# Dependency Locality Influences Word Order During Production in SOV Languages: Evidence from Hindi

**Mudafia Zafar (mudafia@hss.iitd.ac.in)**
Indian Institute of Technology Delhi, New Delhi, India

**Samar Husain (samar@hss.iitd.ac.in)**
Indian Institute of Technology Delhi, New Delhi, India

### Abstract

When two arguments of a sentence vary in length, speakers of SOV languages prefer to place the longer argument before the shorter one leading to a long-before-short word order. The functional motivation behind such an ordering choice has been provided by two opposing accounts. According to the first account, long-before-short order makes production efficient by keeping syntactic heads and dependents close to each other (dependency locality). The alternate account argues that long-before-short order is a product of increased conceptual accessibility of long arguments during production. In this work, we test the predictions made by the two accounts by comparing ordering choices in Hindi transitive sentences containing object-modifying post-nominal relative clauses (RCs). Results reveal that it is efficiency and not accessibility that determines word order during production. These findings add to the body of work that argues for an overarching influence of working memory constraints on both comprehension and production.

**Keywords:** language production; word order; long before short; Hindi; phrasal length; dependency locality; accessibility

## Introduction

Linearization during sentence production has been an important area of research in psycholinguistics (e.g., Bock & Warren, 1985; F. Ferreira & Henderson, 1998; Tanaka, Branigan, & Pickering, 2011). Understanding how speakers order syntactic constituents of a sentence informs us not only about the cognitive processes that unfold during speech but also shed light on the language production architecture (Levelt, 1993). One of the factors found to influence linearization during production is constituent length which is usually operationalized as the number of words (Hawkins, 1994). Length-guided shifts from a canonical to a non-canonical word order have been observed across languages of the world (Hawkins, 2004, 2014).

In Subject-Object-Verb (SOV) languages, this leads to a long-before-short word order preference (Faghiri & Samvelian, 2020; Ranjan, Rajkumar, & Agarwal, 2022; Ros, Santesteban, Fukumura, & Laka, 2015; Yamashita & Chang, 2001). In other words, between two arguments of a sentence, one that is long and the other that is short, speakers of SOV languages prefer to place the long argument before the short argument. Consider the Japanese sentences in examples 1a-1b from Yamashita and Chang (2001) consisting of a transitive verb (V) with a short subject (S) and a long modified object (O). Yamashita and Chang (2001) found that speakers produced a significantly higher proportion of non-canonical OSV order in 1b (where the long O appears before the short S), more than the canonical SOV order in 1a.

(1) a. [S keezi-ga] [O se-ga takakute gassiri
detective-NOM height-NOM tall and
sita hannin-o] [V oikaketa]
big-boned suspect-ACC chased

'The detective chased the suspect who is tall and big-boned.'

b. [O se-ga takakute gassiri sita hannin-o] [S keezi-ga] [V oikaketa]

Two major accounts exist in the literature which explain the cognitive underpinnings of a long-before-short word order such as the one discussed above.

According to the first account, a long-before-short order is motivated by the *efficiency principle* of dependency locality under which languages tend to keep arguments, and verbs closer to each other in a sentence (Futrell, Levy, & Gibson, 2020; Gibson, 2000). Due to such proximity, speakers need to retain exact representations of the arguments for a lesser amount of time in working memory. This prevents forgetting/interference and reduces the cost associated with re-retrieving them when they need to be linked with the verb later in the sentence.[1] This is illustrated schematically in Figure 1.

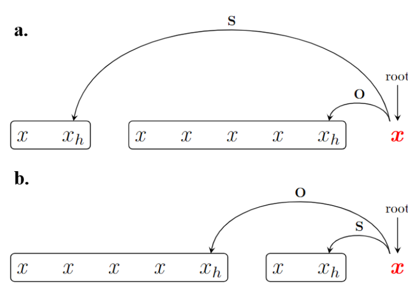


Figure 1: Two possible ordering patterns for a short subject (S) and a long object (O) in SOV languages. In (a) the short S is placed first while in (b), the long O is placed first. Order (b) has a shorter dependency length than (a).

According to the second account, long-before-short order

[1] A similar proposition by Hawkins (2014) posits that long-before-short minimizes the total Constituent Recognition Domain of a sentence.

is *accessibility* driven (Yamashita & Chang, 2001). Long arguments are semantically more salient than short ones and this salience increases their conceptual accessibility. On the other hand, short arguments are less complex and contain fewer words which makes them lexically more accessible. During incremental speech planning, the constituent that is more accessible gets planned first and thereby ends up occupying early positions in sentence structure. Yamashita and Chang (2001) posit that speakers of SOV languages prioritize conceptual accessibility over lexical accessibility, thus the long-before-short order.

## Efficiency or Accessibility?

While the phenomenon of long-before-short ordering has been widely observed in spoken as well as corpus experiments, there is little consensus on what motivates this order. For example, results from Persian, by Faghiri and Samvelian (2020), favor the accessibility account. The authors controlled for dependency length in ditransitive constructions and found that speakers produced sentences in the non-canonical DO-IO-V order when the direct object (DO) was semantically more salient than the Indirect Object (IO).

In contrast, results from Korean by Dennison (2007) favor the efficiency account. The author compared ordering choices between native speakers of Korean and Korean-English bilinguals while manipulating animacy and length. While they found a universal animate-first preference,[2] they did not observe a universal preference for a long-before-short order. Instead, the direction of length-guided shifts was modulated by the language in which the speakers were more proficient.

In most previous studies, accessibility and efficiency accounts were either not directly tested, or they made similar predictions (a long-before-short order). The role of these factors (viz. accessibility and phrasal length) on preverbal word order is thus unclear, and a direct comparison between the two accounts is currently lacking. In this work, we use Hindi, an SOV language, to investigate this. For this purpose, we use postnominal relative clause (RC) constructions.

In Hindi, a transitive sentence with a postnominal (RC) that modifies the object can be ordered in several ways. These include the canonical [***S O RC V***] order as in 2a; object-fronted [***O RC S V***] order as in 2b; or the RC can be right-extraposed leading to a [***S O V RC***] order as in 2c.[3]

(2) a. [$_{\mathbf{S}}$ pari-ne] [$_{\mathbf{O}}$ bachchon-ko] [$_{\mathbf{RC}}$ jo
fairy.SG.F-ERG children.PL.M-ACC REL
subah-se ro rahein thein]
morning-INST cry PROG.PL.M be.PAST
[$_{\mathbf{V}}$ behlaa-yaa]
console-PFV
'The fairy consoled the children who were crying since morning'

b. [$_{\mathbf{O}}$ bachchon-ko] [$_{\mathbf{RC}}$ jo subah se ro rahein thein] [$_{\mathbf{S}}$ pari-ne] [$_{\mathbf{V}}$ behlaya]

c. [$_{\mathbf{S}}$ pari-ne [$_{\mathbf{O}}$ bachchon-ko [$_{\mathbf{V}}$ behlaya [$_{\mathbf{RC}}$ jo subah se ro rahein then]

Critically, efficiency and accessibility accounts make different predictions about the ordering choices for sentence in 2a. In particular, if it is the case that long phrases become conceptually more accessible and therefore are placed in early positions in the sentence, then speakers should produce an object-fronted sentence like 2b.

The total dependency length for sentence 2b is 22. See Fig 2.[4] Alternatively, if ordering choices during production are motivated by efficiency considerations, then speakers should produce a sentence like 2c. Here, the total dependency length is 18, which is less than that of sentence 2b. See Fig 3. Note that both 2b and 2c have smaller dependency lengths compared to the canonical order in 2a (dependency length of 27).

# Experiment

In order to test the predictions of the efficiency and accessibility accounts, we conducted a sentence recall experiment which we report below.

## Task and Procedure

Following previous literature, we used the sentence recall task (Ros et al., 2015; Yamashita & Chang, 2001) to investigate word order choices during sentence production. The sentence recall task has been used to investigate ordering preferences because it can capture naturalistic production under controlled settings (V. S. Ferreira & Dell, 2000).

The experiment was designed using the jsPsych framework (De Leeuw, 2015). Participants viewed the subject, object, relative clause, and verb in different boxes on the screen and were asked to prepare a sentence with them (Screen 1 in Fig 4). The positions of the arguments and RC in Screen 1 were randomized, and the verb always appeared in the top left corner. Following an arithmetic problem (Screen 2 in Fig 4), they were given a cue (Screen 3 in Fig 4) to speak out loud the sentence they had prepared, and their responses were recorded. After they finished speaking, they were prompted (Screen 4 in Fig 4) to answer whether the presented sentence had appeared previously during the course of the experiment or not. For this purpose, forty percent of fillers were repeated during the experiment.

## Materials and Design

The experiment had a single-factor within-subjects design. Twenty-one sets of sentences constituted the critical items. Each set had three sentence variants (see Table 1) that constituted the three levels. Level 1 (All-Short) consisted of

[2] Animacy has been well researched as a factor increasing conceptual accessibility across typologically different languages (e.g., McDonald, Bock, & Kelly, 1993).

[3] In addition, there are other possible orders, but they are not relevant to our research question here.

[4] We chose the Surface Syntactic Universal Dependencies (SUD) annotation scheme to determine the dependency structure for all the sentences (Gerdes, Guillaume, Kahane, & Perrier, 2018). Note that the predictions laid out do not change even with the Universal Dependencies (UD) annotation scheme (De Marneffe, Manning, Nivre, & Zeman, 2021).

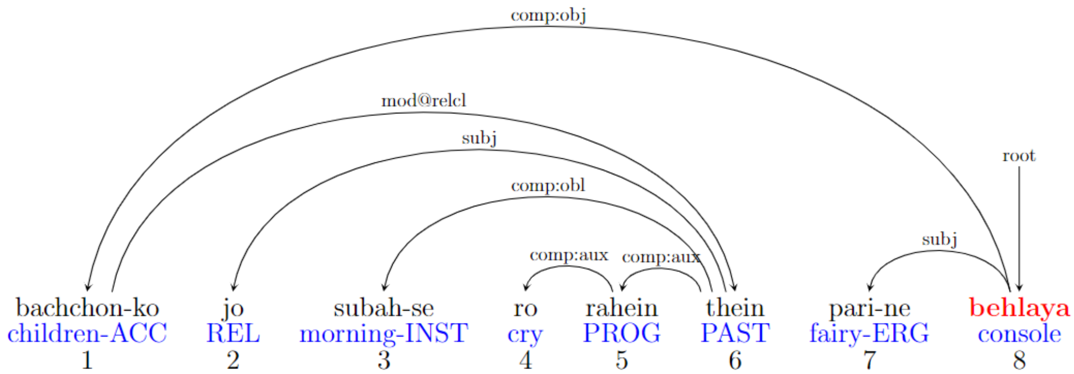


Figure 2: Dependency tree for [***O RC S V***] sentence; Total Dependency Length = 22.

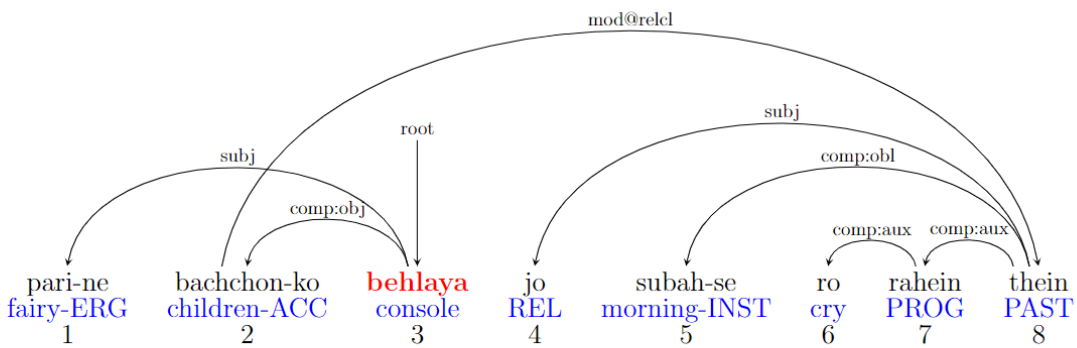


Figure 3: Dependency tree for [***S O V RC***] sentence; Total Dependency Length= 18

a simple transitive sentence with no modifiers; Level 2 (Dependency-Neutral) consisted of an intransitive sentence with an RC modifying the subject; Level 3 (O-Long) consisted of a transitive sentence with an RC modifying the object. Conditions All-Short and Dependency-Neutral acted as baselines against which ordering patterns in the O-Long condition were compared. The number and gender features between the subject and object were varied so that the RC verb could agree only with the object (Kachru, 2006). Animacy was controlled by keeping the arguments of all sentences animate.

A total of three lists were created following the Latin Square Design. Each list contained eighty-one sentences (one of the three variants from the critical manipulation along with fifty fillers). The order of trials in the lists was pseudorandomized such that at least two fillers separated each critical sentence, and there were two fillers at the beginning and at the end of the experiment. Filler trials consisted of sentences with intransitive, transitive, ditransitive, and imperative construction types. Additionally, each session started with eight practice trials.

## Participants

45 native speakers of Hindi participated in the experiment. Each experimental session lasted for fifty minutes, and they were paid INR 250 for participation.

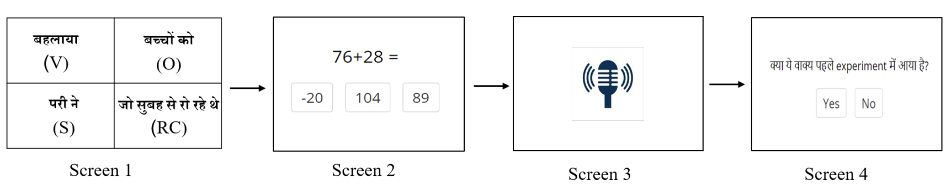


Figure 4: Trial sequence for the sentence recall task. Note that the glossings in Screen 1 (e.g., S, V, etc) are for illustrative purpose only; they did not appear on the screen during the experiment.

Table 1: Sample sentences for the three conditions of the experiment in the canonical order.

| **Condition** | **Sample Item** |
|---|---|
| (a) All-Short (S O V) | pari-ne bachchon-ko behlaya<br>(The fairy consoled the children) |
| (b) Dependency-Neutral (S RC V) | vo pari jo dukhi thi ro rahi hai<br>(The fairy who was upset is crying) |
| (c) O-Long (S O RC V) | pari-ne bachchon-ko jo subah-se ro rahein thein behlaya<br>(The fairy consoled the children who were crying) |

## Predictions

The All-Short condition served as a baseline for testing predictions of the accessibility account. Sentences in the object-fronted order in the O-Long condition [***O RC S V***] were compared with the number of object-fronted sentences in the All-Short condition [***O S V***] where fronting of the object could not be motivated by accessibility. If accessibility drives length-guided ordering in SOV languages, then we predicted that speakers would produce more object-fronted sentences in the O-Long condition when compared to the All-Short condition.

The Dependency-Neutral condition served as a baseline for testing predictions of the efficiency account. Sentences with the right-extraposed order in the O-Long condition [***S O V RC***] were compared with the number of right-extraposed sentences in the Dependency-Neutral condition [***S V RC***] where right-extraposition could not be motivated by dependency length minimization. This is because dependency length does not change between the canonical [***S RC V***] and the right-extraposed [***S V RC***] orders in this condition (see Fig 5). If length-guided shifts result from efficiency, then we predicted that speakers would produce more right-extraposed sentences in the O-Long condition when compared to the Dependency-Neutral condition.[5]

## Scoring and Analysis Procedure

In order to determine the order of the arguments, the audio responses were manually transcribed by a research assistant who was blind to the conditions of the experiment.

We excluded trials where participants gave incorrect or incomplete responses or did not respond at all. This constituted about 3% of the data. The orders in which participants articulated the critical sentences were scored in the following fashion:

***Canonical*** for [S O V] order in the All-Short condition; [S RC V] order in the Dependency-Neutral condition; and [S O RC V] order in the O-Long condition

***Right-extraposed*** for [S V RC] order in the Dependency-Neutral condition; and [S O V RC] order in the O-Long condition

***O-fronted*** for [O S V] order in the All-Short condition and [O RC S V] order in the O-Long condition.

[5]The predictions for the experiment, research question, design and analysis plan were pre-registered (`https://osf.io/u2mtw`) before accessing the data for analysis.

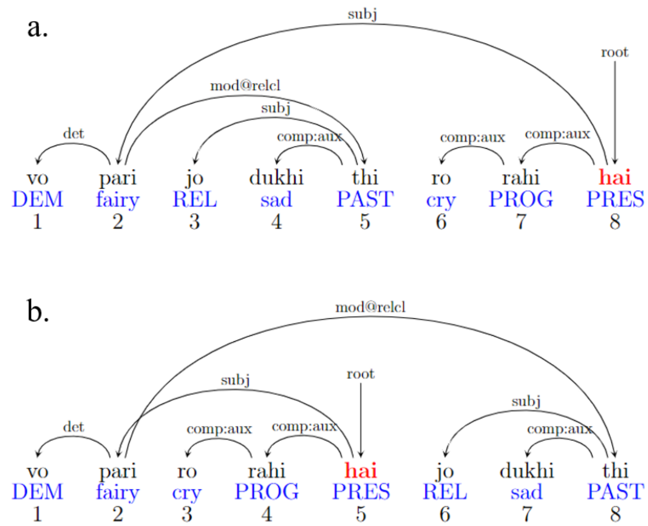


Figure 5: Dependency trees for the canonical (a) and right-extraposed (b) orders in the Dependency-Neutral condition; Total dependency length for both the orders is 15

***Miscellaneous*** for responses in any other order [6].

All analysis was done using R (version 4.2.1) programming language (R Core Team, 2022). Generalized linear mixed effects models with the logit link function (Jaeger, 2008) were fit to the data using the *lme4* package (Bates, Mächler, Bolker, & Walker, 2015) in R. The random effects component of the model was kept maximal to the extent supported by the data (Barr, Levy, Scheepers, & Tily, 2013). Two sets of analyses were performed to test the predictions of the two accounts.

For the first analysis, which tested whether efficiency drives ordering choices or not, we subset the data to include responses only from the Dependency-Neutral and O-Long conditions. Responses in the canonical order were coded as 0, while responses in the right-extraposed order were coded as 1. The model then tested whether the number of right-extraposed orders (outcome variable) was significantly different between the Dependency-Neutral and the O-Long conditions (predictor variable).

[6]*Miscellaneous* responses constituted 1.9% of the coded data and were removed from further analysis

Similarly, to test whether accessibility drives ordering or not, we subset the data to include responses only from the All-Short and O-Long conditions. Responses in the canonical order were coded as 0 and in the O-fronted order were coded as 1. The model then tested whether the number of O-fronted orders (outcome variable) was significantly different between the All-Short and the O-Long conditions (predictor variable).

## Results

Table 2 summarizes the percentages of various orders spoken by participants in the different conditions.

Table 2: Percentage of different sentence orders observed in the experiment

| Condition | Response Order | Percentage |
|---|---|---|
| All-Short | SOV | 92.18 |
| | OSV | 7.82 |
| Dependency-Neutral | SRCV | 81.35 |
| | SVRC | 15.43 |
| | RCSV | 3.22 |
| O-Long | SOVRC | 67.46 |
| | SORCV | 28.14 |
| | ORCSV | 1.69 |
| | OSVRC | 1.36 |
| | OSRCV | 0.68 |
| | SRCOV | 0.34 |
| | SVORC | 0.34 |

**Analysis 1:** In this analysis, we compared the number of right-extraposed responses between the Dependency-Neutral and the O-Long conditions. Results show that right-extraposition was significantly higher in the O-Long condition compared to the Dependency-Neutral condition ($p<0.001$). See Table 3 and Fig 6.

Table 3: glmer model output for right-extraposition from Analysis 1. Treatment contrast was used with Dependency-Neutral as the baseline condition.

| | Estimate | Std.Error | *z value* | *p value* |
|---|---|---|---|---|
| Intercept | -6.87 | 1.87 | -3.67 | <0.001 |
| O-Long | 15.48 | 2.57 | **6.02** | <0.001 |

**Analysis 2:** In this analysis we compared the number of O-fronted responses between the All-Short and the O-Long conditions. Results reveal no significant difference in the number of O-fronted responses between the two conditions ($p$=0.34). In fact, O-fronted responses were relatively low in both the conditions. See Table 4 and Fig 7.

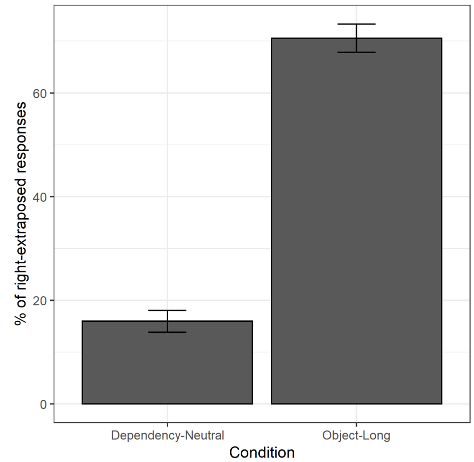


Figure 6: Percentage of right-extraposed responses in the Dependency-Neutral and O-Long conditions

Table 4: glmer model output for O-fronting from Analysis 2. Treatment contrast was used with All-Short as the baseline condition

| | Estimate | Std.Error | *z value* | *p value* |
|---|---|---|---|---|
| Intercept | -3.32 | 0.58 | -5.69 | <0.001 |
| O-Long | -2.05 | 2.15 | -0.95 | 0.34 |

## Discussion

The key finding of the current work is that Hindi native speakers are influenced by dependency length minimization considerations in their choice of word order. We found that participants were more likely to produce sentences with right-extraposed relative clauses when such an extraposition minimized the total dependency length of the sentence. When such extrapositions did not affect dependency length, there were markedly fewer right-extrapositions. In addition, results also reveal that Hindi native speakers did not prefer to front the long object early in the sentence. Together, these results provide compelling evidence in favor of the efficiency account and go against the accessibility account in explaining word order choices in Hindi.

The current findings have interesting implications for the notion of incrementality during speech planning (F. Ferreira & Swets, 2002). Ordering sentences such that heads and dependents are proximate to each other could entail that speakers engage in advance syntactic planning. This suggests that planning during production is more structural than sequential (Momma, 2021). We hope to investigate these issues in future experiments.

While the results go against the accessibility account, it is important to note that we are not arguing against the role of accessibility per se during production. There is independent evidence for accessibility-increasing factors such as animacy (e.g., McDonald et al., 1993), givenness (e.g., V. S. Ferreira & Yoshita, 2003), and imageability (e.g., Bock & Warren, 1985) which cause lemmas to appear early in the sentence. However, this work does suggest that the notion of acces-

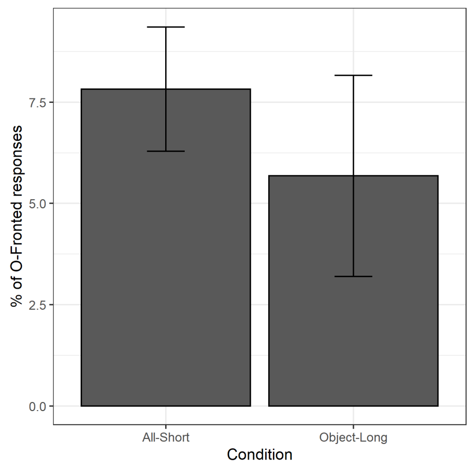


Figure 7: Percentage of O-fronted responses in the All-Short and O-Long conditions

sibility as operationalized by Yamashita and Chang (2001) does not explain length-guided word order variations during production (also see, Ros et al., 2015). This is not surprising because unlike notions such as animacy, givenness, etc. that are generalizations cross-linguistically, it is unclear why SOV language in particular should prioritize conceptual accessibility over lexical accessibility as suggested by Yamashita and Chang (2001).

One possible explanation for the right-extraposition pattern found in the data could be the *easy-first* principle of the Production-Distribution-Comprehension account (MacDonald, 2013). On this account, the production system prefers to plan and produce those sentential elements first that put less strain on cognitive resources. The high [***S O V RC***] pattern in the O-Long condition could be explained as a preference for producing the shorter and therefore easy arguments first. However, this account would also predict a [***S V RC***] pattern in the Dependency-Neutral condition. This prediction is not borne out in the data where the [***S RC V***] is the most frequent order for this condition.

The current findings, therefore, provide novel evidence for dependency locality as a critical constraint underlying language production in an SOV language like Hindi. These findings add to the body of work that argues for an overarching influence of working memory constraints on both comprehension and production (Scontras, Badecker, Shank, Lim, & Fedorenko, 2015; Gibson et al., 2019).

## Conclusion

In a sentence production experiment, we tested the predictions of two accounts that explain length-guided word order variation during production by manipulating the length of the arguments of a transitive sentence. We found compelling evidence in favor of the efficiency account, operationalized as the pressure to keep syntactic heads and dependents close to each other, in influencing ordering choices. On the other hand, we did not find evidence for the accessibility account, operationalized as making longer arguments more conceptually accessible, in determining word order. The results point to the generalizability and universality of dependency locality as a principle explaining the cross-linguistic differences in word order.

## Acknowledgments

We would like to thank Harsh Jain and Pranay Sinha for their help with designing the experiment on JsPsych; Niralee Gupta for helping with data collection and transcribing; Eashani Sharma for helping with stimuli preparation; members of the Psycholinguistics Lab at IIT Delhi for their suggestions; and the anonymous reviewers of CogSci ‘23 for their comments.

## References

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syntactic structure in sentence formulation. *Cognition*, *21*(1), 47–67.

De Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

De Marneffe, M.-C., Manning, C. D., Nivre, J., & Zeman, D. (2021). Universal dependencies. *Computational linguistics*, *47*(2), 255–308.

Dennison, H. (2007). Universal versus language-specific conceptual effects on shifted word order production in Korean: Evidence from bilinguals. In *Poster presented at the 20th CUNY Sentence Processing Conference, 2007.*

Faghiri, P., & Samvelian, P. (2020). Word order preferences and the effect of phrasal length in SOV languages: evidence from sentence production in Persian. *Glossa: a journal of general linguistics*.

Ferreira, F., & Henderson, J. M. (1998). Linearization strategies during language production. *Memory & Cognition*, *26*(1), 88–96.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language*, *46*(1), 57–84.

Ferreira, V. S., & Dell, G. S. (2000). Effect of ambiguity and lexical availability on syntactic and lexical production. *Cognitive psychology*, *40*(4), 296–340.

Ferreira, V. S., & Yoshita, H. (2003). Given-new ordering effects on the production of scrambled sentences in Japanese. *Journal of psycholinguistic research*, *32*(6), 669–692.

Futrell, R., Levy, R. P., & Gibson, E. (2020). Dependency locality as an explanatory principle for word order. *Language*, *96*, 371 - 412.

Gerdes, K., Guillaume, B., Kahane, S., & Perrier, G. (2018). SUD or Surface-Syntactic Universal Dependencies: An annotation scheme near-isomorphic to UD. In *Proceedings of the second workshop on Universal Dependencies (UDW 2018)* (pp. 66–74).

Gibson, E. (2000). The dependency locality theory: A distance-based theory of linguistic complexity. *Image, language, brain*, *2000*, 95–126.

Gibson, E., Futrell, R., Piantadosi, S. P., Dautriche, I., Mahowald, K., Bergen, L., & Levy, R. (2019). How efficiency shapes human language. *Trends in Cognitive Sciences*, *23*(5), 389–407.

Hawkins, J. A. (1994). *A performance theory of order and constituency* (No. 73). Cambridge University Press.

Hawkins, J. A. (2004). *Efficiency and complexity in grammars*. OUP Oxford.

Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. OUP Oxford.

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of memory and language*, *59*(4), 434–446.

Kachru, Y. (2006). *Hindi* (Vol. 12). John Benjamins Publishing.

Levelt, W. J. (1993). *Speaking: From intention to articulation* (Vol. 1). MIT press.

MacDonald, M. C. (2013). How language production shapes language form and comprehension. *Frontiers in Psychology*, *4*, 226.

McDonald, J. L., Bock, K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive psychology*, *25*(2), 188–230.

Momma, S. (2021). Filling the gap in gap-filling: Long-distance dependency formation in sentence production. *Cognitive Psychology*, *129*, 101411.

R Core Team. (2022). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria.

Ranjan, S., Rajkumar, R., & Agarwal, S. (2022). Locality and expectation effects in Hindi preverbal constituent ordering. *Cognition*, *223*, 104959.

Ros, I., Santesteban, M., Fukumura, K., & Laka, I. (2015). Aiming at shorter dependencies: the role of agreement morphology. *Language, Cognition and Neuroscience*, *30*(9), 1156–1174.

Scontras, G., Badecker, W., Shank, L., Lim, E., & Fedorenko, E. (2015). Syntactic complexity effects in sentence production. *Cognitive Science*, *39*(3), 559–583.

Tanaka, M. N., Branigan, H. P., & Pickering, M. J. (2011). The production of head-initial and head-final languages. *Processing and producing head-final structures*, 113–129.

Yamashita, H., & Chang, F. (2001). "Long before short" preference in the production of a head-final language. *Cognition*, *81*(2), B45–B55.