To appear in Proc. of ICASSP 2025

# Psycholinguistic Features Predict Word Duration in Hindi Read Aloud Speech

Rajakrishnan Rajkumar[1], Sneha Raman[2], Aadya Ranjan[1], Mildred Pereira[2], Nagesh Nayak[2] and Preeti Rao[2]
*[1]LTRC IIIT Hyderabad [2]Department of Electrical Engineering, Indian Institute of Technology Bombay, Mumbai, India*

***Abstract*—Reliable assessment of oral reading fluency (ORF) is of great importance in foundational literacy missions globally. For the design of level appropriate testing passages, text difficulty has traditionally been based on coarse-grained measures of readability like the Flesch–Kincaid score. We present a novel study where we deploy psycholinguistic measures of reading difficulty from Natural Language Processing to predict the duration of words in Hindi read-aloud speech. We test the hypotheses that expectation-based measures of linguistic complexity are significant predictors of word duration in Hindi read-aloud speech. We validate the stated hypotheses by estimating surprisal measures inspired from Surprisal Theory of sentence comprehension and introduce a novel measure of orthographic complexity to model the intricacies of the Hindi script. Cognitive modelling experiments were conducted on a dataset of six Hindi short stories read aloud by 5 expert readers, containing 2 measures of word duration. Our results show that both surprisal as well as the orthographic complexity measures are significant predictors of word duration. In contrast with long words, we find duration reducing with increased orthographic complexity in the case of short words. The variation between individual speakers in terms of word duration is very low and the variance in the data is caused by the properties of the words used in the text. Finally, we reflect on the implications of our work for cognitive models of language production and for ORF assessment.**

***Index Terms*—Orthographic complexity, Word duration, Surprisal Theory, Reading aloud**

## I. Introduction

Readability measurement, or the evaluation of text difficulty, is important across education settings. For instance, the assessment of oral reading fluency (ORF) - a component of foundational literacy drives globally - requires the leveling of text passages to the expected proficiency in a given language [1]. Traditional methods capture simple lexico-syntactic features such as sentence lengths and syllables per word, neglecting aspects such as vocabulary, syntax, and cohesion. More recently, NLP features corresponding to predictability from word frequency and context have been used to quantify text difficulty, which in turn is known to be linked to reading speed and comprehension [2]. The present work addresses the potential of computational features obtained from text processing in predicting speech production attributes such as word duration in reading aloud (refer to the edited volume [3] to get a comprehensive overview of the cognitive processes in reading aloud). We extend the large body of work linking text features with eye-tracking derived reading times obtained in silent reading to the less researched context of word durations from recorded speech in reading aloud. Motivated by prior work, we test the following hypotheses:

1) *High values of surprisal measures predict high word durations*: A substantial body of prior work has shown that such complexity measures account for word duration in silent reading [4], [5] as well as spontaneous speech [6]–[8]. However, the use of surprisal measures to model word duration in the reading aloud paradigm (combining both silent reading and spoken language production) is an under explored trajectory. To the best of our knowledge, [9] is the sole work deploying surprisal measures to model word duration in Hindi read-aloud speech. We extend this work by using a better experimental design and a novel measure of orthographic complexity, as described below.
2) *High values of orthographic complexity correspond to high word duration*: Orthographic complexity has been investigated for various languages that have complex elements in their scripts. Some examples of orthographic complexity include a larger number of strokes per character in Chinese [10] and the presence of diacritics and consonant clusters in Kannada, Malayalam and Hindi [11], [12] [13]. In the context of reading aloud and reading assessment, [14] reported that *longer words and words with more maatras (diacritics) and samyuktaksharas (conjunct consonants) often make words more difficult to decode*. These observations motivate us to investigate the effects of the orthographic complexity of words on their reading duration with a suitably defined measure of complexity.

We test the stated hypotheses by conducting cognitive modelling experiments on a dataset of six Hindi short stories read aloud by five expert readers to predict two measures of duration. Our main contribution is that we extend the prior work motivating our hypotheses (as cited above) by validating them in the presence of a comprehensive host of factors in a language other than English. To the best of our knowledge, this is the first work that explores reading aloud production times in Hindi using a repeated measures design (i.e. every passage is read by 5 speakers). The only known previous work on Hindi [9] uses a corpus where 2 distinct passages are read aloud by separate speakers.

## II. Background

The following subsections provide essential background on the Hindi language and its orthography, and Surprisal Theory.

The first author acknowledges support from IIIT Hyderabad's seed grant number IIIT/R&D Office/Seed-Grant/2024-25/007

### A. *Hindi Language and Script*

Hindi is written in the Devanagari alphasyllabary-based script, consisting of 47 characters in total (33 consonants in total) along with conjunct consonants and diacritics [13]. It is a transparent script i.e, exhibits a high degree of letter-sound correspondence. In scripts such as the Japanese Kanji and Korean Hanja, a "greater cost of processing the complex orthographies was attributed to their relatively opaque orthography-to-phonology mapping, rather than their complex visuospatial layout" [15]. The effects of orthographic complexity in orthographically transparent scripts like Hindi are underexplored. The few studies we encountered included an exploration of the neurocognitive correlates of visually complex and simple Hindi words [15], and the effect of orthographic complexity on word durations derived from isolated word naming tasks [16] and eye-tracking [17].

### B. *Surprisal Theory*

Surprisal Theory [18], [19] is an information-theoretic characterization of language comprehension which defines the *surprisal* of the $(k+1)^{th}$ word, $w_{k+1}$, as the negative logarithm of conditional probability of word, $w_{k+1}$ given the preceding context, which can be either sequence of words or a syntactic tree. Both these kinds of surprisal have been shown to predict eye movement durations in language processing [4], [8], [20].

## III. Read Aloud Dataset and Audio Processing

Motivated by our long-term goal of measuring text difficulty for elementary school assessments, we enlist the assistance of language teachers to select six Hindi language grade 3 level passages (equal numbers of narrative and descriptive texts), each containing 130-150 words. The number of unique words (i.e. word types) is 329. Passages are available here: https://github.com/mildredpereira/Canonical-text-hindi-stories

For this pilot study, we had five model readers (3F, 2M) - experienced language teachers teaching Hindi in English medium schools - read each passage aloud by recording the audio on their personal mobile devices to obtain a total of 30 utterances. The model readers read at a moderate pace with the expected expressiveness, a characteristic of fluent reading, or reading with comprehension [21]. After checking for consistency with the text prompts, the audio was forced-aligned using an available acoustic model trained on adult native Hindi read speech to achieve segmentation at the word level. Thus, for each of the audio recordings, the duration of each word was obtained together with the duration of any preceding silence. Silence during reading aloud is on account of effects terms as a spillover [22] as well as parafoveal preview [23]. Spillover is the impact of cognitive processing of previous words, while parafoveal preview induces delays on account of eyes looking ahead to absorb information about upcoming words. The average word duration was similar across speakers (mean value of 340 ms).

## IV. Methodology

### A. *Linguistic/lexical metrics from text*

*1) word length:* Word length was calculated simply as the total number of consonants, vowels and matras in the word. For example, इसलिए (isliye) has a word length of 5 since it has 2 consonants (स, ल), two vowels (इ, ए) and one matra (ि). See Figure 1 for more examples.

*2) Word frequency:* Word frequency, the count of each target word, was obtained from the EMILLE Hindi corpus [24].

*3) Trigram surprisal:* Trigram surprisal is defined as the negative log of the probability of target word $w_{k+1}$ given two preceding words: $S_{k+1} = -\log P(w_{k+1}|w_{k-1}, w_k)$

For each word in a sentence, we computed this measure using a trigram language model trained on the EMILLE corpus of written text with mixed genre [24] using the SRILM toolkit [25] with Good-Turing discounting smoothing algorithm.

*4) PCFG surprisal:* Surprisal estimates using a Probabilistic Context Free Grammar (PCFG) is defined as the negative log probability of target word $w_{k+1}$given contextual syntactic tree (T):

$$S_{k+1} = -\log P(w_{k+1}|T) = \log \frac{\sum_T P(T, w_1 \ldots w_k)}{\sum_T P(T, w_1 \ldots w_{k+1})}$$

For each word, PCFG surprisal was estimated by training an incremental probabilistic left-corner parser [26] on 13,000 phrase structure trees from the Hindi-Urdu Treebank corpus [27] of newswire text using the ModelBlocks toolkit (https://github.com/modelblocks).

| Word | IPA | Complex Elements | Orthographic Complexity | Word length |
|---|---|---|---|---|
| एक | ekɐ | No complex elements | 0 | 2 |
| कि | kɪ | **1** diacritic (ि) | 10 | 2 |
| राज्य | ɾadʒjɐ | **1** diacritic (ा),<br>**1** conjunct consonant (ज्य) | 20<br>(10+10) | 3.5 |
| बगीचे | bɐgitʃe | **2** diacritics (ी, े) | 20 | 5 |
| मंत्रियों | mẽt̪ɾɪjõ | **1** conjunct consonant (त्र),<br>**4** diacritics (ं, ि, ो, ं) | 50<br>(10+40) | 7 |

Fig. 1. Examples of orthographic complexity computation

*5) Orthographic complexity:* Orthographic complexity, as distinct from the word length, was computed by counting the occurrences of complex elements in a word such as diacritics and conjunct consonants. The word in the Devanagari script was decoded into its Unicode form, which breaks it down into its constituent elements as separate characters. A penalty of 10 was assigned every time a complex character element was encountered. Please see Figure 1 for examples.

### B. *Statistical Analyses*

We trained Linear Mixed Models (LMMs) to predict per-word duration measures (transformed to a logarithmic scale following previous work). All the independent variables were normalized to $z$-scores. The $lme4$ package in R was used to perform our regression experiments using a very basic model, given below in R GLM format (independent variable $\sim$ dependent variables + 1| random intercept terms to model random effects pertaining to speakers and items):

$$Duration \sim Orthographic\ complexity + PCFG\ surprisal + Trigram\ surprisal + Word\ length + Word\ frequency + (1|Speaker) + (1|word)$$

Word frequency and word length are conceived as control factors in Linear Mixed Models [28, LMMs] based on long-standing findings from the liteature [29]–[31].

## V. Results

We now present our correlation and regression results.

### A. Correlation Results

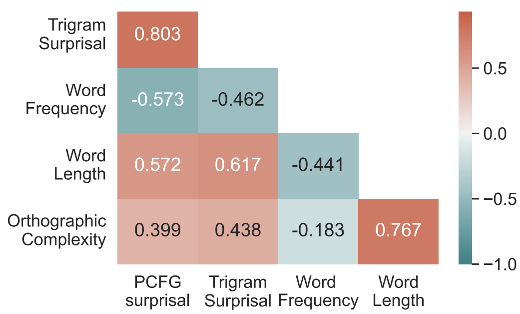


Fig. 2. Pearson's correlation coefficients amongst the different predictors. All correlations shown above are significant p<0.001

Correlation results (Figure 2) show that both surprisal measures are highly correlated with each other. The high degree of lexicalization in the PCFG parser explains this result. Orthographic complexity and length are directly correlated, along expected lines. Word frequency and word length are inversely correlated *i.e.*, short words are more frequent than long words, as known for long in linguistics theory [31].

### B. Regression Experiments

TABLE I
FIXED EFFECTS OF AN LMM PREDICTING WORD DURATIONS (3790 DATA POINTS; SIGNIFICANCE FOR |T|=2 THRESHOLD SHOWN IN BOLD)

| Predictors | Estimate | Std. Error | t-value |
|---|---|---|---|
| (Intercept) | 0.069 | 0.084 | 0.820 |
| **PCFG surprisal** | 0.079 | 0.0268 | 2.968 |
| Trigram surprisal | 0.006 | 0.024 | 0.264 |
| **Frequency** | -0.185 | 0.037 | -4.987 |
| **Length** | 0.554 | 0.035 | 15.394 |
| **Orthographic complexity** | -0.127 | 0.031 | -3.981 |
| *Model with interaction term* | | | |
| **Word length x Orthographic complexity** | -0.078 | 0.016 | -4.981 |

We used two separate LMMs to predict word duration as well as word duration including previous silence (Tables I and III show the results). Both our LMMs show that high-frequency words have shorter duration compared to their lower-frequency counterparts (negative coefficient). These controls are based on long-standing findings in the literature that high-frequency words are read out faster than their lower-frequency counterparts as their activation requires lower input from the visual features of the input letters [29]. Thus activation induces faster word retrieval due to swifter search time [30]. Longer words are produced with higher duration compared to shorter words in both settings.

The results of our first LMM to predict the duration of words (Table I) show that words with high values of PCFG surprisal have longer duration compared to their counterparts with lower values. (positive regression coefficient). The impact of word length and PCFG surprisal mirrors previous studies on English spontaneous speech production [7] and Hindi comprehension studies [5], [17]. Trigram surprisal is not a significant predictor of word duration. The high degree of lexicalization of PCFG surprisal, as it includes word forms as features, can account for this effect. We also found that the random effects of the model indicate that between-speakers variance accounts for only 7.4% of the total variance, while words (items) account for around 29.21% of the variance. This indicates that we have successfully controlled for aspects such as reading proficiency by using language teachers, who can be assumed to be following normative and prescriptive standards.

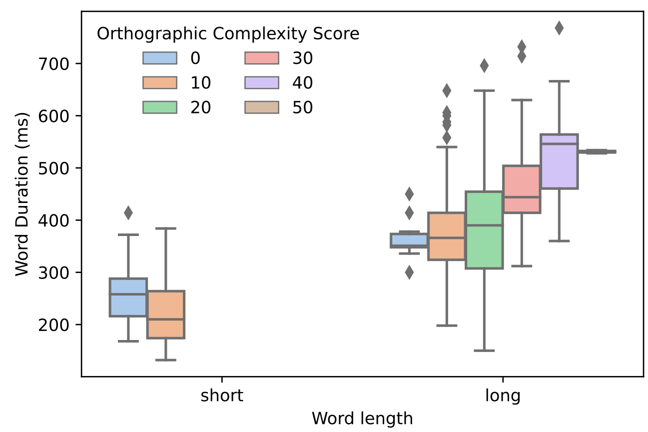


Fig. 3. Covariation of ortho complexity and duration, binned by word length

An unexpected finding is that orthographic complexity has a negative coefficient in the LMM prediction of word duration. Length and orthographic complexity display a high positive correlation (Figure 2). Therefore, both features contribute to word duration, but in different ways. A separate model with all other features and containing an interaction term between length and orthographic complexity shows that the effect of orthographic complexity on word duration decreases by 0.078 with every unit increase in word length.

We explored this further by placing the words into two bins: short (word length $<3$) and long words (word length $>=3$). As seen in Figure 3, for short words, increased orthographic complexity corresponded with shorter word duration. For instance, short, high orthographic complexity words such as था, सा, का, और, की, से, जो etc. were read faster compared to low orthographic complexity words such as एक, जब, वह, कर, गए, सब etc. This is opposite to the effect observed for long words. In addition, for the short length words, increased orthographic complexity also corresponded with lower trigram and PCFG surprisal, and higher word frequency.

TABLE II
LOG LIKELIHOOD TEST RESULTS COMPARING MODELS PREDICTING WORD DURATION (MODEL ON EACH ROW COMPARED TO THAT IN THE PREVIOUS ROW AFTER ADDING FEATURES)

| **Model** | **Features** | $\mathbf{R^2}$ | **Log-likelihood** | **Chi-square** | **p-value** |
|---|---|---|---|---|---|
| Baseline | Frequency, length | 0.485 | -3152.9 | | |
| Model-1 | +Orthographic complexity | 0.505 | -3145.6 | 14.693 | $< .001$ |
| Full model | +Surprisal measures | 0.508 | -3140.1 | 10.906 | $< .05$ |

We also did comparisons of models by doing log likelihood tests by adding various features to a baseline model (word frequency and length only). Table II shows that adding orthographic complexity to the baseline model results in a significantly different model with greater log-likelihood. Adding the surprisal measures results in a significantly better model over the previous model which contained baseline features and orthographic complexity. The $R^2$ values depicted in the table also show how adding more predictors to the baseline model results in models which account for the proportion of the variance of the dependent variable better.

TABLE III
FIXED EFFECTS OF AN LMM PREDICTING WORD DURATION INCLUDING PREVIOUS SILENCE (3790 DATA POINTS; SIGNIFICANCE FOR |T|=2 THRESHOLD SHOW IN BOLD)

| **Predictors** | **Estimate** | **Std. Error** | **t-value** |
|---|---|---|---|
| (Intercept) | 0.055 | 0.074 | 0.745 |
| **PCFG surprisal** | -0.107 | 0.033 | -3.246 |
| **Trigram surprisal** | 0.136 | 0.029 | 4.589 |
| **Frequency** | -0.144 | 0.049 | -2.896 |
| **Length** | 0.381 | 0.049 | 7.775 |
| **Orthographic complexity** | -0.127 | 0.044 | -2.883 |
| *Model with interaction terms* | | | |
| **PCFG surprisal x trigram surprisal** | -0.053 | 0.0228 | -2.308 |
| Word length x Orthographic complexity | 0.0123 | 0.022 | 0.569 |

Now we turn to the results of a LMM to predict word duration including preceding silence (Table III). Words with high values of length and trigram surprisal have longer duration compared to their counterparts with lower values. (positive regression coefficient). Low frequency and low PCFG surprisal words have longer duration compared to their counterparts with higher values of these measures. As for the case of word duration prediction, lower orthographic complexity words tend to have higher duration compared to more complex words (negative coefficient). Another model with all other features and containing an interaction term shows that the effect of trigram surprisal on reading-aloud times decreases by 0.05 with every unit increase in PCFG surprisal, while the word length vs orthographic complexity interaction is not significant.

## VI. DISCUSSION AND CONCLUSIONS

Our study modelling read aloud duration using psycholinguistically motivated features shows that both syntactic and trigram surprisal are significant predictors of the two word duration measures in our study. High PCFG surprisal words have longer actual duration compared to their counterparts with lower values, while trigram surprisal is not a significant predictor at all. The lexicalization inherent in the parser we use (as both word and POS information is included) to estimate surprisal is overriding the effect of the trigram model.

Orthographic complexity has an effect independent of word length, *i.e.* reduced duration for increased orthographic complexity for short words and increased word duration for increased orthographic complexity for long words. For short words, the effects of increased orthographic complexity may have been overridden by their lower surprisal and higher word frequency, which led to shorter reading duration. Previous studies that investigated the effect of orthographic complexity in Hindi either involved simple stimuli (isolated words) [16] or did not find an effect [17]. Husain et. al [17] reflect that one of the reasons for this could be that their "complexity metric may not characterize the sources of difficulty correctly". Our study presents a novel orthographic complexity metric, and is the first to find an effect of orthographic complexity on read aloud word durations of Hindi connected text.

In the condition of word duration including previous silence, low PCFG surprisal but high trigram surprisal words have longer duration compared to their counterparts with high PCFG and low trigram surprisal. Here, orthographic complexity behaves exactly the same as in the actual word duration condition. In both conditions, the variation between individual speakers in terms of word duration is very low and the variance in the data is caused by the properties of the words used in the text.

Further, our results can be connected to the Dual Route Cascaded (DRC) Model of reading aloud and visual word recognition [32], which posits 3 routes for word processing: 1. Grapheme-Phoneme Correspondence. (GPC) 2. Lexical-semantic 3. Lexical non-semantic routes. Orthographic complexity effects are prominent when the GPC route is activated. The lexical-semantic route could be activated for high syntactic surprisal words resulting in high word duration. In contrast, the lexical non-semantic route might be activated in situations where speakers are pausing just before words with low syntactic complexity so as to put meaning construction processes on hold, in order to focus on word recognition aspects of reading. This conjecture is corroborated by the co-occurrence of high trigram surprisal words after silences. This is also consistent with the phenomenon of pre-focal pauses used in Indian languages before new information words [33].

The presented work demonstrates the validity of text measures in the estimation of speech production difficulty as represented by the word durations for the selected grade level material. Apart from word length, frequency and surprisal, we see how orthographic complexity plays an important role. By using expert readers, we have tried to keep the focus on text-intrinsic factors. Future work will involve extending the study to school grade 3 readers where extrinsic factors such as reading proficiency need to be included in the model [34]. Our work can be extended to other languages with transparent scripts and enhance practical applications like oral reading fluency assessment and text to speech synthesis.

# References


[1] R. G. Benjamin and P. J. Schwanenflugel, "Text complexity and oral reading prosody in young readers," *Reading Research Quarterly*, vol. 45, no. 4, pp. 388–404, 2010.

[2] S. Vajjala and D. Meurers, "On improving the accuracy of readability classification using insights from second language acquisition," in *Proceedings of the seventh workshop on building educational applications using NLP*, 2012, pp. 163–173.

[3] S. Sulpizio and S. Kinoshita, "Bridging reading aloud and speech production," *Frontiers in psychology*, vol. 7, p. 661, 2016.

[4] V. Demberg and F. Keller, "Data from eye-tracking corpora as evidence for theories of syntactic processing complexity," *Cognition*, vol. 109, no. 2, pp. 193–210, 2008. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0010027708001741

[5] A. Agrawal, S. Agarwal, and S. Husain, "Role of expectation and working memory constraints in Hindi comprehension: An eyetracking corpus analysis," *Journal of Eye Movement Research*, vol. 10, no. 2, 2017. [Online]. Available: https://bop.unibe.ch/index.php/JEMR/article/view/2968

[6] A. Bell, D. Jurafsky, E. Fosler-Lussier, C. Girand, M. Gregory, and D. Gildea, "Effects of disfluencies, predictability, and utterance position on word form variation in English conversation," *The Journal of the Acoustical Society of America*, vol. 113, no. 2, pp. 1001–1024, 2003.

[7] A. Bell, J. M. Brenier, M. Gregory, C. Girand, and D. Jurafsky, "Predictability effects on durations of content and function words in conversational English," *Journal of Memory and Language*, vol. 60, no. 1, pp. 92–111, 2009.

[8] V. Demberg, A. B. Sayeed, P. J. Gorinski, and N. Engonopoulos, "Syntactic surprisal affects spoken word duration in conversational contexts," in *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, ser. EMNLP-CoNLL '12. Stroudsburg, PA, USA: Association for Computational Linguistics, 2012, pp. 356–367. [Online]. Available: http://dl.acm.org/citation.cfm?id=2390948.2390992

[9] S. Ranjan, R. Rajkumar, and S. Agarwal, "Linguistic complexity and planning effects on word duration in Hindi read aloud speech," *Society for Computation in Linguistics*, vol. 5, no. 1, 2022.

[10] S. P. Liversedge, C. Zang, M. Zhang, X. Bai, G. Yan, and D. Drieghe, "The effect of visual complexity and word frequency on eye movements during Chinese reading," *Visual Cognition*, vol. 22, no. 3-4, pp. 441–457, 2014.

[11] P. Karanth, A. Mathew, and P. Kurien, "Orthography and reading speed: Data from native readers of Kannada," *Reading and Writing*, vol. 17, pp. 101–120, 2004.

[12] R. Shallam and A. Vaidya, "Towards measuring lexical complexity in Malayalam," in *ICON*, 2019. [Online]. Available: https://api.semanticscholar.org/CorpusID:234345291

[13] J. Vaid and A. Gupta, "Exploring word recognition in a semi-alphabetic script: The case of Devanagari," *Brain and Language*, vol. 81, pp. 679–90, 04 2002.

[14] S. Menon, R. Krishnamurthy, S. Sajitha, N. Apte, A. Basargekar, S. Subramaniam, M. Nalkamani, and M. Modugala, "Literacy research in Indian languages (LiRIL): Report of a three-year longitudinal study on early reading and writing in Marathi and Kannada," *Bangalore: Azim Premji University, New Delhi: Tata Trusts*, 2017.

[15] C. Rao and N. C. Singh, "Visuospatial complexity modulates reading in the brain," *Brain and Language*, vol. 141, pp. 50–61, 2015.

[16] J. Vaid and A. Gupta, "Exploring word recognition in a semi-alphabetic script: The case of Devanagari," *Brain and Language*, vol. 81, no. 1-3, pp. 679–690, 2002.

[17] S. Husain, S. Vasishth, and N. Srinivasan, "Integration and prediction difficulty in Hindi sentence comprehension: Evidence from an eye-tracking corpus," *Journal of Eye Movement Research*, vol. 8, no. 2, 2015. [Online]. Available: https://bop.unibe.ch/index.php/JEMR/article/view/2400

[18] J. Hale, "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*, ser. NAACL '01. Pittsburgh, Pennsylvania: Association for Computational Linguistics, 2001, pp. 1–8. [Online]. Available: http://dx.doi.org/10.3115/1073336.1073357

[19] R. Levy, "Expectation-based syntactic comprehension," *Cognition*, vol. 106, no. 3, pp. 1126 – 1177, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0010027707001436

[20] N. J. Smith and R. Levy, "The effect of word predictability on reading time is logarithmic," *Cognition*, vol. 128, no. 3, p. 302–319, 2013.

[21] L. S. Fuchs, D. Fuchs, M. K. Hosp, and J. R. Jenkins, "Oral reading fluency as an indicator of reading competence: A theoretical, empirical, and historical analysis," in *The role of fluency in reading competence, assessment, and instruction*. Routledge, 2001, pp. 239–256.

[22] D. Mitchell, *Mitchell, D.C. (1984) An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading.* Hillsdale, N.J.: Erlbaum., 01 1984.

[23] E. R. Schotter, B. Angele, and K. Rayner, "Parafoveal processing in reading," *Attention, Perception, & Psychophysics*, vol. 74, no. 1, pp. 5–35, 2012.

[24] P. Baker, A. Hardie, T. McEnery, H. Cunningham, and R. Gaizauskas, "EMILLE, a 67-million word corpus of Indic languages: Data collection, mark-up and harmonisation," in *Proceedings of the Third International Conference on Language Resources and Evaluation (LREC'02)*, M. González Rodríguez and C. P. Suarez Araujo, Eds. Las Palmas, Canary Islands - Spain: European Language Resources Association (ELRA), May 2002. [Online]. Available: http://www.lrec-conf.org/proceedings/lrec2002/pdf/319.pdf

[25] A. Stolcke, "SRILM — An extensible language modeling toolkit," in *Proc. ICSLP-02*, 2002.

[26] M. van Schijndel, A. Exley, and W. Schuler, "A model of language processing as hierarchic sequential prediction," *Topics in Cognitive Science*, vol. 5, no. 3, pp. 522–540, 2013.

[27] R. Bhatt, B. Narasimhan, M. Palmer, O. Rambow, D. M. Sharma, and F. Xia, "A multi-representational and multi-layered treebank for Hindi/urdu," in *Proceedings of the Third Linguistic Annotation Workshop*, ser. ACL-IJCNLP '09. Stroudsburg, PA, USA: Association for Computational Linguistics, 2009, pp. 186–189. [Online]. Available: http://dl.acm.org/citation.cfm?id=1698381.1698417

[28] J. C. Pinheiro and D. M. Bates, "Linear mixed-effects models: basic concepts and examples," *Mixed-effects models in S and S-Plus*, pp. 3–56, 2000.

[29] J. Morton, "Interaction of information in word recognition." *Psychological review*, vol. 76, no. 2, p. 165, 1969.

[30] K. I. Forster and S. M. Chambers, "Lexical access and naming time," *Journal of Verbal Learning and Verbal Behavior*, vol. 12, no. 6, pp. 627–635, 1973. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0022537173800428

[31] G. K. Zipf, *Human Behaviour and the Principle of Least Effort*. Addison-Wesley, 1949.

[32] M. Coltheart, K. Rastle, C. Perry, R. Langdon, and J. Ziegler, "Drc: a dual route cascaded model of visual word recognition and reading aloud." *Psychological review*, vol. 108, no. 1, p. 204, 2001.

[33] C. Féry, P. Pandey, and G. Kentner, "The prosody of focus and givenness in Hindi and Indian English," *Studies in Language*, vol. 40, no. 2, pp. 302–339, 2016.

[34] K. Sabu and P. Rao, "Predicting children's perceived reading proficiency with prosody modeling," *Computer Speech & Language*, vol. 84, p. 101557, 2024.