UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

Production Efficiency in Written Hindi Falsehoods

Permalink

https://escholarship.org/uc/item/8ft8t0x7

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 48(0)

Authors

Arun, Ishita

Husain, Samar

Publication Date

2026

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Production Efficiency in Written Hindi Falsehoods

Ishita Arun (huz248352@iitd.ac.in)[1] **& Samar Husain (samar@iitd.ac.in)**[2]
[1]Department of Humanities and Social Sciences, Indian Institute of Technology Delhi
[2]Department of Humanities and Social Sciences, Indian Institute of Technology Delhi

Abstract

This study examines linguistic differences between truthful and false written narratives in Hindi using psycholinguistically motivated measures of language processing complexity. Adopting a corpus-based, processing-oriented approach, we investigate how the cognitive demands of producing falsehoods shape lexical and syntactic choices beyond surface-level cues. Participants produced written opinions that were either truthful or intentionally false, which were analyzed using metrics including surprisal, dependency distance, part-of-speech distributions, and type–token ratio. Our results show that lower mean surprisal, along with higher verb counts and reduced use of nouns and adverbs, significantly predicts false narratives. These patterns suggest that falsehood production in Hindi relies on linguistically efficient strategies characterized by predictable constructions and reduced informational specificity to ease planning and monitoring under cognitive load. By presenting evidence from an underrepresented South Asian language, this study advances cross-linguistic research on deception and highlights the value of processing-based metrics for understanding deceptive language production.

Keywords: falsehood; deception; false opinion; verbal cues; psycholinguistics; corpus analysis; Hindi; language processing

Introduction

Human communication often involves deception, as language enables speakers to intentionally mislead others by expressing ideas, beliefs, or opinions that they do not genuinely hold (Austin, 1975; Bok, 2011; Carson, 2006; Levinson, 1983; Meibauer, 2019). Verbal deception can take several forms, including lying, exaggeration, and misdirection (Levine, 2014; Meibauer, 2019; Zuckerman et al., 1981). Lying typically involves presenting false information as true (Arciuli et al., 2010; Bok, 2011; Gillings, 2024; Mahon, 2008; Sarzynska-Wawer et al., 2023; Vrij et al., 2010). In this paper, we operationalize *falsehood* as the intentional verbal expression of an opinion that the speaker does not believe. Under this definition, falsehood is treated as a form of lying, with an emphasis placed on subjective opinions rather than objective factual claims.

Research on deception has largely centered on the identification of cues that can reliably signal deceptive behavior. Early work primarily emphasized non-verbal indicators, such as facial expressions, gestures, and bodily movements, while subsequent research expanded this focus to include verbal and paralinguistic cues, including speech errors, hesitations, narrative inconsistencies, and variations in vocal pitch (DePaulo et al., 1982, 2003; Ekman, 1989; Ekman & Friesen, 1969; Loy et al., 2018; Masip et al., 2012; Vrij et al., 2010; Zuckerman et al., 1982).

Linguistic cues of deception are increasingly important for detecting deceptive behavior in technology-mediated and text-based interactions, where non-verbal signals are absent. Prior research has shown that deceivers tend to use language differently from truth-tellers (Burgoon et al., 2003; Hartwig et al., 2010; Newman et al., 2003). Because deceptive narratives are not grounded in actual experience, they are often shorter and more abstract, contain fewer concrete and contextual details, exhibit reduced use of first-person pronouns and exclusivizers, and employ less complex vocabulary (Arciuli et al., 2010; Gillings, 2024; HaCohen-Kerner et al., 2015; Johnson & Raye, 1981; Newman et al., 2003; Sarzynska-Wawer et al., 2023).

However, these surface-level, count-based linguistic cues are not universal and vary across languages. For instance, truthful accounts in Spanish contain more prepositions than deceptive ones, whereas in Polish, deceptive accounts more frequently employ infinitive forms (Masip et al., 2012; Sarzynska-Wawer et al., 2023). Such findings suggest that deception is manifested through language-specific realizations, limiting the cross-linguistic generalizability of cues that solely rely on lexical and morphosyntactic features.

Despite these cross-linguistic differences, a common underlying theoretical account holds that the differential use of language in truthful versus deceptive accounts reflects the greater cognitive effort required during deception (Burgoon & Buller, 1994; Ekman & Friesen, 1969; Pennebaker & Chew, 1985; Zuckerman et al., 1981). To lie successfully, speakers must maintain internal consistency in their fabrication while ensuring external consistency with what others already know (Zuckerman et al., 1981). Because deception is an interactive process, liars must continuously monitor and regulate their behavior to avoid "leakages" that may reveal the lie, making it mentally taxing (Burgoon & Buller, 1994; Zuckerman et al., 1981).

The increased cognitive demands are reflected in both the content and timing of deceptive language. Deceptive responses often exhibit longer response times, indicating increased planning before communication (Blandón-Gitlin et al., 2014; Gombos, 2006; Walczyk et al., 2003). Sentence production involves conceptual preparation, lexical selection, syntactic planning, and active monitoring (Bock & Levelt,

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society* ©2026 the author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1994; Levelt, 1983; Willem, 1989). When lying, speakers must additionally inhibit truthful information, monitor consistency, and anticipate the listener's expectations, which increases cognitive load (Gombos, 2006; Lane & Wegner, 1995; Walczyk et al., 2003; Zuckerman et al., 1981). This extra effort often results in shorter, simpler, and more constrained utterances, mirroring the linguistic patterns observed in deceptive narratives (DePaulo et al., 2003; Leal et al., 2010; Loconte et al., 2023; Newman et al., 2003; Walczyk et al., 2003). Models such as the Preoccupation Model of Secrecy (Lane & Wegner, 1995) and the Activation-Decision-Construction Model (ADCM) (Walczyk et al., 2003) attempt to formalize some of these processes from a Cognitive Psychology perspective, emphasizing the role of inhibition and working memory in lie fabrication.

Building on this cognitive account, Psycholinguistics offers a novel perspective for analyzing linguistic complexity in deceptive narratives. It defines linguistic complexity as "a measure of the cognitive difficulty of human language processing" (Frazier, 1985; Liu, 2008; Nichols, 2009). Within this framework, different theoretical accounts have operationalized processing difficulty.

The *dependency distance minimization hypothesis* posits that, because language processing is constrained by the limited capacity of working memory, languages tend to avoid complex structures, particularly those involving long distances between syntactic heads and their dependents, thereby motivating dependency distance as a measure of linguistic complexity (Futrell et al., 2020; Gibson, 2000). Thus, dependency distance measures the linear distance between syntactically linked words, with longer distances reflecting higher processing demands during sentence processing (Futrell et al., 2015; Gibson, 2000; Liu, 2008; Rajkumar et al., 2016; Scontras et al., 2015; Yadav et al., 2020; Zafar & Husain, 2023).

By contrast, the adaptability hypothesis characterizes complexity in terms of distributional expectations: structures that occur more frequently are easier to process than less probable ones, motivating surprisal as a measure of processing difficulty (Christiansen & MacDonald, 2009; Hale, 2001; Levy, 2008; Vasishth et al., 2010). Surprisal estimates the predictability of each word given its preceding context, with higher surprisal indicating a greater processing effort required to integrate unexpected or less predictable words (Demberg & Keller, 2008; Hale, 2001; Levy, 2008; Patil et al., 2008).

Approaches grounded in semantic coherence view processing difficulty as arising from the effort required to integrate meaning across discourse (Carter & Hoffman, 2024). From this perspective, semantic similarity captures the degree of conceptual alignment between a word and its surrounding context, with lower similarity indicating increased integration difficulty and greater cognitive effort during comprehension and production (Carter & Hoffman, 2024; Corley & Mihalcea, 2005; Kun et al., 2023; Miller & Charles, 1991; Sravanthi & Srinivasu, 2017).

Psycholinguistic measures of complexity, such as dependency distance, surprisal, and semantic similarity, provide quantifiable insights into the cognitive effort involved in language planning, production, and comprehension (Futrell et al., 2020; Jaeger & Tily, 2011; Levy, 2008; McDonough & Trofimovich, 2011; Rajkumar et al., 2016). Together, these metrics systematically capture subtle variations in sentence structure, lexical choice, and conceptual organization that arise from the increased cognitive load during deception, offering a fine-grained perspective that complements traditional cue-based approaches. Crucially, because they target underlying cognitive and informational properties of language rather than surface-level features, these measures have been shown to generalize across typologically diverse languages (Futrell et al., 2015; Wilcox et al., 2023). In other words, the patterns they reveal are consistent across languages, suggesting they capture universal properties of language production rather than features specific to one language, making them well-suited for investigating deception from a cross-linguistic perspective.

While cues of deception have been extensively studied, much of the research is grounded in psychological or forensic-linguistic frameworks. Applying psycholinguistic measures of complexity to corpus data provides a novel toolkit for analyzing deceptive discourse at scale, revealing context-sensitive patterns (Gillings, 2024). Building on these insights, our study proposes a corpus-based comparison of deceptive and truthful accounts using psycholinguistic measures of processing complexity alongside count-based features. Count-based features capture what is produced, while psycholinguistic measures provide insights into how language is planned and processed (Futrell et al., 2020; Jaeger & Tily, 2011; Levy, 2008; McDonough & Trofimovich, 2011; Newman et al., 2003; Rajkumar et al., 2016; Sarzynska-Wawer et al., 2023; Scontras et al., 2015). Taken together, these complementary perspectives offer a more comprehensive account of linguistic variation in falsehoods.

Specifically, we aim to address three gaps: (1) comparing linguistic patterns between true and false narratives in terms of processing efficiency; (2) applying psycholinguistic metrics, such as dependency distance, surprisal, and semantic similarity, to analyze the verbal content of the narratives while grounding the findings in psychological theories of deception; and (3) extending the analysis cross-linguistically to include Hindi[1] (a Subject-Object-Verb language).

To compare the linguistic efficiency of truthful narratives and falsehoods in Hindi, we first construct a dataset of speakers' true and opposing opinions using the false-opinion paradigm (Ekman, 1989; Leal et al., 2010; Mehrabian, 1971; Mihalcea & Strapparava, 2009; Newman et al., 2003; Sarzynska-Wawer et al., 2023). Using this dataset, we compare narratives along both simple count-based measures (for example, the frequency of nouns, verbs, adjectives, and

[1]Hindi belongs to the Indo-Aryan language family and is spoken primarily in northern-western and central India. It has over 600 million speakers, making it one of the most widely spoken languages in the world (Gordon Jr, 2005).

adverbs) and psycholinguistically motivated complexity metrics, including mean dependency distance, surprisal, and semantic similarity. Our results demonstrate several predictors that reliably distinguish between truthful and false opinions in the current dataset. This work contributes a novel, cognitively grounded approach to the study of deception. Future research can extend these findings to other languages and to speech-based corpora to assess the robustness of these metrics across modalities and linguistic contexts.

Methods

Participants

Participants were 46 native adult speakers of Hindi recruited through the authors' personal networks. All participants reported normal or corrected-to-normal vision and hearing, with no history of speech, language, or neurological disorders. Participation was voluntary; informed consent was obtained prior to participation, and all collected responses were anonymized.

Procedure

To construct a dataset comprising both truthful and false opinions, we employed the "false opinion" paradigm, frequently used in deception literature (Arciuli et al., 2010; Capuozzo et al., 2020; Ekman, 1989; Leal et al., 2010; Mehrabian, 1971; Mihalcea & Strapparava, 2009; Newman et al., 2003; Pérez-Rosas & Mihalcea, 2014; Sarzynska-Wawer et al., 2023), to elicit deceptive and truthful narratives from native Hindi speakers across three socially controversial topics: abortion, the death penalty, and best friends. Data were collected using Google Forms, a platform also used by Capuozzo et al. (2020). Each participant was permitted a single submission. Responses were provided in Romanized Hindi and were subsequently converted into Devanagari script by the authors for analysis.

For the topics of abortion and the death penalty, participants were first asked to write a 4–5 line narrative expressing their genuine opinion on whether the practice should be legal, including as many details as possible, as if they were participating in a debate. They were then instructed to produce a second narrative of comparable length and detail that argued the opposite position, thereby generating a falsehood. For the best friend topic, participants were instructed to describe the reasons for their friendship with their actual best friend, which served as the truthful narrative. To elicit a false opinion, they were asked to think of someone they strongly disliked and describe that person as if they were their best friend. Instructions regarding the length of the narrative, level of detail, and the debate scenario were kept identical across topics for both truthful and falsehood conditions.

For each topic, we collected 46 truthful and 46 false opinions, yielding a total of 276 texts. Following a manual quality check, 28 narratives were excluded from the analysis. These narratives were either written in English or consisted of refusals or meta-statements indicating an inability or unwillingness to provide an answer. The final dataset, therefore, consists of 248 narratives. The distribution of narratives across topics and conditions is reported in Table 1.

Topic	Total Statements	Truthful	Falsehood
Abortion	87	42	45
Death Penalty	86	43	43
Best Friend	75	44	31
Overall	**248**	**129**	**119**

Table 1: Distribution of truthful and falsehood statements by topic.

Analysis

To investigate the linguistic differences between truthful narratives and falsehoods, we extracted a range of features using Python. Generative AI tools were employed minimally, solely to assist with code debugging. Following prior work (Loconte et al., 2023; Newman et al., 2003; Sarzynska-Wawer et al., 2023; Vrij et al., 2000), we computed a set of count-based features, described below, using the Stanza NLP library (Qi et al., 2020).

- *Narrative length*: the total number of tokens in a narrative.
- *Sentences per narrative*: the total number of sentences in a narrative.
- *Type-token ratio*: the ratio of unique words (types) to total words (tokens) in a narrative, reflecting lexical diversity.
- *Part-of-speech counts*: the frequency of nouns, verbs, adjectives, adverbs, auxiliaries, and first-person pronouns in a narrative. For this paper, verb count was operationalized as the total number of main verbs, with no distinction made for light verbs. Nouns were counted irrespective of their syntactic position within the sentence.
- *Negation markers*: the frequency of negation markers (e.g.[2], *nahi* 'not', *na* 'no/not') in a narrative.
- *Discourse connectives*: the frequency of discourse connectives (e.g., *kyunki* 'because', *jaise* 'for example/like') in a narrative.

The operational definitions of the psycholinguistic features, along with the procedures used to compute them for each narrative, are detailed below.

- *Mean dependency distance* is a measure of syntactic complexity proposed by Futrell et al., 2015; Liu, 2008; Yadav et al., 2020. It captures the linear distance between a head and its dependent in a sentence. The dependency distance (DD) for each head-dependent pair is calculated as the absolute difference between their positions in the sentence. The mean dependency distance for a sentence is computed as the mean of the absolute distances across all dependency pairs.
  We computed the mean dependency distance for each narrative by averaging the values across all sentences using

[2]Hindi examples are presented using a simplified Latin transliteration without diacritics to ensure readability. Glosses are provided in English for clarity.

the Stanza NLP library for Hindi (Qi et al., 2020). Higher values indicate longer syntactic dependencies and are associated with greater planning and processing effort during sentence production (Hawkins, 2014; Scontras et al., 2015; Zafar & Husain, 2023).

- *Mean surprisal* measures the cognitive effort during sentence processing by computing the predictability of a word in a sentence given its preceding context (Hale, 2001). It is computed using the formula: negative log probability of a word given its preceding context (Hale, 2001). Less predictable words in a given context have higher surprisal values and are associated with an increased processing cost (Demberg & Keller, 2008; Levy, 2008; Patil et al., 2008). In this study, we estimated token-level surprisal for each narrative using Nvidia's Nemotron-4-Mini-Hindi-4B-Base model, a pretrained language model for Hindi, English, and Hinglish (Joshi et al., 2024). Token-level surprisal values were summed to obtain sentence-level surprisal, which was normalized by sentence length. Lastly, narrative-level surprisal was derived by adding sentence-level surprisal values and normalizing by sentence count. Higher mean surprisal values suggest that a narrative is syntactically or semantically less predictable.
- *Mean semantic similarity* captures the conceptual coherence among consecutive structures (Miller & Charles, 1991; Sravanthi & Srinivasu, 2017). Lower scores indicate that two consecutive structures are less semantically related. In this study, we computed semantic similarity for each narrative as the average cosine similarity between vector embeddings of consecutive sentences obtained from Google's LaBSE model (Feng et al., 2022).

Feature extraction in Python produced a CSV file in which each narrative was represented by a set of linguistic features encoded as separate columns. This dataset was subsequently analyzed in R. To address multicollinearity, we examined variance inflation factor (VIF) scores and selected a generalized linear model (GLM) that excluded predictors with high collinearity. The final GLM included condition (truthful vs. falsehood) as a binary dependent variable and the following linguistic features as fixed effects: mean dependency distance, mean surprisal, mean semantic similarity, type–token ratio, counts of nouns, verbs, adverbs, adjectives, first-person pronouns, auxiliaries, negation markers, and discourse connectives.

Predictions

Based on prior research, we hypothesized that falsehood narratives would exhibit markers of increased cognitive effort relative to truthful narratives (Blandón-Gitlin et al., 2014; Gombos, 2006; Walczyk et al., 2003; Zuckerman et al., 1981). Specifically, we predicted lower mean dependency distance, reflecting reduced syntactic complexity to maintain simpler sentence structures; higher mean surprisal, reflecting the production of less predictable linguistic sequences when events are not grounded in actual experience; and lower semantic similarity, reflecting reduced coherence across sentences.

We also anticipated differences in count-based features, with falsehood narratives showing fewer first-person pronouns, consistent with increased psychological distancing, fewer adjectival and adverbial modifiers to maintain structural simplicity, and fewer content words overall to reduce the cognitive demands of fabrication.

Results

We found that participants produced an average of 63 words to express their true opinions and 55 words to express their false opinions. A linear mixed-effects model with random intercepts for participants and topics revealed a significant effect of condition on narrative length, with falsehood narratives being shorter than truthful narratives ($\beta = -8.997$, $SE = 2.69$, $t = -3.35$).

Since narrative length exhibited substantial collinearity with several linguistic predictors, it was excluded from the final model. The final model included scaled values of mean dependency distance, mean surprisal, mean semantic similarity, type–token ratio, and counts of nouns, verbs, adverbs, adjectives, first-person pronouns, auxiliaries, negation markers, and discourse connectives as fixed effects. Table 2 reports the parameter estimates from this model.

The analysis revealed a significant positive effect of verb count ($\beta = 1.04$, $SE = 0.31$, $p < 0.001$), indicating that narratives with a higher number of verbs were more likely to be classified as falsehoods. In contrast, noun count ($\beta = -1.04$, $SE = 0.35$, $p < 0.01$) and adverb count ($\beta = -0.35$, $SE = 0.17$, $p < 0.05$) showed significant negative effects, such that increases in these features were associated with a lower likelihood of a narrative being a falsehood. Mean surprisal also exhibited a significant negative effect ($\beta = -0.32$, $SE = 0.16$, $p < 0.05$), with higher surprisal values corresponding to reduced odds of falsehood. No other predictors reached statistical significance in the present dataset. Figure 1 visualizes the effects of predictors that showed significant contributions to the model.

Discussion

Our study investigated linguistic differences between written truthful narratives and falsehoods in Hindi, using measures of language processing complexity drawn from Psycholinguistics and applied within a corpus-based framework. Our goal was to extend the literature on deception by introducing processing-oriented metrics that offer insight into the cognitive mechanisms underlying deceptive language production. In doing so, we contribute empirical evidence on how planning and production demands shape deceptive communication, particularly in a language that has been underrepresented in this literature.

A central assumption in deception research is that speakers strategically modulate their behavior to appear credible (DePaulo et al., 2003; Hartwig et al., 2010). Such modula-

Predictor	β	SE	z	p
Intercept	−0.115	0.136	−0.85	.397
Mean dependency distance	−0.066	0.156	−0.42	.674
Semantic similarity	−0.019	0.168	−0.12	.908
Sentence surprisal	−0.321	0.156	−2.06	.040*
Type–token ratio	−0.242	0.241	−1.00	.315
Verb count	1.038	0.314	3.31	< .001***
Noun count	−1.038	0.354	−2.94	.003**
Auxiliary count	−0.365	0.270	−1.35	.177
Adjective count	0.023	0.265	0.09	.930
Adverb count	−0.348	0.171	−2.03	.043*
First-person pronoun count	−0.066	0.169	−0.39	.695
Negation marker count	−0.260	0.170	−1.53	.126
Discourse connective count	0.032	0.221	0.15	.885
Honorificity marker count	−0.163	0.151	−1.08	.280

Table 2: Results of the logistic regression model predicting narrative condition (truthful vs. falsehood). All predictors were standardized. Positive coefficients indicate higher log-odds of a narrative being classified as a falsehood. Rows highlighted in light grey indicate statistically significant effects.

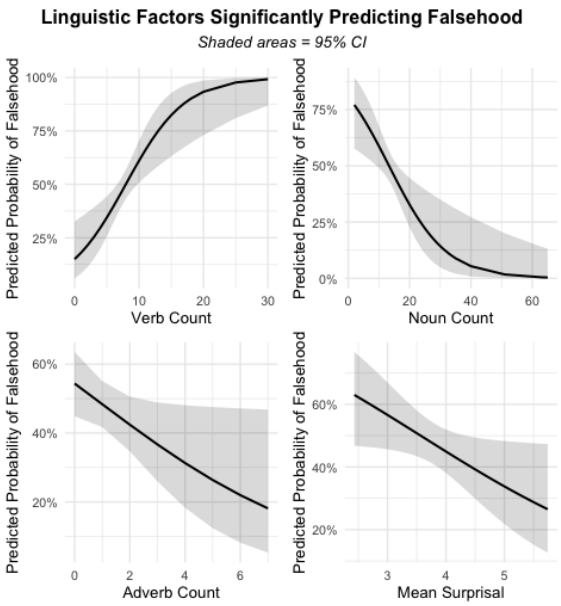

Figure 1: Effects of significant linguistic predictors on the probability of falsehood narratives. Shaded bands represent 95% confidence intervals.

tion may manifest in subtle differences in how messages are planned and produced. These processes are not necessarily universal; they may vary across languages due to typological properties and discourse norms (Futrell et al., 2015; Gibson et al., 2019; Husain et al., 2014). Identifying reliable cues of deception requires examining both language-specific patterns and factors that hold across languages. By focusing on Hindi, a South Asian Subject-Object-Verb (SOV) language with distinct syntactic and discourse characteristics, this study addresses a significant gap in cross-linguistic deception research. By leveraging psycholinguistic measures of linguistic complexity that have been shown to generalize across languages (Futrell et al., 2015; Wilcox et al., 2023), we aimed to uncover universal markers of deceptive language that are not tied to any single linguistic system.

Our results show that verb count, noun count, adverb count, and mean surprisal significantly predict falsehood in Hindi narratives. Falsehood narratives showed a higher verb count in our study, suggesting a greater focus on actions rather than entities. This finding aligns with reports that deceptive texts often emphasize motion or action verbs, which allow speakers to construct plausible narratives without committing to specific, verifiable details (Nahari et al., 2014; Newman et al., 2003; Sarzynska-Wawer et al., 2023). Such action-focused constructions may be easier to maintain and monitor under cognitive load. Additionally, the lower adverb count in falsehood narratives suggests reduced elaboration, consistent with a strategy of maintaining simplicity and coherence when cognitive resources are constrained.

The higher noun count observed in truthful narratives suggests greater specificity and the inclusion of more concrete, potentially verifiable details such as entities, locations, or names. This pattern aligns with the verifiability approach, which posits that truthful accounts contain more details that can, in principle, be checked (Kleinberg et al., 2018; Nahari et al., 2014). The differences in noun count, along with the absence of differences in type–token ratio, are also consistent with Porter and Yuille (1996), indicating that specificity may be driven more by content selection that favors more concrete entities than by lexical diversity.

Although we expected false opinions to exhibit higher surprisal due to their lack of grounding in reality, our results revealed the opposite pattern: false opinions were associated with lower surprisal. Our findings suggest that falsehoods in Hindi rely on more predictable lexical sequences. Lower surprisal corresponds to routinized, highly predictable constructions that are easier to plan, generate, and monitor (Demberg & Keller, 2008; Levy, 2008; Rajkumar et al., 2016). The predictability of these constructions may reduce processing effort for listeners (Lowder et al., 2018; Wilcox et al., 2023), giving rise to an illusion of truthfulness (Goupil et al., 2021; Mac Giolla et al., 2013).

We did not observe differences in mean dependency distance between truthful and deceptive narratives, indicating comparable levels of syntactic complexity across conditions. This finding mirrors results reported for Polish (Sarzynska-Wawer et al., 2023), suggesting that deception may not necessarily reduce syntactic complexity (as operationalized by dependency distance) in written opinion tasks. Similarly, like Sarzynska-Wawer et al. (2023), we found no differences in the use of negation markers. However, whereas Sarzynska-Wawer et al. (2023) reported differences in pronouns and infinitives, our results point to differences in the use of nouns, verbs, and adverbs in Hindi. These contrasts underscore the importance of language-specific analysis and suggest that the linguistic realization of deception may depend on typological

and discourse-level factors.

Our findings align with prior research, showing that linguistic simplification is consistent with cognitive accounts of deception (Gombos, 2006; Newman et al., 2003; Sarzynska-Wawer et al., 2023; Vrij et al., 2010; Walczyk et al., 2003; Zuckerman et al., 1981). Vrij et al. (2010) outlines several reasons why lying is cognitively more demanding than truth-telling, including the need to maintain internal consistency, align with what others know, actively monitor credibility, assess the listener's reactions, suppress truthful information, and sustain a fabricated mental representation. Some of these demands are reduced in written communication. For example, the absence of real-time monitoring of non-verbal cues and limited feedback from the interlocutor. In the present study, eliciting true and false opinions in a written format enabled us to isolate linguistic cues that reflect differences in planning and production processes specifically associated with the fabrication of false opinions.

Overall, our results can be explained by two underlying mechanisms: strategic adaptation and structural constraints in language production. First, deceptive speakers may strategically adapt their language to align with culturally salient norms of truthfulness, such as sounding fluent, coherent, and natural (DePaulo et al., 2003; Goupil et al., 2021; Hartwig et al., 2010; Mac Giolla et al., 2013). Therefore, to sound credible, speakers strategically produce simpler and more routinized linguistic structures. From this perspective, linguistic simplification is not merely a byproduct of deception but a strategic adjustment designed to project honesty and credibility.

Second, linguistic simplicity in deceptive narratives can arise from the working memory constraints imposed by the language production system. This approach aligns with production-based accounts of language use, suggesting that speakers favor linguistically conservative strategies that minimize processing demands during conceptualization and formulation when under cognitive load (Bock & Levelt, 1994; Ferreira & Swets, 2002; Hartsuiker & Barkhuysen, 2006; Ivanova & Ferreira, 2019; Slevc, 2011). Fabricating false opinions demands additional resources for truth inhibition and behavioral monitoring (Gombos, 2006; Lane & Wegner, 1995; Vrij et al., 2010; Walczyk et al., 2003; Zuckerman et al., 1981), thereby constraining planning and production. Under these pressures, writers produce lower-surprisal texts with reduced informational specificity, reflecting a shift toward computational efficiency in lexical selection and grammatical encoding. In contrast, truthful narratives face fewer conceptual and monitoring constraints, enabling greater elaboration and entity-specific detail despite higher processing costs. This pattern underscores that linguistic efficiency in falsehood production arises from inherent constraints in the language production system rather than strategic stylistic choices.

We acknowledge several limitations of this study. The use of online data collection and the false-opinion paradigm may constrain ecological validity. Responses to socially controversial topics may be influenced by social desirability biases and the consistent ordering of the elicitation procedure, whereby participants produced truthful accounts before deceptive ones, may have introduced order effects. The limited range of topics further restricts generalizability. Future research should replicate these findings using larger datasets, a broader range of communicative contexts, and additional languages. Extending these metrics to spoken data would also enable the examination of how linguistic complexity interacts with prosodic and temporal cues.

Conclusion

We examined the linguistic cues of falsehood in written Hindi narratives through the lens of language processing complexity. By applying metrics informed by Psycholinguistics literature to a corpus of truthful and false opinions in a typologically different language, we sought to move beyond traditional surface cues and uncover the cognitive mechanisms underlying the planning and production of false messages. Our findings show that false narratives in Hindi are characterized by lower surprisal and systematic differences in lexical category usage, particularly involving verbs, nouns, and adverbs.

Our results suggest that predictability and production ease play a central role in written falsehood. Lower surprisal in false texts points to a reliance on more predictable word sequences, supporting the notion that writers employ linguistically conservative strategies to sound more fluent. This aligns with prior work linking deception to the cognitive load account while challenging the view that truth-telling necessarily yields more coherent discourse. The higher noun count in truthful narratives indicates greater specificity and inclusion of verifiable details, consistent with the verifiability approach. In contrast, false narratives were marked by actions over entities, evident in higher verb frequencies and reduced adverbial elaboration, reflecting a strategy aimed at preserving plausibility while minimizing detail.

Notably, no differences emerged in syntactic complexity as measured by mean dependency distance, suggesting that truthful and false narratives may be structurally comparable in this context. Cross-linguistic comparisons reveal both convergences and divergences, highlighting the need for language-specific and typologically sensitive approaches to the study of falsehood in discourse.

Despite some limitations concerning task design and topic scope, this study provides novel psycholinguistic evidence from Hindi and demonstrates the utility of processing-based metrics for analyzing falsehoods. Future research should extend this framework to spoken data, larger and more heterogeneous datasets, and additional languages, as well as explore how these metrics can inform computational models of falsehood detection.

References

Arciuli, J., Mallard, D., & Villar, G. (2010). "um, i can tell you're lying": Linguistic markers of deception versus truth-

telling in speech. *Applied psycholinguistics*, *31*(3), 397–411.

Austin, J. L. (1975). *How to do things with words*. Harvard university press.

Blandón-Gitlin, I., Fenn, E., Masip, J., & Yoo, A. H. (2014). Cognitive-load approaches to detect deception: Searching for cognitive mechanisms. *Trends in cognitive sciences*, *18*(9), 441–444.

Bock, J., & Levelt, W. (1994). Sentence production. *M. Gernsbacher (Ed.)*

Bok, S. (2011). *Lying: Moral choice in public and private life*. Vintage.

Burgoon, J. K., Blair, J. P., Qin, T., & Nunamaker Jr, J. F. (2003). Detecting deception through linguistic analysis. *International Conference on Intelligence and Security Informatics*, 91–101.

Burgoon, J. K., & Buller, D. B. (1994). Interpersonal deception: Iii. effects of deceit on perceived communication and nonverbal behavior dynamics. *Journal of Nonverbal Behavior*, *18*(2), 155–184.

Capuozzo, P., Lauriola, I., Strapparava, C., Aiolli, F., Sartori, G., et al. (2020). Decop: A multilingual and multi-domain corpus for detecting deception in typed text. *Proceedings of the 12th language resources and evaluation conference*, *12*, 1423–1430.

Carson, T. L. (2006). The definition of lying. *Noûs*, *40*(2), 284–306.

Carter, G.-A., & Hoffman, P. (2024). Discourse coherence modulates use of predictive processing during sentence comprehension. *Cognition*, *242*, 105637.

Christiansen, M. H., & MacDonald, M. C. (2009). A usage-based approach to recursion in sentence processing. *Language Learning*, *59*, 126–161.

Corley, C. D., & Mihalcea, R. (2005). Measuring the semantic similarity of texts. *Proceedings of the ACL workshop on empirical modeling of semantic equivalence and entailment*, 13–18.

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

DePaulo, B. M., Lassiter, G. D., & Stone, J. L. (1982). Attentional determinants of success at detecting deception and truth. *Personality and social psychology bulletin*, *8*(2), 273–279.

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin*, *129*(1), 74.

Ekman, P. (1989). Why lies fail and what behaviors betray a lie. In *Credibility assessment* (pp. 71–81). Springer.

Ekman, P., & Friesen, W. V. (1969). Nonverbal leakage and clues to deception. *Psychiatry*, *32*(1), 88–106.

Feng, F., Yang, Y., Cer, D., Arivazhagan, N., & Wang, W. (2022). Language-agnostic bert sentence embedding. *Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 878–891.

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.

Frazier, L. (1985). Syntactic complexity. *Natural language parsing: Psychological, computational, and theoretical perspectives*, 129–189.

Futrell, R., Levy, R. P., & Gibson, E. (2020). Dependency locality as an explanatory principle for word order. *Language*, *96*(2), 371–412.

Futrell, R., Mahowald, K., & Gibson, E. (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences*, *112*(33), 10336–10341.

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.

Gillings, M. (2024). *Corpus linguistic approaches to deception detection*. Routledge.

Gombos, V. A. (2006). The cognition of deception: The role of executive processes in producing lies. *Genetic, social, and general psychology monographs*, *132*(3), 197–214.

Gordon Jr, R. G. (2005). Ethnologue, languages of the world. *http://www. ethnologue. com/*.

Goupil, L., Ponsot, E., Richardson, D., Reyes, G., & Aucouturier, J.-J. (2021). Listeners' perceptions of the certainty and honesty of a speaker are associated with a common prosodic signature. *Nature communications*, *12*(1), 861.

HaCohen-Kerner, Y., Dilmon, R., Friedlich, S., & Cohen, D. N. (2015). Distinguishing between true and false stories using various linguistic features. *Proceedings of the 29th Pacific Asia Conference on Language, Information and Computation: Posters*, 176–186.

Hale, J. (2001). A probabilistic earley parser as a psycholinguistic model. *Second meeting of the north american chapter of the association for computational linguistics*.

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(1-3), 181–204.

Hartwig, M., Granhag, P. A., Strömwall, L. A., & Doering, N. (2010). Impression and information management: On the strategic self-regulation of innocent and guilty suspects. *The Open Criminology Journal*, *3*(1), 10–16.

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

Husain, S., Vasishth, S., & Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from hindi. *PloS one*, *9*(7), e100986.

Ivanova, I., & Ferreira, V. S. (2019). The role of working memory for syntactic formulation in language production. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *45*(10), 1791.

Jaeger, T. F., & Tily, H. (2011). On language 'utility': Processing complexity and communicative efficiency. *Wiley Interdisciplinary Reviews: Cognitive Science*, *2*(3), 323–335.

Johnson, M. K., & Raye, C. L. (1981). Reality monitoring. *Psychological review*, *88*(1), 67.

Joshi, R., Singla, K., Kamath, A., Kalani, R., Paul, R., Vaidya, U., Chauhan, S. S., Wartikar, N., & Long, E. (2024). Adapting multilingual llms to low-resource languages using continued pre-training and synthetic corpus. *arXiv preprint arXiv:2410.14815*. https://arxiv.org/abs/2410.14815

Kleinberg, B., Mozes, M., Arntz, A., & Verschuere, B. (2018). Using named entities for computer-automated verbal deception detection. *Journal of forensic sciences*, *63*(3), 714–723.

Kun, S., Qiuying, W., & Xiaofei, L. (2023). An interpretable measure of semantic similarity for predicting eye movements in reading. *Psychonomic Bulletin & Review*, *30*(4), 1227–1242.

Lane, J. D., & Wegner, D. M. (1995). The cognitive consequences of secrecy. *Journal of personality and social psychology*, *69*(2), 237.

Leal, S., Vrij, A., Mann, S., & Fisher, R. P. (2010). Detecting true and false opinions: The devil's advocate approach as a lie detection aid. *Acta psychologica*, *134*(3), 323–329.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, *14*(1), 41–104.

Levine, T. R. (2014). Truth-default theory (tdt) a theory of human deception and deception detection. *Journal of Language and Social Psychology*, *33*(4), 378–392.

Levinson, S. C. (1983). Pragmatics. *Cambridge UP*.

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

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

Loconte, R., Russo, R., Capuozzo, P., Pietrini, P., & Sartori, G. (2023). Verbal lie detection using large language models. *Scientific reports*, *13*(1), 22849.

Lowder, M. W., Choi, W., Ferreira, F., & Henderson, J. M. (2018). Lexical predictability during natural reading: Effects of surprisal and entropy reduction. *Cognitive science*, *42*, 1166–1183.

Loy, J. E., Rohde, H., & Corley, M. (2018). Cues to lying may be deceptive: Speaker and listener behaviour in an interactive game of deception. *Journal of Cognition*, *1*(1), 42.

Mac Giolla, E., Granhag, P., & Liu-Jönsson, M. (2013). Markers of good planning behavior as a cue for separating true and false intent. psych journal, 2 (3), 183–189.

Mahon, J. E. (2008). The definition of lying and deception.

Masip, J., Bethencourt, M., Lucas, G., SEGUNDO, M. S.-S., & Herrero, C. (2012). Deception detection from written accounts. *Scandinavian Journal of Psychology*, *53*(2), 103–111.

McDonough, K., & Trofimovich, P. (2011). How to use psycholinguistic methodologies for comprehension and production. *Research methods in second language acquisition: A practical guide*, 117–138.

Mehrabian, A. (1971). Nonverbal betrayal of feeling. *Journal of Experimental Research in Personality*.

Meibauer, J. (2019). *The oxford handbook of lying*. Academic.

Mihalcea, R., & Strapparava, C. (2009). The lie detector: Explorations in the automatic recognition of deceptive language. *Proceedings of the ACL-IJCNLP 2009 conference short papers*, 309–312.

Miller, G. A., & Charles, W. G. (1991). Contextual correlates of semantic similarity. *Language and cognitive processes*, *6*(1), 1–28.

Nahari, G., Vrij, A., & Fisher, R. P. (2014). Exploiting liars' verbal strategies by examining the verifiability of details. *Legal and Criminological Psychology*, *19*(2), 227–239.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and social psychology bulletin*, *29*(5), 665–675.

Nichols, J. (2009). Linguistic complexity: A comprehensive definition and survey. In *Language complexity as an evolving variable* (pp. 110–125). Oxford University Press.

Patil, U., Vasishth, S., Bartek, B., & Lewis, R. (2008). The interplay of locality and surprisal. *CUNY Sentence Processing Conference, Chapel Hill, NC*.

Pennebaker, J. W., & Chew, C. H. (1985). Behavioral inhibition and electrodermal activity during deception. *Journal of personality and social psychology*, *49*(5), 1427.

Pérez-Rosas, V., & Mihalcea, R. (2014). Cross-cultural deception detection. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, 440–445.

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior*, *20*(4), 443–458.

Qi, P., Zhang, Y., Zhang, Y., Bolton, J., & Manning, C. D. (2020). Stanza: A Python natural language processing toolkit for many human languages. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics: System Demonstrations*. https://nlp.stanford.edu/pubs/qi2020stanza.pdf

Rajkumar, R., Van Schijndel, M., White, M., & Schuler, W. (2016). Investigating locality effects and surprisal in written english syntactic choice phenomena. *Cognition*, *155*, 204–232.

Sarzynska-Wawer, J., Pawlak, A., Szymanowska, J., Hanusz, K., & Wawer, A. (2023). Truth or lie: Exploring the language of deception. *Plos one*, *18*(2), e0281179.

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

Slevc, L. R. (2011). Saying what's on your mind: Working memory effects on sentence production. *Journal of experimental psychology: Learning, memory, and cognition*, *37*(6), 1503.

Sravanthi, P., & Srinivasu, B. (2017). Semantic similarity between sentences. *International Research Journal of Engineering and Technology (IRJET)*, *4*(1), 156–161.

Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from verb-final structures. *Language and Cognitive Processes*, *25*(4), 533–567.

Vrij, A., Edward, K., Roberts, K. P., & Bull, R. (2000). Detecting deceit via analysis of verbal and nonverbal behavior. *Journal of Nonverbal behavior*, *24*(4), 239–263.

Vrij, A., Granhag, P. A., & Porter, S. (2010). Pitfalls and opportunities in nonverbal and verbal lie detection. *Psychological science in the public interest*, *11*(3), 89–121.

Walczyk, J. J., Roper, K. S., Seemann, E., & Humphrey, A. M. (2003). Cognitive mechanisms underlying lying to questions: Response time as a cue to deception. *Applied Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition*, *17*(7), 755–774.

Wilcox, E. G., Pimentel, T., Meister, C., Cotterell, R., & Levy, R. P. (2023). Testing the predictions of surprisal theory in 11 languages. *Transactions of the Association for Computational Linguistics*, *11*, 1451–1470.

Willem, J. M. L. (1989). *Speaking: From intention to articulation*. MIT Press.

Yadav, H., Vaidya, A., Shukla, V., & Husain, S. (2020). Word order typology interacts with linguistic complexity: A cross-linguistic corpus study. *Cognitive science*, *44*(4), e12822.

Zafar, M., & Husain, S. (2023). Dependency locality influences word order during production in sov languages: Evidence from hindi. *Proceedings of the annual meeting of the cognitive science society*, *45*(45).

Zuckerman, M., Amidon, M. D., Bishop, S. E., & Pomerantz, S. D. (1982). Face and tone of voice in the communication of deception. *Journal of Personality and Social Psychology*, *43*(2), 347.

Zuckerman, M., DePaulo, B. M., & Rosenthal, R. (1981). Verbal and nonverbal communication of deception. In *Advances in experimental social psychology* (pp. 1–59, Vol. 14). Elsevier.