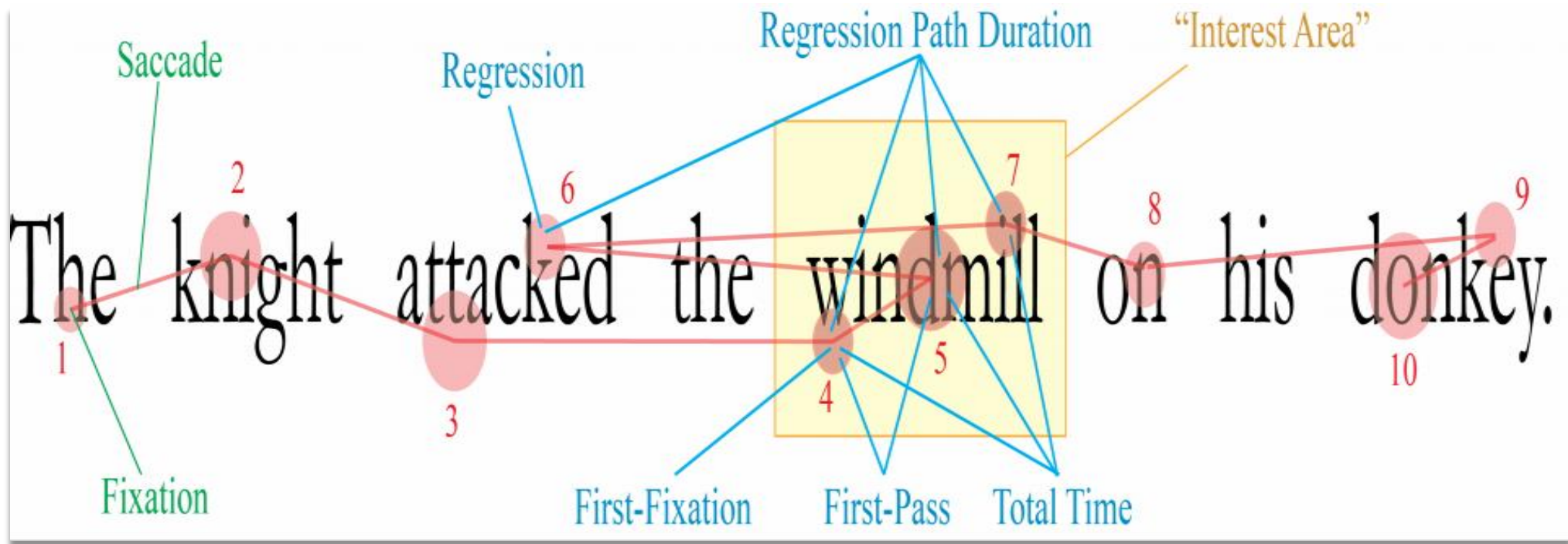


# INFREQUENT WORDS ARE MORE DIFFICULT TO COMPREHEND :

Analysis using gaze tracking By Shashank Sonkar

## Saccade & Terminology



$$\mathbf{Fixation\_time}_{\text{word}} = f(\text{frequency}_{\text{word}})$$

- Frequency less, fixation more
- With more of less frequent, fixation<sub>time</sub> rapidly decreases
- Frequency more, fixation less
- With more of more frequent, fixation<sub>time</sub> remains constant
- Short within word regression
- Parafoveal processing

Experiment on hindi language.

# The हिन्दी paragraph

हिंदी के आख्यान पढ़ने का हमेशा से मुझे बहुत शौक था। ये शौक मुझे धरोहर में अपनी माँ से लिया है, जो कि स्वयं हिंदी के आख्यान किताबी कीड़ो कि तरह चाट डालती थी। परन्तु हिंदी के आख्यान ढूँढ पाना बहुत कष्टदायक कार्य है। माँ ने कभी पुस्तके खरीदी ही नहीं थी, वे ग्रंथालय से पुस्तके ला ला के पढ़तीं और उन्हें लौटा देती थी। इस कारणवश पुस्तको का कभी संग्रह न हो सका। जैसे तैसे मैंने रबिन्द्रनाथ टैगोर की "काबुलीवाला" पढ़ी थी और आनंद से प्रफुल्लित हो उठी थी। परन्तु उसके बाद से कोई अवसर ही न मिल सका। लेकिन भगवान के आगे कहीं कुछ रुक पाया है? कुछ ही दिनों पहले मुझे एक ऐसा ही स्वर्णिम अवसर मिल गया। ऐसा जिसकी मुझे हमेशा से तलाश थी। लैंडमार्क में हिंदी पुस्तको का ढेर लगा था जिसे मैं उठा लायी ! अब एक एक कर के मैं वो सारी किताबें पढ़ूंगी, जो पहले न पढ़ सकी थी। इनमे से सब से ऊपर थी मुंशी प्रेमचंद की "नैराश्य लीला" और हरिवंश राइ बच्चन की "बचपन के साथ"। इस उपन्यास के बारे में मैंने माँ से बहुत सुना था।

# Expectations with the Design

- Considered Hindi corpus with 2861431 words
  - More fixation on infrequent words
    - आख्यान - 28 times
    - नैराश्य - 4 times
  - Less fixation on frequent words
    - भगवान - 500 times
    - पुस्तक - 300 times
    - NEW\* - टैगोर though infrequent(22) but Famous person.
      - Parafoveal processing should play a role here if the subject is familiar with the poet
  - आख्यान repeated 3 times - fixation time should decrease
  - पुस्तक repeated 3 times - fixation time should remain same
  - Short within-word regression on infrequent words like नैराश्य
- \* Frequent words and infrequent words have been kept of same length.

# Gaze Results 1 :

हिंदी के आख्यान पढ़ने का हमेशा से मुझे बहुत शौक था। ये शौक मुझे धरोहर में अपनी माँ से लिया है, जो कि स्वयं हिंदी के आख्यान किताबी कीड़ों कि तरह चाट डालती थी। परन्तु हिंदी के आख्यान ढूँढ पाना बहुत कष्टदायक कार्य है। माँ ने कभी पुस्तकें खरीदी ही नहीं थी, वे ग्रंथालय से पुस्तकें ला ला के पढ़तीं और उन्हें लौटा देती थी। इस कारणवश पुस्तकों का कभी संग्रह नहीं हो सका। जैसे जैसे मैंने रबिन्द्रनाथ टैगोर की "काबुलीवाला" पढ़ी थी और आनंद से प्रफुल्लित हो उठी थी। परन्तु उसके बाद से कोई अवसर ही न मिल सका। लेकिन भगवान के आगे कहीं कुछ रुक पाया है? कुछ ही दिनों पहले मुझे एक ऐसा ही स्वर्णिम अवसर मिल गया। ऐसा जिसकी मुझे हमेशा से तलाश थी। लैंडमार्क में हिंदी पुस्तकों का ढेर लगा था जिसे मैं उठा लायी। अब एक एक कर के मैंने सारी किताबें पढ़ूंगी, जो पहले न पढ़ सकी थी। इनमें से सब से ऊपर थी मुंशी प्रेमचंद की "बचपन के साथ" और हरिवंश राइ बच्चन की "बचपन के साथ"। इस उपन्यास के बारे में मैंने माँ से बहुत सुना था।

## Gaze Results 2:

हिंदी के आख्यान पढ़ने का हमेशा से मुझे बहुत शौक था। ये शौक मुझे धरोहर में अपनी माँ से लिया है, जो कि स्वयं हिंदी के आख्यान किताबी कीड़ों की तरह घाट डालती थी। परन्तु हिंदी के आख्यान छूंट पाना बहुत कष्टदायक कार्य है। माँ ने कभी पुस्तकें खरीदी ही नहीं थीं, वे ग्रंथालय से पुस्तकें ला ला के पढ़तीं और उन्हें लौटा देती थीं। इस कारणवश पुस्तकों का कभी संग्रह न हो सका। जैसे तैसे मैंने रबिन्द्रनाथ टैगोर की "काबुलीवाला" पढ़ी थी और आनंद से प्रफुल्लित हो उठी थी। परन्तु उसके बाद से कोई अवसर ही न मिल सका। लेकिन भगवान के आगे कहीं कुछ रुक पाया है? कुछ ही दिनों पहले मुझे एक ऐसा ही स्वर्णिम अवसर मिल गया। ऐसा जिसकी मुझे हमेशा से तलाश थी। लैंडमार्क में हिंदी पुस्तकों का ढेर लगा था जिसे मैं उठा लायी। अब एक एक कर के मैं वो सारी किताबें पढ़ूंगी, जो पहले न पढ़ सकी थी। इनमें से सब से ऊपर थी मुंशी प्रेमचंद की "नैराश्य लीला" और हारिवंश राइ बच्चन की "बचपन के साथ"। इस उपन्यास के बारे में मैंने माँ से बहुत सुना था।

# Observations :

- In 1:
  - The circle ( radius directly proportional to fixation) on आख्यान decreases
  - The circle on पुस्तक remains roughly of same radius
  - Short within word regression on नैराश्य
  - टैगोर & बच्चन skipped though frequency less
- In 2:
  - नैराश्य – very big circle
  - Both प्रेमचंद, बच्चन & टैगोर skipped
    - **Parafoveal processing of proper nouns. Reading and comprehension of text takes place simultaneously.**
- Fixation times on frequent words less, sometimes skipped.

# References:

- [1]K. Rayner. Eye movements in reading and information processing:20 years of research. Psychological bulletin, 124:372-422, 1998.
- [2]Matthew S. Starr and Keith Rayner. Eye movements during reading: some current controversies. Trends in Cognitive Science, 5 (2001), pp. 156-163.
- [3]Inhoff, A.W. and Rayner, K. (1986) Parafoveal word processing during eye fixations in reading:effects of word frequency. Percept. Psychophys. 40, 431-439.
- [4]Duffy, S. A., Morris, R. K.,& Rayner, K. (1988). Lexical ambiguity and fixation times in reading. Journal of Memory and Language, 27, 429-446.
- [5]Images from - [http://www.scholarpedia.org/article/Eye\\_movements](http://www.scholarpedia.org/article/Eye_movements)