fillers.
- Average comprehension according to the second se

#### Comprehension Acc

- conditions (z=-1.6).
- Comprehension accuracy in German [1]. Grammatical = 65%, Ungrammatical = 71.5%.
- Comprehension accuracy for questions about NP2 in French [7]. Grammatical = 68%, Ungrammatical = 49%.

#### **Discussion**:

- i. the parser is making correct structural integrations
- ii. the parser is using a surface cue (e.g., counting the number of Relpros since these clearly mark clause boundaries) to track the upcoming heads.
- The comprehension data for this experiment makes (ii) seem more likely.

## CONCLUSION

- RT results: Hindi native speakers are not susceptible to verb forgetting effects in doubly center-embedded structures.
- This result patterns with other head-final languages.
- This result is consistent with the predictions of the language adaptability hypothesis [1].
- However, the low comprehension accuracy suggests a shallow parsing strategy where the required structural integrations may not be taking place in spite of the successful tracking of the number of verbal heads. - The low comprehension accuracy in [1] also points to this.
- Given that in Hindi prediction errors are frequent [5] and predictions can be forgotton [6], the role of robust prediction and its maintenance as an explanation for the lack of forgetting effects in head-final languages needs to be further probed.

## Acknowledgments

#### References

2009.



Comparing participant responses across items and fillers:

curacy	for exper	imental	items i	s not high
	<b>C</b>		1.1.1	

	Grammatical	Ungrammatical	
curacy (%)	71	68	

Non significant difference between Grammatical and Ungrammatical

The reading time data (Fig.1) is compatible with two underlying states:

- Thanks to Vasundhara Srivastava for helping run this experiment. This study was supported by an IIT Delhi Faculty Grant to the first author.
- [1] Vasishth et al., 2010. [2] Gibson & Thomas, 1999. [3] Frank et al., 2016. [4] Häussler & Bader, 2015. [5] Apurva & Husain, 2018. [6] Bhatia & Husain, 2018. [7] Gimenes et al.,

samar@hss.iitd.ac.in; sakshibhatia@linguist.umass.edu