Interference Predicts Locality: Evidence from an SOV language

Sidharth Ranjan
University of Stuttgart
sidharth.ranjan03@gmail.com

Sumeet Agarwal
IIT Delhi
sumeet@iitd.ac.in

Rajakrishnan Rajkumar
IIIT Hyderabad
raja@iiit.ac.in

Abstract

LOCALITY and INTERFERENCE are two mechanisms which are attested to drive sentence comprehension. However, the relationship between them remains unclear—are they alternative explanations or do they operate independently? To answer this question, we test the hypothesis that in Hindi, interference effects (measured by semantic similarity and case markers) significantly predict locality effects (modelled using dependency length quantifying distance between syntactic heads and their dependents) within a sentence, while controlling for expectation-based measures and discourse givenness. Using data from the Hindi-Urdu Treebank corpus (HUTB), we validate the stated hypothesis. We demonstrate that sentences with longer dependency length consistently have semantically similar preverbal dependents, more case markers, greater syntactic surprisal, and violate intra-sentential givenness considerations. Overall, our findings point towards the conclusion that locality effects are reducible to broader memory interference effects rather than being distinct manifestations of locality in syntax. Finally, we discuss the implications of our findings for the theories of interference in comprehension.

1 Introduction

The language comprehension system is long known to be constrained by working memory considerations (Ebbinghaus, 1885; Yngve, 1960; Ebbinghaus, 2013). Several theories and mechanisms of *syntactic complexity* have been proposed in sentence comprehension literature to account for processing difficulties. LOCALITY and INTERFERENCE are two such mechanisms by which online, incremental processing happens in language comprehension (Vasishth, 2011). As defined in that work, locality is the claim that the distance between syntactically related words (*i.e*, dependent and head) determines the difficulty of integrating a dependent with its head in a syntactic structure, owing to limited working memory capacity during comprehension. In contrast, the notion of interference denotes situation wherein linguistic elements sharing common characteristics, such as form, meaning, animacy, or concreteness, result in processing difficulties when they are situated nearby or in close proximity (Lewis, 1996). Notably, interference effects causing forgetting has a long history in cognitive psychology literature, suggesting that interference could be a manifestation of memory overload during retrieval (Baddeley and Hitch, 1977; Roediger III and Abel, 2022).

In an extensive survey of locality and interference effects and their interplay, Vasishth (2011) proposed that locality and interference may represent two sides of the same underlying memory effects. For example, locality represented in terms of the number of discourse referents between head and dependent in theories like Dependency Locality Theory (DLT, Gibson, 2000), primarily reflecting the accessibility or availability of intervening elements. Conversely, interference emerges from the similarity among intervening materials, thereby impeding the dependent's integration at the head. Therefore, it is not clear yet if locality and interference are two alternative explanations or do both the factors operate independently? Vasishth characterizes this point as the *locality-interference* debate and advocates an empirical investigation to disentangle their relative impact.

Our work addresses this gap in the literature by investigating the relationship between locality and interference effects using observational data from naturally occurring sentences in Hindi. We test the hypothesis that interference effects significantly predict dependency length within a sentence, while we statistically control for predictability measures and givenness discourse considerations as potential confounds. We quantify *locality effects* using dependency length, which is inspired from integra-

Proceedings of the Society for Computation in Linguistics (SCiL) 2024, pages 240-256.
Irvine, California, June 27-29, 2024

tion costs posited by DLT, and *interference effects* using semantic similarity among preverbal heads and metrics based on case density (number of case markers per sentence). Hindi, an Indo-Aryan language within the Indo-European family, exhibits a robust case-marking system and flexible word order with Subject-Object-Verb (SOV) as its canonical structure (Kachru, 2006):

(1) amar ujala-**ko** yah sukravar-**ko** daak-**se**
Amar Ujala-ACC it friday-on post-INST
prapt hua
receive be.PST.SG

Amar Ujala received it by post on Friday.

Our hypothesis is inspired by a substantial body of evidence from the studies on dependency locality and interference effects across languages (Staub, 2010; Vasishth, 2011; Vasishth and Drenhaus, 2011; Jäger et al., 2015; Ranjan et al., 2019; Stone et al., 2020). These studies suggest that language processor concurrently exhibits both dependency distance and interference minimization to overcome pressure related to working memory load. In contrast to English, verb-final languages like Hindi lack strong empirical support for locality effects (Vasishth and Lewis, 2006; Husain et al., 2014, 2015; Ranjan et al., 2022a,b; Ranjan and van Schijndel, 2024, cf. Ranjan and von der Malsburg, 2023; Ranjan and von der Malsburg, 2024) and numerous instances of anti-locality effects have been reported. Levy (2008) demonstrated that anti-locality patterns can be effectively explained using expectation-based accounts such as surprisal theory, with a view that introduction of more intervening words sharpens expectations at the verbal integration site within the sentence, thereby aiding comprehension. However, Vasishth and Lewis (2006) proposed an alternative unified explanation that explains both locality and anti-locality effects in Hindi. Based on Adaptive Control of Thought—Rational (ACT-R) framework (Anderson and Paulson, 1977), Vasishth and colleagues suggest that these effects can either be on account of activation decay in memory (anti-locality) or due to interference of intervening elements (locality). Subsequent studies in the literature advocate for a comprehensive theory of syntactic complexity encompassing both expectation-based and memory-based theories (Levy et al., 2013; Husain et al., 2014; Ranjan et al., 2022b).

To test the stated hypothesis, we deploy data from Hindi-Urdu Treebank (Bhatt et al., 2009, HUTB) corpus containing written text from newswire domain. We compute sentence-level cognitive measures: dependency length as *locality* measure, trigram surprisal and PCFG surprisal as *predictability* measures, semantic similarity and case-marker features as *interference* measures, and lastly, information status as *discourse* measure. We then compute their averaged values throughout each sentence and subsequently, fit a linear regression model to predict the average dependency length of corpus sentences. This approach is well motivated from the previous studies that have tried explaining dependency locality in natural languages (Futrell, 2019; Sharma et al., 2020). Our results demonstrate that similarity-based interference (as modelled using semantic similarity of preverbal dependents) is a significant predictor of dependency length for the entire dataset as well as for specific constructions of interest, *viz.,* non-canonical OSV orders and conjunct verbs. Our analysis of different bins of increasing dependency lengths consistently revealed higher occurrences of case markers and semantically similar elements. Overall, our findings suggest that dependency length, indicative of locality effects, is modulated by more general memory interference effects. Finally, we discuss the implications of our findings for the theories of interference in comprehension.

Our main contribution is that we provide an empirical basis for the Vasishth (2011)'s theoretical proposal on the *interference-locality* debate using broad-coverage study in Hindi. Moreover, we make use of naturally occurring sentences as opposed to artificially crafted sentences in a controlled laboratory experiments, thereby providing broader significance to the presented findings. Finally, our work extends the scope of psycholinguistic research beyond an anglocentric focus, allowing for a broader typological base for theory development (Jaeger and Norcliffe, 2009; Norcliffe et al., 2015).

2 Measures of Processing Difficulty

In this section, we present details of the theories alluded to in the introduction and measures derived from them for our experiments.

2.1 Locality measure

Dependency Locality Theory (Gibson, 1998, 2000, DLT) has been successfully shown to empirically predict the source of comprehension difficulty within sentences. DLT quantifies memory load

during sentence comprehension in two ways: a) by counting the number of new discourse referents introduced between heads and dependents (INTEGRATION COST), b) by counting the number of incomplete dependencies (upcoming heads) at a given word that needs to be stored in memory (STORAGE COST). The theory assumes *decay* as its underlying cognitive construct, suggesting that as the distance between head-dependent pairs increases, the information fades away, leading to *forgetting*. As a result, language users strive to minimize the distance between syntactically related units in the sentence. The aforementioned DLT metrics have been successfully shown to account for greater complexity (measured in terms of reading times) of object relative clauses compared to subject relatives in English, and more generally have been influential in shaping natural languages (Liu, 2008; Futrell et al., 2015, 2020; Ranjan and von der Malsburg, 2023, 2024).

Inspired by Gibson's word-by-word integration cost, we define the dependency length as the count of intervening words between head and dependent units within a dependency graph (Temperley, 2007). Figure 4 in Appendix A illustrates the calculation of dependency length for Example sentence 1. The average dependency length for each sentence was computed by dividing the total dependency length (sum of per-word dependency lengths) by the sentence length (count of words in the sentence).

2.2 Interference measure

Previous studies on *similarity-based interference* have quantified the comprehension difficulty by examining the intervening materials between syntactically related units, and the elements retrieved at the integration site in the sentence (Gordon et al., 2006, 2002, 2001; Van Dyke and Lewis, 2003; Van Dyke and McElree, 2006; Jager et al., 2017; Lewis, 1998; Lee et al., 2005; Van Dyke and Johns, 2012). These studies reported that comprehender tends to make more retrieval errors when they experience similar items that need to be retrieved from the working memory. This happens because similar items share common feature attributes in the memory and cause undesired confusion while retrieving the correct target element. For instance, both Traxler et al. (2002) and Staub (2010) contend that the processing difficulties associated with English relative clause examples shown in Figure 1 can be explained using interference effects in addition to distance effects. Sentences containing

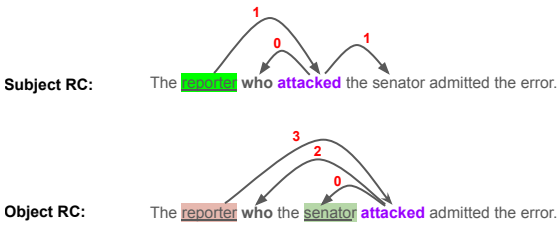


Figure 1: Interference-locality debate

objective relative clauses (ORC) are typically more challenging to process and comprehend than subject relative clauses (SRC) due to their greater dependency length. However, this behavior can also be attributed to interference phenomena. The ORC structures involve more interference compared to SRC structures. The nouns like 'reporter' and 'senator' in ORCs, both falling into the same category, induce greater interference during retrieval at the inner verb 'attacked,' unlike their SRC counterparts.

As pointed previously, the interference explanation has its independent motivation from the theory of working memory retrieval—*cue-based retrieval model* derived from the ACT-R cognitive architecture, which includes *decay, re-activation, cue-matching,* and *interference* (Lewis and Vasishth, 2005; Anderson et al., 2004). Contrasting DLT's decay and retrieval interference mechanisms, Vasishth (2011) expounds that decay is a lack of focused attention over *to-be-retrieved* information when the processor is engaged with interpreting the intervening elements. On the contrary, interference is about attention being shared unnecessarily to multiple units of information leading to *unavailability* of the required information. Therefore, under this logical space, Vasishth posits that DLT and interference could be the two manifestations of the same phenomenon that we intend to probe in the current work.

We operationalize these insights by estimating the semantic similarity between adjacent preverbal heads (directly linked to the main-verb) in the sentence. The similarity scores $sim(d_i)$ for each preverbal head was estimated by computing the cosine similarity (Salton, 1972) of the target head word with the adjacent head-word (see Equation 1). For instance, consider the sentence Example 1 and corresponding dependency graph shown in Figure 2, we computed the cosine similarity between following pairs: (ujala, yah); (yah, sukravar); (sukravar, daak); (daak, prapt).

$$sim(d_i) = \frac{wv(d_i) \cdot wv(d_{i+1})}{\|wv(d_i)\| \ \|wv(d_{i+1})\|} \quad (1)$$

$wv(d)$ and $wv(d_{i+1})$ denote the word vectors of the head-word d_i and d_{i+1}, respectively. These word vectors were obtained from the pre-trained *word2vec* model for Hindi (Grave et al., 2018). We then calculated the average semantic similarity by summing the similarity score over all the preverbal heads and then divided it by total preverbal heads in the sentence. As a sanity check, we deployed this method to understand how well the cosine similarity predicts the human judgment ratings[1] of word-similarity in Hindi (Bhatia et al., 2021). Notably, we found that the cosine similarity had a Spearman's rank correlation of 0.75 with the human judgment ratings, signifying its capability to model interference effects. Prior work has shown that cosine similarity metric is effective in modeling interference phenomenon (Sharma et al., 2020; Smith and Vasishth, 2020) and reading time (Frank, 2017; Salicchi et al., 2021) but also see Merlo and Ackermann (2018) and De Deyne et al. (2016).

2.3 Predictability measures

Surprisal theory (Hale, 2001; Levy, 2008) posits that language knowledge (or grammar) is probabilistic in nature, shaped by prior linguistic experiences and language learning. The cited authors suggest that the cognitive effort required to comprehend a word w_k in its context can be quantified using Shannon's information-theoretic measure of the log of the inverse of word's conditional probability given the preceding context ($w_{1...k-1}$):

$$\textbf{Effort } (\mathbf{w_k}) \propto \log \frac{1}{P(w_k|w_{1...k-1})} \quad (2)$$

$$\mathbf{S_k} = -\log P(w_k|w_{1...k-1}) = \log \frac{P(w_1...w_{k-1})}{P(w_1...w_k)} \quad (3)$$

Subsequently, the *surprisal* of the k^{th} word, w_k, is defined as the negative log probability of w_k given the preceding intra-sentential ($w_{1...k-1}$) context (see Equation 3). These probabilities can be computed either over word sequences or syntactic configurations and reflect the information load (or predictability) of w_k. The theory is supported by a large body of empirical evidence from behavioural as well as broad-coverage corpus data (Demberg and Keller, 2008; Boston et al., 2008; Roark et al., 2009; Agrawal et al., 2017; Dammalapati et al., 2021; Ranjan et al., 2022a). The cited studies suggest that words with high surprisal tend to have high reading time. In this work, we estimated the per-word lexical trigram surprisal using n-gram language model and syntactic surprisal using probabilistic context-free grammar (PCFG) parser as described below.

- **Trigram surprisal**: We estimated the lexical surprisal for each word in the sentence using a 3-gram language model (LM) trained on the written section of the EMILLE Hindi Corpus (Baker et al., 2002), consisting of 1 million sentences, using the SRILM toolkit (Stolcke, 2002) with Good-Turing discounting.

- **PCFG surprisal**: We estimated the syntactic surprisal of each word in a sentence using the Berkeley latent-variable PCFG parser[2] (Petrov et al., 2006). We trained the parser using 12000 phrase structure trees obtained by converting Bhatt et al.'s HUTB dependency trees into constituency trees using the approach described in Yadav et al. (2017). We adopted the 5-fold cross-validation approach to compute the surprisal scores from the PCFG parser i.e., surprisal for each test sentence was estimated by training a PCFG LM on four folds of the phrase structure trees and then testing on a fifth held-out fold.

For both measures above, we computed the average surprisal for each sentence by summing the word-level surprisal of all the words in the sentence and then divided it by sentence length.

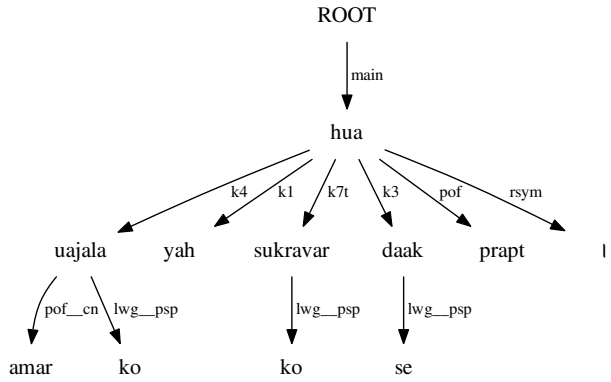
2.4 Information status measure

Languages are known to obey *given-before-new* principle by producing elements that are previously mentioned in the discourse prior to the new content in the sentence (Clark and Haviland, 1977; Chafe, 1976; Kaiser and Trueswell, 2004).

In this work, we annotate each sentence in our dataset with GIVEN-NEW ordering. The preverbal subject and object constituents of a target sentence were assigned *Given* tag if any content word within the phrase appeared in the preceding sentence in the corpus or the head of the phrase was a pronoun; else, the phrases were tagged as *New*. We then assigned scores to each ordering as per the following

[1] https://github.com/ashwinivd/similarity_hindi

[2] 5-fold cross-validated parser training and testing F1-score metrics were 90.82% and 84.95%, respectively.



(a) Dependency tree

Label	Dependency relation
Invariant syntactic relations	
k1	subject/agent
k2	object/patient
k3	instrument
k4	object/recipient
k7t	location in time
Complex predicate relation	
pof	parts of conjunct verb
pof_cn	parts of compound noun
Local word group (lwg)	
lwg_psp	postposition
lwg_vaux	auxilliary verb
Symbols	
rsym	symbol relation

(b) Dependency relations

Figure 2: HUTB dependency tree and corresponding dependency relation labels for Example 1

scheme: a) New-Given = -1 b) Given-New = +1 c) Given-Given and New-New = 0. See Appendix B for an illustration.

2.5 Case markers

In Hindi, case markers are identified as postpositions,[3] crucial for conveying grammatical relationships within sentences (Kachru, 2006; Agnihotri, 2007). Case markers influence comprehension via the mechanisms of either predictability (Avetisyan et al., 2020) or interference effects (Tily, 2010; Ranjan et al., 2019). They have been shown to predict the upcoming verb (Husain et al., 2014; Grissom II et al., 2016) and effectively reduce interference in memory by correctly distinguishing between subject and objects, thereby enhancing retrieval at the verb. Moreover, there is extensive research on case marker interference in sentence comprehension in SOV languages like Japanese (Lewis and Nakayama, 2001), Korean (Lee et al., 2005) and Hindi (Vasishth, 2003). Inspired by these insights, we compute following measures to quantify the distinct effects of case markers:

- **Case density**: Ratio of the number of case markers to the word counts in the sentence.
- **Same case bigrams**: Total number of identical case marker sequences associated with pairs of adjacent preverbal constituents.

[3]Table 4 in Appendix C outlines Hindi case markers and their functions.

3 Data and Methods

Our dataset consists of 1996 declarative sentences with well-defined subject and object phrases from the Hindi-Urdu Treebank (HUTB) corpus of the written text belonging to the newswire domain (Bhatt et al., 2009). For each sentence, we first compute average values of various cognitive measures. We then fit a linear regression model to predict the average dependency length of sentences in our dataset, using the remaining cognitive measures discussed in the preceding section as independent variables. All the independent variables were normalized to z-scores, *i.e.*, the predictor's value (centered around its mean) was divided by its standard deviation. We used the `glm` function in R to perform our regression experiments expressed using the `glm` equation below:

$$\text{Dependency length} \sim \begin{cases} \text{similarity + same case bigram +} \\ \text{trigram surprisal + PCFG surprisal +} \\ \text{case density + IS score} \end{cases} \quad (4)$$

The pairwise Pearson correlation coefficients among these measures are shown in Appendix D, Figure 5. We observed a moderate correlation of 0.31 between dependency length and semantic similarity, whereas the remaining predictors exhibited weaker correlations with dependency length.

4 Results

In this section, we test the hypothesis that locality effects as captured by dependency length are

Predictor	$\hat{\beta}$	$\hat{\sigma}$	t
Intercept	1.87	0.016	118.81
IS Score	-0.04	0.013	-3.13
PCFG surprisal	0.06	0.019	2.99
3.gram surprisal	-0.05	0.019	-2.53
same case bigram	-0.01	0.015	-0.27
case density	0.09	0.015	5.94
similarity	0.19	0.015	13.16
PCFG x 3.gram surp	-0.05	0.011	-4.08

Table 1: Linear regression model predicting average dependency length on full data set (1996); significant predictors denoted in bold

reducible to more general memory-interference effects as captured by semantic similarity and case marker features while controlling for expectation-based measures and discourse givenness. We, therefore, expect that interference-based features should have positive regression coefficients in the regression model. In other words, sentences with longer dependency length in Hindi should exhibit greater interference effects as quantified by more case markers, and interfering noun phrases (NPs) with similar featural attributes. Our results are discussed in the remaining subsections.

4.1 Predicting dependency length

We first performed regression analyses on the entire data set to investigate the influence of predictability, interference, and givenness measures on the dependency length. We then reported the statistical analyses on different bins of dependency lengths. Table 1 displays the regression results over the entire data set. All interference measures other than same-case bigram counts are significant predictors of dependency length, thus validating our proposed hypothesis. The positive regression coefficient for semantic similarity indicates that with every unit increase in its score, the value of dependency length also increases, thus shedding light on how locality effects are modulated in Hindi. Moreover, adding similarity score into a model containing all other predictors significantly improved the fit of our regression model (χ^2 = 166.75; p $<$ 0.001). The positive regression coefficient of case density suggests that sentences with more case markers tend to have higher dependency length, consequently highlighting both predictability and interference effects of case markers as discussed in the comprehension literature (Husain et al., 2014; Avetisyan et al., 2020). The negative regression coefficient of the IS score suggests that sentences with longer dependency length have NEW-GIVEN ordering. Finally, syntactic PCFG and lexical trigram surprisal measures have positive and negative regression coefficients, respectively, with a significant interaction between the two while predicting dependency length, suggesting that syntactically complex sentences may have more probable word sequences.

We investigated the relationship between case marking and dependency length in more detail. For each of the 25 most frequent verbs in the HUTB, we plotted the average case density of all sentences having that verb as the root of the sentence against the average dependency length of those sentences (refer to Figure 3). Many of the high-frequency verbs have an average dependency length greater than the average value for all verbs. Such verbs also have higher a case density value compared to the average value for verbs in the entire dataset. Almost all these verbs are perfective verbs which are transitive in nature. In Hindi, it is well known that the ergative marker *ne* indicates the presence of an upcoming transitive verb with perfective aspect (Choudhary et al., 2009; Husain et al., 2014). Vice-versa, we observe verbs having lower-than-average values for both average dependency length and case density. The verbs in this set are mostly auxiliary verbs like *hai* and *tha*. Thus root verbs with longer dependencies are associated with dependents marked by more case markers and conversely, verbs involved in shorter dependencies are linked to fewer case-marked heads.

In a recent work, Ranjan et al. (2022b) conjectured that the presence of semantically similar noun phrases and case markers are a strong factor potentially overriding the pressure for dependency length minimization in determining Hindi word-ordering choices. They further observed that the dependency length was an effective predictor of human choices only at very high dependency length values, consistent with prior work in the literature denoting interference effects in long-distance dependency resolution (Van Dyke and McElree, 2006; Van Dyke, 2007; Lewis, 1996). To substantiate these insights further, we fitted separate regression models containing all the predictors on four different bins of the dependency length; we report the results in Table 2. Our results at high dependency length suggest that long-distance dependency resolution is indeed driven by interference effects. However, case-marker effects are significant in all bins except

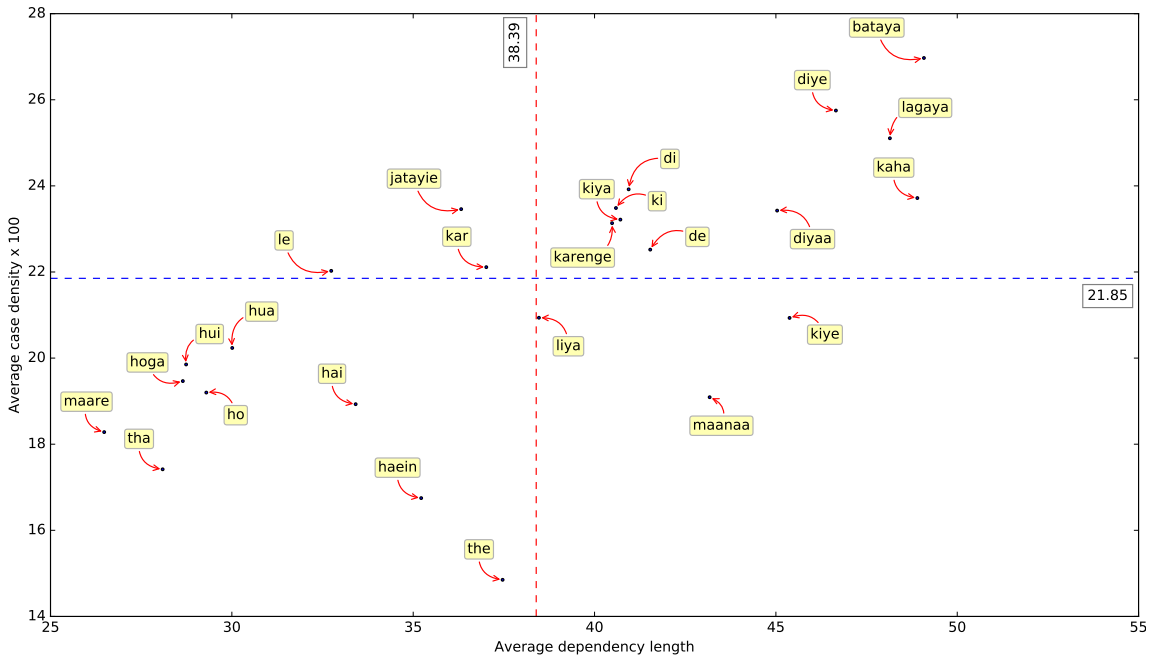


Figure 3: Average case density and dependency length for the 25 most-frequent HUTB verbs (average dependency length and case density values for the entire dataset depicted as dotted lines parallel to X and Y axes respectively)

the final one, a finding which requires more exploration factoring in the possibility of the interplay between case-based facilitation (Logačev and Vasishth, 2012) and cosine similarity. In other words, interference (whether proactive or retroactive) on account of a greater number of similar intervening items might be the working mechanism behind the processing difficulty postulated for longer dependency distances. This finding is also consistent with prior work in the literature, which argues that decay, the underlying cognitive construct behind locality, does not have robust empirical evidence supporting it (Engelmann et al., 2019; Oberauer and Lewandowsky, 2013, 2014; Stone et al., 2020; Berman et al., 2009, cf. Hardt et al., 2013).

4.2 Construction Analysis

In this section, we examined two Hindi syntactic constructions studied in the sentence processing literature, *viz.,* object-fronted (non-canonical) word orders, *i.e.,* direct (DO) and indirect object (IO) fronting (Vasishth, 2004), and conjunct verb constructions (Husain et al., 2014).

4.2.1 Non-canonical word order

We analyzed HUTB sentences that displayed Object-Subject-Verb (OSV) order, as illustrated in Example 1. The fronted objects could be direct or indirect. Vasishth (2004) showed that dependency length effectively predicts the processing difficulties associated with OSV orders. Husain et al. (2014) demonstrated that sentences with conjunct verbs exhibit anti-locality effects. Table 3 displays the regression results for DO-/IO-fronted subsets and conjunct-verb constructions. For both DO- and IO-fronted sentences, our results revealed that the similarity measure is the only feature that significantly predicts the dependency length, while other effects are non-significant. The regression coefficient is also in the expected direction. Vasishth (2004) in his investigation of OSV order in Hindi reported that unlike IO-fronted sentences, the DO-fronted sentences still remained difficult to comprehend when provided with appropriate discourse context as compared to their canonical counterparts. He attributed the difficulty to greater dependency length, and thereby greater memory load, associated with DO-fronted sentences.

Example 2 illustrates a DO-fronted sentence from our dataset where all preverbal heads (*gundon (henchmen), hathiyaar (wepons), police, kshetra (area), giraftaar (arrest)*) directly linked to the main verb (*kiya*) are highlighted. This non-canonical sentence exhibits a greater dependency length (1.91) than average (1.83), indicating that

Predictor	**dl <= 1.36** (#495)	**1.36 <dl <= 1.80** (#555)	**1.80 <dl <= 2.20** (#448)	**dl >2.20** (#498)
Intercept	1.13	1.59	1.99	2.70
case density	0.03	0.01	0.02	NS
similarity	0.06	NS	0.01	0.07
IS Score	NS	-0.01	NS	-0.04
PCFG x 3.gram surprisal	-0.02	NS	NS	NS

Table 2: Four different regression models predicting average dependency length in binned data sets with the no. of data points in each indicated in column headers; column values represent regression coefficient of different predictors in the regression model; **dl = Average dependency length**; Bin-wise number of data points in parentheses; trigram and PCFG surprisal, and same case bigram features not shown as they are not significant (NS) in the models; Avg dl (Min, 1^{st} Quartile, Mean, 3^{rd} Quartile, Max) = 0.37, 1.36, 1.83, 2.20, 6.20

Predictor	$\hat{\beta}$	$\hat{\sigma}$	t
DO-FRONTED SUBSET			
Intercept	1.66	0.053	31.32
similarity	0.27	0.052	5.11
IO-FRONTED SUBSET			
Intercept	1.78	0.065	27.62
similarity	0.22	0.054	4.03
CONJUNCT-VERB SUBSET			
Intercept	1.93	0.021	92.24
PCFG surprisal	0.08	0.027	3.32
case density	0.09	0.022	4.13
similarity	0.18	0.019	9.34

Table 3: Linear regression model predicting average dependency length on DO-fronted (133), IO-fronted (101), and conjunct-verb (1158) data sets; significant predictors denoted in bold; non-significant predictors not shown here but see Appendix E for full details)

OSV sentences generally impose a higher memory load. Additionally, this sentence also has a higher semantic similarity (0.18) than the average (0.08) due to confusability among the four aforementioned preverbal head nouns (two animate and two inanimate nouns) when retrieved at the main verb. Therefore, these results suggest that the observed difficulty, as captured by dependency length in these OSV constructions, can be effectively explained by examining the semantic similarity among the preverbal heads within a sentence.

(2) [kukhyaat sargana chhota rajan giroh-ke
Infamous gangster chhota raja gang-GEN
chaar **gundon**-ko]$_{DO}$ **hathiyaar** sahit
four goons-ACC weapons along with
[**police**-ne]$_{S}$ sehar-ke uttari paschami **kshetra**-se
police-ERG city-GEN north-western area-LOC
[**giraftaar**]$_{POF}$ kiya
arrest do.PST.PFV.SG.M

The police arrested four henchmen from the notorious gangster Chhota Rajan gang, along with weapons, from the north-western area of the city.

4.2.2 Conjunct verbs

We focused on sentences in the corpus that contained noun-verb complex predicates, commonly referred to as conjunct verbs (Kachru, 1982; Butt, 1995; Mohanan, 1994). A conjunct verb consists of a complex predicate composed of a noun and a subsequent verb; these are annotated with the POF dependency relation in the HUTB corpus (See Example 5 in Appendix F).

For conjunct verb constructions (bottom block in Table 3), our analysis revealed that semantic similarity, case density, and PCFG surprisal emerged as significant predictors of dependency length. Notably, these predictors displayed positive regression coefficients, affirming the validity of our proposed hypothesis. In a self-paced reading study, Husain et al. (2014) found no significant reading time differences at the final verb in non-compositional sequences (e.g., *khyaal rakhna*) when the noun-verb distance increased with intervening adverbials. They observed locality effects only in simple predicates where the final verb was not predictable from its noun counterpart (e.g., *guitar rakhna*). Table 8 in Appendix G depicts construction-wise average feature values. In comparison to sentences with non-canonical word orders (11.72%), the conjunct verbs constructions (58.02%) are very frequent in our dataset and have higher average dependency length, number of constituents, similarity, and case density. Thus the differential impact of various features across the three constructions can be explained by the variation in these basic properties.

Thus, these construction-specific results further corroborate the view that the underlying reason behind locality effects may not be decay but rather more general memory retrieval and interference

effects as captured by semantic similarity.

5 Discussion

Our results show that our proposed interference measures, *viz.* semantic similarity and case density, model locality effects (as captured by dependency length) in Hindi. Their effects remain consistent at high dependency length, suggesting that dependency locality may be driven not just by decay of information but also by proactive and retroactive interference. Our findings also highlight that long dependencies involve a greater proportion of case markers. This reinforces the idea that within a natural corpus, the processing load on account of longer dependencies is due to increased memory load caused by interference and predictability effects arising from case markers. Additionally, we found that sentences with longer dependency lengths consistently exhibited high PCFG syntactic surprisal but low lexical trigram surprisal, with a notable interaction between the two. This hints at the possibility that syntactically complex sentences (as denoted by longer dependencies or greater syntactic surprisal) perhaps feature more probable word sequences, potentially mitigating the memory load. Finally, we noted that the interference due to same-case bigrams (*i.e.,* adjacent NPs marked with the same case marker) is insignificant in predicting dependency length, and further analyses confirmed that their effects were already accounted by our semantic similarity measure.

For non-canonical OSV orders, we found that semantic similarity was the only significant positive predictor of dependency length. In contrast, for the sentences with conjunct verbs, in addition to PCFG surprisal, both semantic similarity and case density were significant predictors of dependency length, thereby validating our initial hypothesis across both constructions. These findings provide further insights for retrieval interference as an explanatory mechanism underlying locality effects that have been observed across various constructions in Hindi (Vasishth, 2004; Vasishth and Lewis, 2006; Husain et al., 2014; Ranjan et al., 2022b; Ranjan and von der Malsburg, 2023, 2024). Our results corroborate previous conjectures suggesting that temporal decay alone may not be the only explanation for the observed locality effects (Berman et al., 2009; Oberauer and Lewandowsky, 2013, 2014; Engelmann et al., 2019; Stone et al., 2020; Ranjan et al., 2022b, cf. Hardt et al., 2013).

More recently, Ranjan and van Schijndel (2024) in their extensive study of non-canonical word orders in a corpus of naturally occurring Hindi text demonstrated that discourse expectations captured by surprisal estimates from neural language models fine-tuned over preceding sentential context primarily govern the production of Hindi sentences. Notably, they report that discourse-enhanced surprisal entirely subsumes the impact of dependency length minimization effects in predicting Hindi OSV orders. Future work needs to investigate how interference, locality and surprisal jointly shape natural languages and human behaviour.

Our results provide an empirical basis for Vasishth (2011)'s theoretical proposal, where it was argued that locality and interference could be different manifestations of the same phenomenon. Vasishth contends that dependency locality instantiates the concept of decay in the form of dependency distance by counting the number of intervening discourse referents. In contrast, interference has no notion of memory limitation (storage) and only exhibits its effect through syntactic and semantic integration during retrieval processes, which get affected by the nature, quality, and specific content of information stored in the memory. Our semantic similarity measure (as quantified by cosine similarity among preverbal heads) significantly predicts the dependency length in Hindi, possibly indicating that interference effects may subsume the predictions of dependency locality. Therefore, we propose that interference effects also need to be factored in while developing a comprehensive theory of sentence processing.

As a part of future work, we plan to investigate the role of interference in presence of various other factors such as surprisal, locality, and discourse considerations. We also intend to tease apart the distance *vs.* interference effects by studying the nature of intervening material between the head and dependent units in a more controlled setup. Finally, future work should explore the impact of these factors on other languages to make cross-linguistic generalizations, as well as on language production using spoken datasets.

In sum, our results suggest a significant association between locality and interference effects, perhaps indicating that locality might be a surface phenomenon whose internal workings are driven by interference during memory retrieval.

Acknowledgements

We thank Marten van Schijndel, Shravan Vasishth, and the audiences of the Sentence Processing Colloquium at University of Potsdam and Cornell's C.Psyd group for their insightful comments on this work. We would also like to express our gratitude to the anonymous reviewers of CogSci-2022, HSP-2023, and SCiL-2024 for their helpful suggestions and feedback. Finally, the last two authors acknowledge the extramural funding provided by the Department of Science and Technology of India through the Cognitive Science Research Initiative (project no. DST/CSRI/2018/263).

References

Rama Kant Agnihotri. 2007. *Hindi: An Essential Grammar*. Essential Grammars. Routledge.

Arpit Agrawal, Sumeet Agarwal, and Samar Husain. 2017. Role of expectation and working memory constraints in Hindi comprehension: An eyetracking corpus analysis. *Journal of Eye Movement Research*, 10(2).

John R. Anderson, Daniel Bothell, Michael D. Byrne, Scott Douglass, Christian Lebiere, and Yulin Qin. 2004. An Integrated Theory of the Mind. *Psychology Review*, 111(4):1036–1060.

John R. Anderson and Rebecca Paulson. 1977. Representation and retention of verbatim information. *Journal of Verbal Learning and Verbal Behavior*, 16(4):439 – 451.

Serine Avetisyan, Sol Lago, and Shravan Vasishth. 2020. Does case marking affect agreement attraction in comprehension? *Journal of Memory and Language*, 112:104087.

AD Baddeley and G Hitch. 1977. Recency re-examined. s. dornic (ed.) attention and performance (vol. 6, pp. 647-667).

Paul Baker, Andrew Hardie, Tony McEnery, Hamish Cunningham, and Robert Gaizauskas. 2002. Emille: a 67-million word corpus of indic languages: data collection, mark-up and harmonization. In *Proceedings of LREC 2002*, pages 819–827. Lancaster University.

Marc G Berman, John Jonides, and Richard L Lewis. 2009. In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(2):317.

Kushagra Bhatia, Divyanshu Aggarwal, and Ashwini Vaidya. 2021. Fine-tuning distributional semantic models for closely-related languages. In *Proceedings of the Eighth Workshop on NLP for Similar Languages, Varieties and Dialects*, pages 60–66, Kiyv, Ukraine. Association for Computational Linguistics.

Rajesh Bhatt, Bhuvana Narasimhan, Martha Palmer, Owen Rambow, Dipti Misra Sharma, and Fei Xia. 2009. A multi-representational and multi-layered treebank for Hindi/urdu. In *Proceedings of the Third Linguistic Annotation Workshop*, ACL-IJCNLP '09, pages 186–189, Stroudsburg, PA, USA. Association for Computational Linguistics.

Marisa Ferrara Boston, John Hale, Reinhold Kliegl, Umesh Patil, and Shravan Vasishth. 2008. Parsing costs as predictors of reading difficulty: An evaluation using the potsdam sentence corpus. *Journal of Eye Movement Research*, 2(1).

Miriam Butt. 1995. *The structure of complex predicates in Urdu*. Center for the Study of Language (CSLI).

Miriam Butt and Tracy Holloway King. 1996. Structural topic and focus without movement. In Miriam Butt and Tracy Holloway King, editors, *Proceedings of the First LFG Conference*. CSLI Publications, Stanford.

Wallace Chafe. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. *Subject and topic*.

Kamal Kumar Choudhary, Matthias Schlesewsky, Dietmar Roehm, and Ina D. Bornkessel-Schlesewsky. 2009. The N400 as a correlate of interpretively relevant linguistic rules: Evidence from Hindi. *Neuropsychologia*, 47(13):3012–3022.

H. H. Clark and S. E. Haviland. 1977. Comprehension and the Given-New Contract. In R. O. Freedle, editor, *Discourse Production and Comprehension*, pages 1–40. Ablex Publishing, Hillsdale, N. J.

Samvit Dammalapati, Rajakrishnan Rajkumar, Sidharth Ranjan, and Sumeet Agarwal. 2021. Effects of duration, locality, and surprisal in speech disfluency prediction in english spontaneous speech. In *Proceedings of the Society for Computation in Linguistics*, volume 4, page 10.

Simon De Deyne, Amy Perfors, and Daniel J Navarro. 2016. Predicting human similarity judgments with distributional models: The value of word associations. In *Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers*, pages 1861–1870.

Vera Demberg and Frank Keller. 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2):193–210.

Herman Ebbinghaus. 1885. *Über das Gedächtnis: Untersuchungen zur experimentellen Psychologie*. Duncker & Humblot.

Hermann Ebbinghaus. 2013. Memory: A contribution to experimental psychology. *Annals of neurosciences*, 20(4):155.

Felix Engelmann, Lena A. Jager, and Shravan Vasishth. 2019. The effect of prominence and cue association on retrieval processes: A computational account. *Cognitive Science*, 43(12):e12800.

Stefan L Frank. 2017. Word embedding distance does not predict word reading time. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society (CogSci)*, pages 385—-390. Cognitive Science Society, Austin.

Richard Futrell. 2019. Information-theoretic locality properties of natural language. In *Proceedings of the First Workshop on Quantitative Syntax (Quasy, SyntaxFest 2019)*, pages 2–15, Paris, France. Association for Computational Linguistics.

Richard Futrell, Roger P Levy, and Edward Gibson. 2020. Dependency locality as an explanatory principle for word order. *Language*, 96(2):371–412.

Richard Futrell, Kyle Mahowald, and Edward Gibson. 2015. Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, 112(33):10336–10341.

Edward Gibson. 1998. Linguistic complexity: Locality of syntactic dependencies. *Cognition*, 68:1–76.

Edward Gibson. 2000. Dependency locality theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita, and Wayne O'Neil, editors, *Image, Language, brain: Papers from the First Mind Articulation Project Symposium*. MIT Press, Cambridge, MA.

Peter C. Gordon, Randall Hendrick, and Marcus Johnson. 2001. Memory interference during language processing. *Journal of Experimental Psychology: Learning Memory and Cognition*, 27(6):1411–1423.

Peter C. Gordon, Randall Hendrick, Marcus Johnson, and Yoonhyoung Lee. 2006. Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning Memory and Cognition*, 32(6):1304–1321.

Peter C. Gordon, Randall Hendrick, and William H. Levine. 2002. Memory-load interference in syntactic processing. *Psychological Science*, 13(5):425–430. PMID: 12219808.

Édouard Grave, Piotr Bojanowski, Prakhar Gupta, Armand Joulin, and Tomáš Mikolov. 2018. Learning word vectors for 157 languages. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*.

Alvin Grissom II, Naho Orita, and Jordan Boyd-Graber. 2016. Incremental prediction of sentence-final verbs: Humans versus machines. In *Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning*, pages 95–104. Association for Computational Linguistics.

John Hale. 2001. A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics on Language technologies*, NAACL '01, pages 1–8, Pittsburgh, Pennsylvania. Association for Computational Linguistics.

Oliver Hardt, Karim Nader, and Lynn Nadel. 2013. Decay happens: the role of active forgetting in memory. *Trends in cognitive sciences*, 17(3):111–120.

Samar Husain, Shravan Vasishth, and Narayanan Srinivasan. 2014. Strong expectations cancel locality effects: Evidence from Hindi. *PLOS ONE*, 9(7):1–14.

Samar Husain, Shravan Vasishth, and Narayanan Srinivasan. 2015. Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus. *Journal of Eye Movement Research*, 8(2).

T. Florian Jaeger and Elizabeth Norcliffe. 2009. The cross-linguistic study of sentence production: State of the art and a call for action. *Language and Linguistic Compass*, 3(4):866–887.

Lena A Jäger, Felix Engelmann, and Shravan Vasishth. 2015. Retrieval interference in reflexive processing: Experimental evidence from mandarin, and computational modeling. *Frontiers in psychology*, 6:617.

Lena A. Jager, Felix Engelmann, and Shravan Vasishth. 2017. Similarity-based interference in sentence comprehension: Literature review and bayesian meta-analysis. *Journal of Memory and Language*, 94:316 – 339.

Y. Kachru. 2006. *Hindi*. London Oriental and African language library. John Benjamins Publishing Company.

Yamuna Kachru. 1982. Conjunct verbs in hindi-urdu and persian. *South Asian Review*, 6(3):117–126.

Elsi Kaiser and John C Trueswell. 2004. The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2):113–147.

Sun-Hee Lee, Mineharu Nakayama, and Richard L. Lewis. 2005. Difficulty of processing Japanese and Korean center-embedding constructions. In M. Minami, H. Kobayashi, M. Nakayama, and H.Sirai, editors, *Studies in Language Science*, volume Volume 4, pages 99–118. Kurosio Publishers, Tokyo, Tokyo.

Roger Levy. 2008. Expectation-based syntactic comprehension. *Cognition*, 106(3):1126 – 1177.

Roger Levy, Evelina Fedorenko, and Edward Gibson. 2013. The syntactic complexity of russian relative clauses. *Journal of Memory and Language*, 69(4):461 – 495.

R. L. Lewis and S. Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29:1–45.

Richard L Lewis. 1996. Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of psycholinguistic research*, 25(1):93–115.

Richard L. Lewis. 1998. Interference in working memory: Retroactive and proactive interference in parsing. CUNY sentence processing conference.

Richard L. Lewis and Mineharu Nakayama. 2001. Syntactic and positional similarity effects in the processing of Japanese embeddings. In *Sentence Processing in East Asian Languages*, pages 85–113, Stanford, CA. CSLI.

Haitao Liu. 2008. Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science*, 9(2):159–191.

Pavel Logačev and Shravan Vasishth. 2012. Case matching and conflicting bindings interference. In Monique Lamers and Peter de Swart, editors, *Case, Word Order and Prominence: Interacting Cues in Language Production and Comprehension*, pages 187–216. Springer Netherlands, Dordrecht.

Paola Merlo and Francesco Ackermann. 2018. Vectorial semantic spaces do not encode human judgments of intervention similarity. In *Proceedings of the 22nd Conference on Computational Natural Language Learning*, pages 392–401.

Tara Mohanan. 1994. *Argument structure in Hindi*. Center for the Study of Language (CSLI).

Elisabeth Norcliffe, Alice C Harris, and T Florian Jaeger. 2015. Cross-linguistic psycholinguistics and its critical role in theory development: Early beginnings and recent advances.

Klaus Oberauer and Stephan Lewandowsky. 2013. Evidence against decay in verbal working memory. *Journal of Experimental Psychology: General*, 142(2):380–411.

Klaus Oberauer and Stephan Lewandowsky. 2014. Further evidence against decay in working memory. *Journal of Memory and Language*, 73:15–30.

Slav Petrov, Leon Barrett, Romain Thibaux, and Dan Klein. 2006. Learning accurate, compact, and interpretable tree annotation. In *Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics*, ACL-44, pages 433–440, Stroudsburg, PA, USA. Association for Computational Linguistics.

Sidharth Ranjan, Sumeet Agarwal, and Rajakrishnan Rajkumar. 2019. Surprisal and interference effects of case markers in Hindi word order. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 30–42, Minneapolis, Minnesota. Association for Computational Linguistics.

Sidharth Ranjan, Rajakrishnan Rajkumar, and Sumeet Agarwal. 2022a. Linguistic Complexity and Planning Effects on Word Duration in Hindi Read Aloud Speech. *In Proceedings of the Society for Computation in Linguistics (SCiL)*, 5:11.

Sidharth Ranjan, Rajakrishnan Rajkumar, and Sumeet Agarwal. 2022b. Locality and expectation effects in Hindi preverbal constituent ordering. *Cognition*, 223:104959.

Sidharth Ranjan and Marten van Schijndel. 2024. Does dependency locality predict non-canonical word order in Hindi? In *Proceedings of the 46th Annual Meeting of the Cognitive Science Society*, Rotterdam, Netherlands. Cognitive Science Society, Cognitive Science Society.

Sidharth Ranjan and Titus von der Malsburg. 2023. A bounded rationality account of dependency length minimization in Hindi. In *Proceedings of the 45th Annual Meeting of the Cognitive Science Society*, Sidney, Australia. Cognitive Science Society, Cognitive Science Society.

Sidharth Ranjan and Titus von der Malsburg. 2024. Work smarter...not harder: Efficient minimization of dependency length in SOV languages. In *Proceedings of the 46th Annual Meeting of the Cognitive Science Society*, Rotterdam, Netherlands. Cognitive Science Society, Cognitive Science Society.

Brian Roark, Asaf Bachrach, Carlos Cardenas, and Christophe Pallier. 2009. Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1*, EMNLP '09, pages 324–333, Stroudsburg, PA, USA. Association for Computational Linguistics.

Henry L Roediger III and Magdalena Abel. 2022. The double-edged sword of memory retrieval. *Nature Reviews Psychology*, 1(12):708–720.

Lavinia Salicchi, Alessandro Lenci, and Emmanuele Chersoni. 2021. Looking for a role for word embeddings in eye-tracking features prediction: Does semantic similarity help? In *Proceedings of the 14th International Conference on Computational Semantics (IWCS)*, pages 87–92, Groningen, The Netherlands (online). Association for Computational Linguistics.

Gerard Salton. 1972. A new comparison between conventional indexing (medlars) and automatic text processing (smart). *Journal of the American Society for Information Science*, 23(2):75–84.

C. E. Shannon. 1948. A mathematical theory of communication. *Bell system technical journal*, 27.

Kartik Sharma, Richard Futrell, and Samar Husain. 2020. What determines the order of verbal dependents in hindi? effects of efficiency in comprehension and production. In *Proceedings of the Workshop on Cognitive Modeling and Computational Linguistics*, pages 1–10.

Garrett Smith and Shravan Vasishth. 2020. A principled approach to feature selection in models of sentence processing. *Cognitive science*, 44(12):e12918.

Adrian Staub. 2010. Eye movements and processing difficulty in object relative clauses. *Cognition*, 116:71–86.

Andreas Stolcke. 2002. SRILM — An extensible language modeling toolkit. In *Proc. ICSLP-02*.

Kate Stone, Titus von der Malsburg, and Shravan Vasishth. 2020. The effect of decay and lexical uncertainty on processing long-distance dependencies in reading. *PeerJ*, 8:e10438.

David Temperley. 2007. Minimization of dependency length in written English. *Cognition*, 105(2):300–333.

Harry Tily. 2010. *The Role of Processing Complexity in Word Order Variation and Change*. Ph.D. thesis, Stanford University. Unpublished thesis.

Matthew J Traxler, Robin K Morris, and Rachel E Seely. 2002. Processing subject and object relative clauses: Evidence from eye movements. *Journal of memory and language*, 47(1):69–90.

Julie A. Van Dyke. 2007. Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33 2:407–30.

Julie A. Van Dyke and Clinton L. Johns. 2012. Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, 6(4):193–211.

Julie A. Van Dyke and R. L. Lewis. 2003. Distinguishing effects of structure and decay on attachment and repair: a retrieval interference theory of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, 49:285–413.

Julie A. Van Dyke and Brian McElree. 2006. Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2):157 – 166.

S. Vasishth. 2003. *Working Memory in Sentence Comprehension: Processing Hindi Center Embeddings*. Outstanding Dissertations in Linguistics. Taylor & Francis.

S. Vasishth. 2004. Discourse context and word order preferences in Hindi. *Yearbook of South Asian Languages*, pages 113–127.

Shravan Vasishth. 2011. Integration and prediction in head-final structures. In Yuki Hirose and Jerome L.Packard, editors, *Processing and Producing Head-final Structures*, pages 349–367. Springer Netherlands, Dordrecht.

Shravan Vasishth and Heiner Drenhaus. 2011. Locality in german. *Dialogue & Discourse*, 2(1):59–82.

Shravan Vasishth and Richard L. Lewis. 2006. Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4):767–794.

Himanshu Yadav, Ashwini Vaidya, and Samar Husain. 2017. Keeping it simple: Generating phrase structure trees from a Hindi dependency treebank. In *TLT*.

Victor H Yngve. 1960. A model and an hypothesis for language structure. *Proceedings of the American philosophical society*, 104(5):444–466.

Appendix

A Dependency length calculation

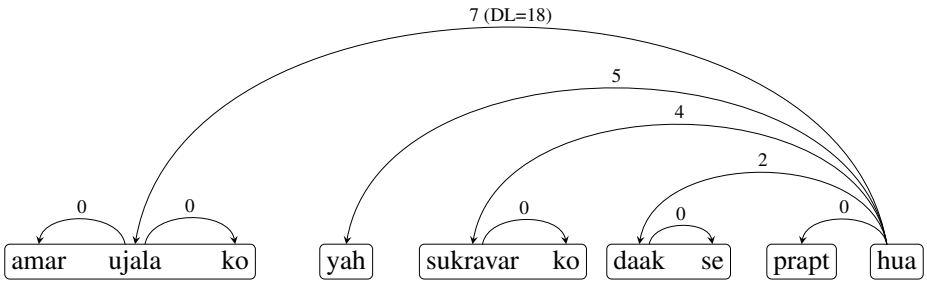


Figure 4: Calculation of dependency length in a dependency tree; Total dependency length (DL) of the structure indicated above the top arc; Word's dependency length is mentioned above each dependency arc

B Information Status Annotation

(3) **Preceding context sentence**

amar ujala-ki bhumika nispaksh rehti hai
Amar Ujala-GEN role unbiased remain be.PRS.SG

Amar Ujala's role remains unbiased.

(4) **Target Sentence**

a. [amar ujala-ko]$_O$ [**yah**]$_S$ *sukravar*-ko daak-se prapt hua [Given-Given = 0]
Amar Ujala-ACC **it** *friday*-on post-INST receive be.PST.SG

Amar Ujala received **it** by post on *Friday*.

In the above target example, the object phrase shares a content word "amar ujala" with the preceding context sentence. Therefore, the object phrase is assigned a GIVEN tag. Additionally, the subject phrase "yah" in the target sentence is a pronoun, so it is also assigned a GIVEN tag. As a result, the target sentence, overall, belongs to GIVEN-GIVEN ordering.

C Hindi Case Markers

Marker	**Case** (Gloss)	**Grammatical Function**
ϕ	nominative (NOM)	subject/object
ne	ergative (ERG)	subject
ko	accusative (ACC)	object
	dative (DAT)	subject/indirect object
se	instrumental (INS)	subject/oblique/adjunct
$ka/ki/ke$	genitive (GEN)	subject (infinitives) specifier
$m\tilde{e}/par/tak$	locative (LOC)	oblique/adjunct

Table 4: Hindi case markers (Butt and King, 1996)

D Pearson Correlation Analysis

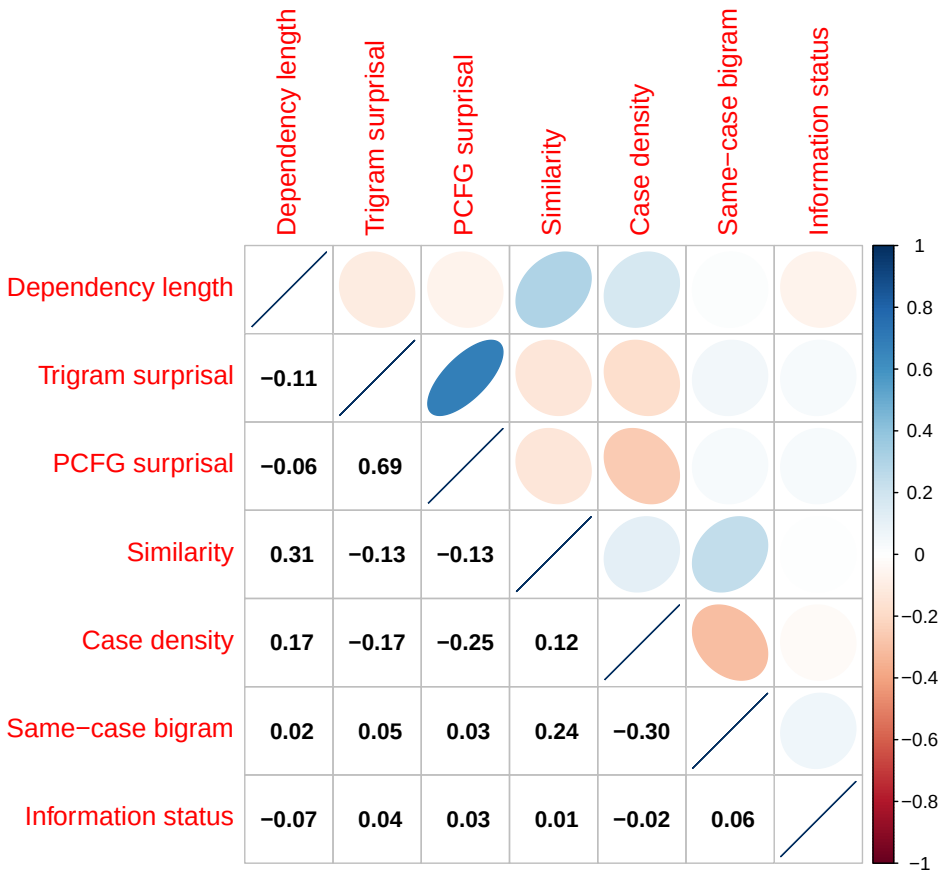


Figure 5: Pearson's correlation coefficient between various cognitive measures

E Construction Analysis

Predictor	$\hat{\beta}$	$\hat{\sigma}$	t
Intercept	1.66	0.053	31.32
IS Score	0.02	0.044	0.35
PCFG surprisal	0.06	0.066	0.88
3.gram surprisal	-0.01	0.064	-0.15
same case bigram	-0.01	0.052	-0.27
case density	0.08	0.052	1.57
similarity	0.27	0.052	5.11
3.gram x PCFG surp	-0.05	0.034	-1.31

Table 5: Linear regression model predicting average dependency length on DO-fronted dataset (133); significant predictors denoted in bold; other predictors not shown as they are not significant in the model

Predictor	$\hat{\beta}$	$\hat{\sigma}$	t
Intercept	1.78	0.065	27.62
IS Score	0.02	0.050	0.40
PCFG surprisal	0.04	0.081	0.48
3.gram surprisal	0.03	0.084	0.40
same case bigram	-0.01	0.060	-0.14
case density	-0.04	0.058	-0.65
similarity	0.22	0.054	4.03
3.gram x PCFG surp	-0.07	0.045	-1.59

Table 6: Linear regression model predicting average dependency length on IO-fronted dataset (101); significant predictors denoted in bold; other predictors not shown as they are not significant in the model

Predictor	$\hat{\beta}$	$\hat{\sigma}$	t
Intercept	1.93	0.021	92.24
IS Score	-0.02	0.018	-1.54
PCFG surprisal	0.08	0.027	3.32
3.gram surprisal	-0.05	0.027	-1.64
same case bigram	-0.03	0.019	-1.57
case density	0.09	0.022	4.13
similarity	0.18	0.019	9.34
3.gram x PCFG surp	-0.03	0.02	-1.72

Table 7: Linear regression model predicting average dependency length on conjunct-verb dataset (1158); significant predictors denoted in bold; other predictors not shown as they are not significant in the model

F Conjunct Verb Construction

In Hindi conjunct verbs, a highly predictable verb follows a nominal element, resulting in a non-compositional meaning such as *khyaal rakhna* ('care keep/put'; 'to take care of') as opposed to *guitar rakhna* ('guitar keep/put'; 'to put down or keep a guitar'). The following example illustrates Hindi conjunct verbs:

(5) baasu chatterjee-ne apne parivaar-ka **[khyaal]**$_{POF}$ **rakha**
baasu chatterjee-ERG his own family-GEN care keep.PST.PFV
Basu Chatterjee took care of his family.

G Dataset distribution

Construction(#cases)	DL	Similarity	Case density	Same-Same Sequence	Trigram surprisal	PCFG surprisal	Sentence length	#Preverbal constituents
Conjunct verbs (1158)	46.40	0.42	0.21	0.49	49.26	138.97	22.42	4.28
IO-fronted orders (101)	38.73	0.35	0.21	0.29	45.71	126.19	20.33	3.66
DO-fronted orders (133)	29.56	0.31	0.19	0.46	41.53	112.54	17.08	3.69
Full data (1996)	40.04	0.38	0.21	0.44	45.03	125.93	20.03	4.04

Table 8: Construction-specific statistics (mean values)