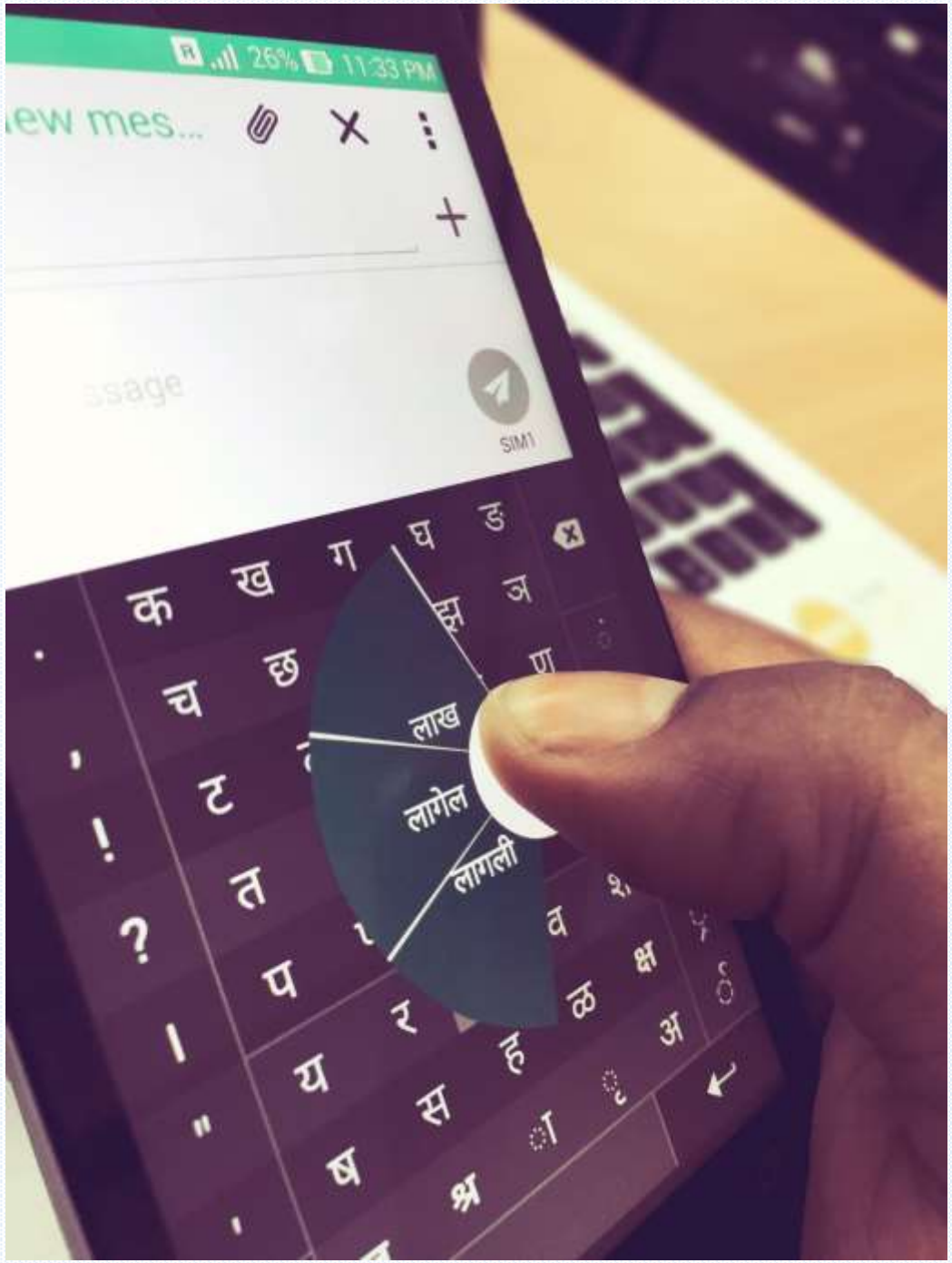


Design of Predictive text input method for Swarachakra

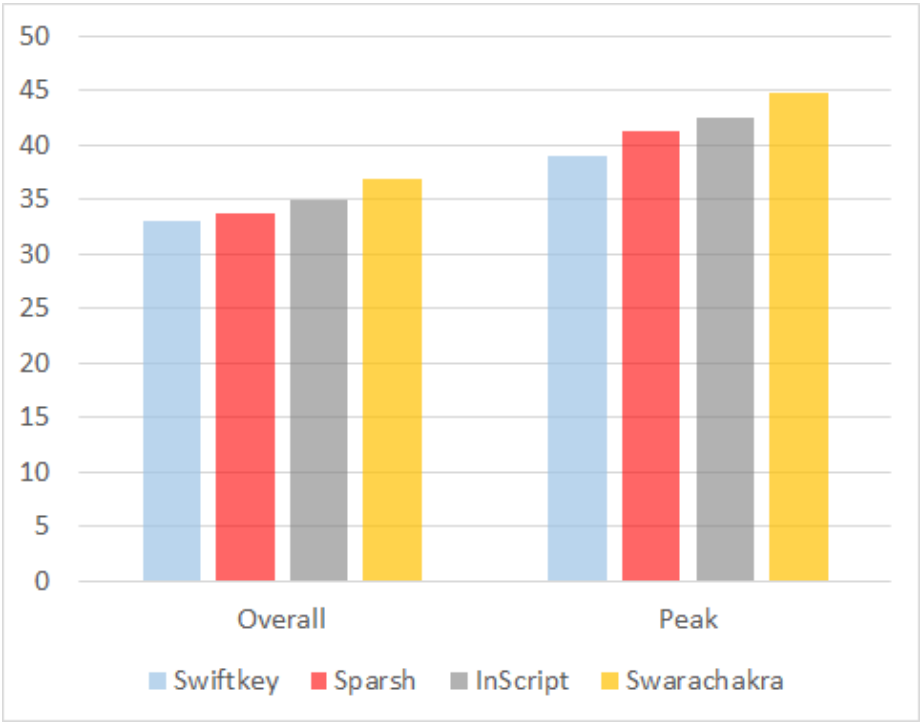
Guided By
Prof. Anirudha Joshi

Prasad Ghone | Interaction Design | 146330010

Indian Language **Predictive keyboards perform worse** than Non-predictive

Keyboard	Model	Taps per Unicode
InScript	Shift	1.10
InScript	Long-press	1.05
Swiftkey	Shift	1.10
Swiftkey	Long-press	1.05
Swiftkey	Long-press + prediction	0.91
Swiftkey	Long-press + Flow	0.68
Swiftkey	Long-press + prediction + Flow	0.67
Sparsh	Without prediction	0.85
Sparsh	Prediction	0.78
Swarachakra	Without prediction	0.84

Summary of the 10 preliminary theoretical effort models



Empirical findings

Indian Language **Predictive keyboards perform worse** than Non-predictive

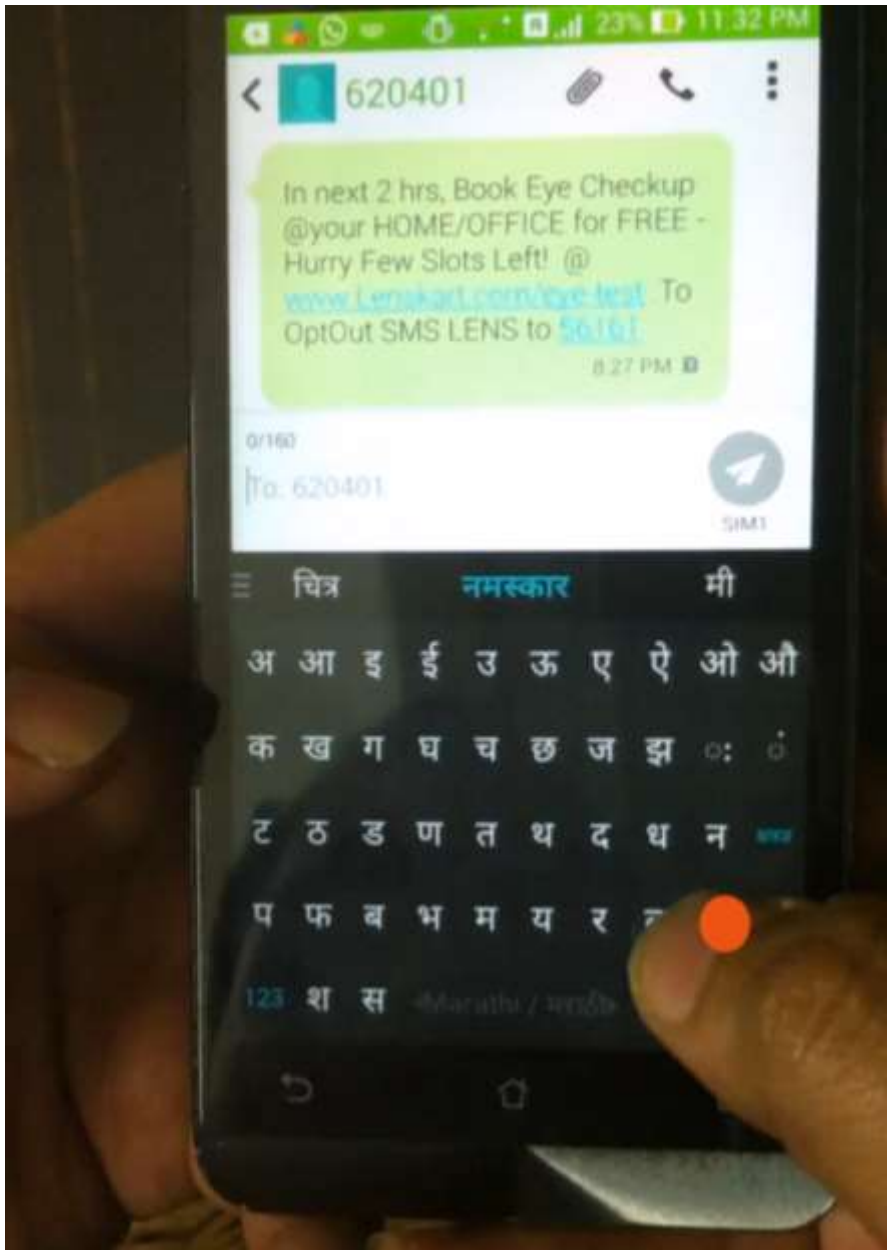
WHY?

Indian Language **Predictive keyboards perform worse** than Non-predictive

- Shift of attention
  - Visual vigilance
  - Visual discontinuity
- Concept model of which words are predicted and which are not.
- Assumption of more words in the shadow
- Does not consider the morphological structure of Indian languages

Indian Language **Predictive keyboards perform worse** than Non-predictive

- **Shift of attention**
  - **Visual vigilance**
  - **Visual discontinuity**
- Concept model of which words are predicted and which are not.
- Assumption of more words in the shadow
- Does not consider the morphological structure of Indian languages



Indian Language **Predictive keyboards perform worse** than Non-predictive

- Shift of attention
  - Visual vigilance
  - Visual discontinuity
- **Concept model of which words are predicted and which are not.**
- Assumption of more words in the shadow
- Does not consider the morphological structure of Indian languages

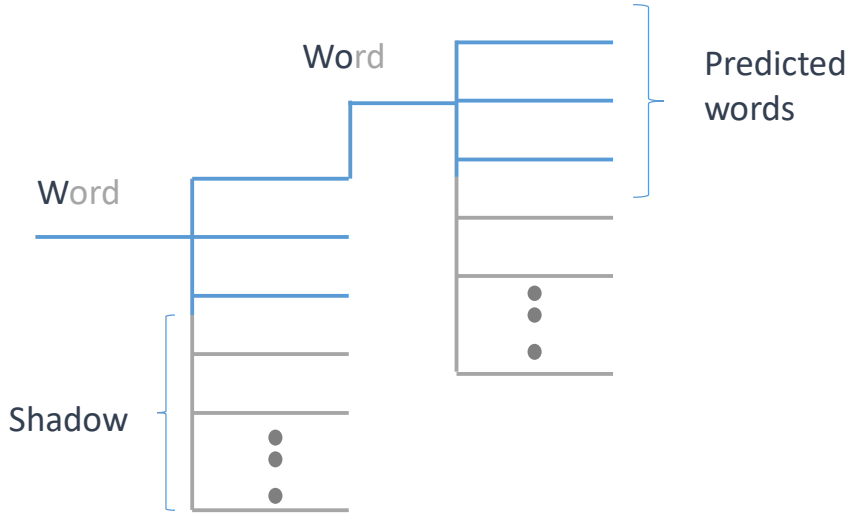
Example

औषध घेतले दुखणे थांबले

जन गण मंगलदायक जय हे

Indian Language **Predictive keyboards perform worse** than Non-predictive

- Shift of attention
  - Visual vigilance
  - Visual discontinuity
- Concept model of which words are predicted and which are not.
- **Assumption of more words in the shadow**
- Does not consider the morphological structure of Indian languages



Prediction Window

आहे | आणि | आता | आज | आहेत | आपण . . .

Indian Language **Predictive keyboards perform worse** than Non-predictive

आ (Typed Word)	Words	Frequency	Relative frequency
Words in prediction window	आहे	5,36,521	0.014956838
	आणि	1,75,624	0.00489595
	आता	1,12,186	0.00312746
Next 3 words in shadow	आज	99,854	0.002783675
	आहेत	69,644	0.001941497
	आपण	49,658	0.001384339
Prediction Window		8,24,331	0.022980248
Complete shadow		15,98,608	0.044565118
Total words in entire corpus		3,58,71,284	
No. of words in shadow		31,694	
Percentage of corpus in prediction window (%)		34.02194607	
Percentage of corpus in shadow (%)		65.97805393	

का (Typed Word)		Frequency	Relative frequency
Words in prediction window	काय	3,34,091	0.009313606
	काही	76,197	0.002124178
	काम	50,204	0.00139956
Next 3 words in shadow	कारण	30,095	0.000838972
	काल	20,090	0.000560058
	काळजी	10,811	0.000301383
Prediction Window		4,60,492	0.012837344
Complete shadow		3,88,418	0.010828104
Total words in entire corpus		3,58,71,284	
No. of words in shadow		1187	
Percentage of corpus in prediction window (%)		54.24509076	
Percentage of corpus in shadow (%)		45.75490924	

Indian Language **Predictive keyboards perform worse** than Non-predictive

- Shift of attention
  - Visual vigilance
  - Visual discontinuity
- Conceptual model of which words are predicted and which are not.
- Higher key hand movements
- **Does not consider the morphological structure of Indian languages**

Morphology of language

- **Agglutinative language**
- घराच्याखाली =

घरा	+ च्या +	खाली
बिल्डिंग	+ च्या +	वर
मंडपा		समोर
झाडा		मागे
		पाठी
		आत
		बरोबर

Why predict complete words?

- Reduce space keystrokes
- Using the same prediction model of English
- Corpus is in that way

What if we don’t predict complete words?

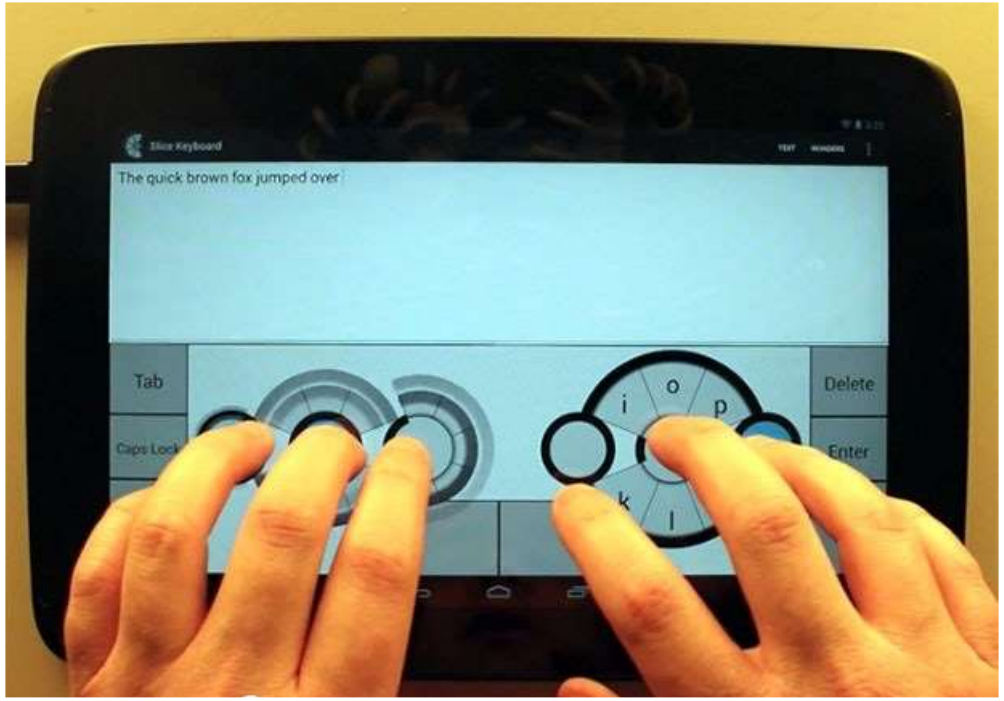
- Extra space keystrokes
- More predictability
- Overall more keystrokes
- Creative language based prediction model and corpus

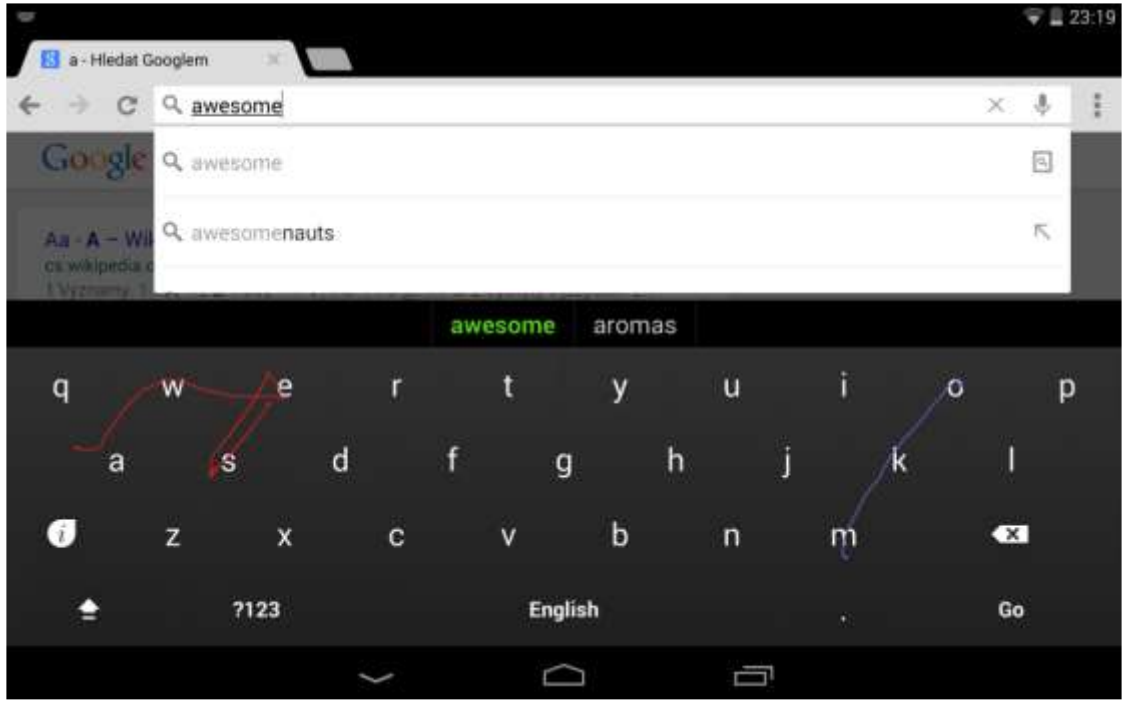
Current prediction mechanisms

Current prediction mechanisms

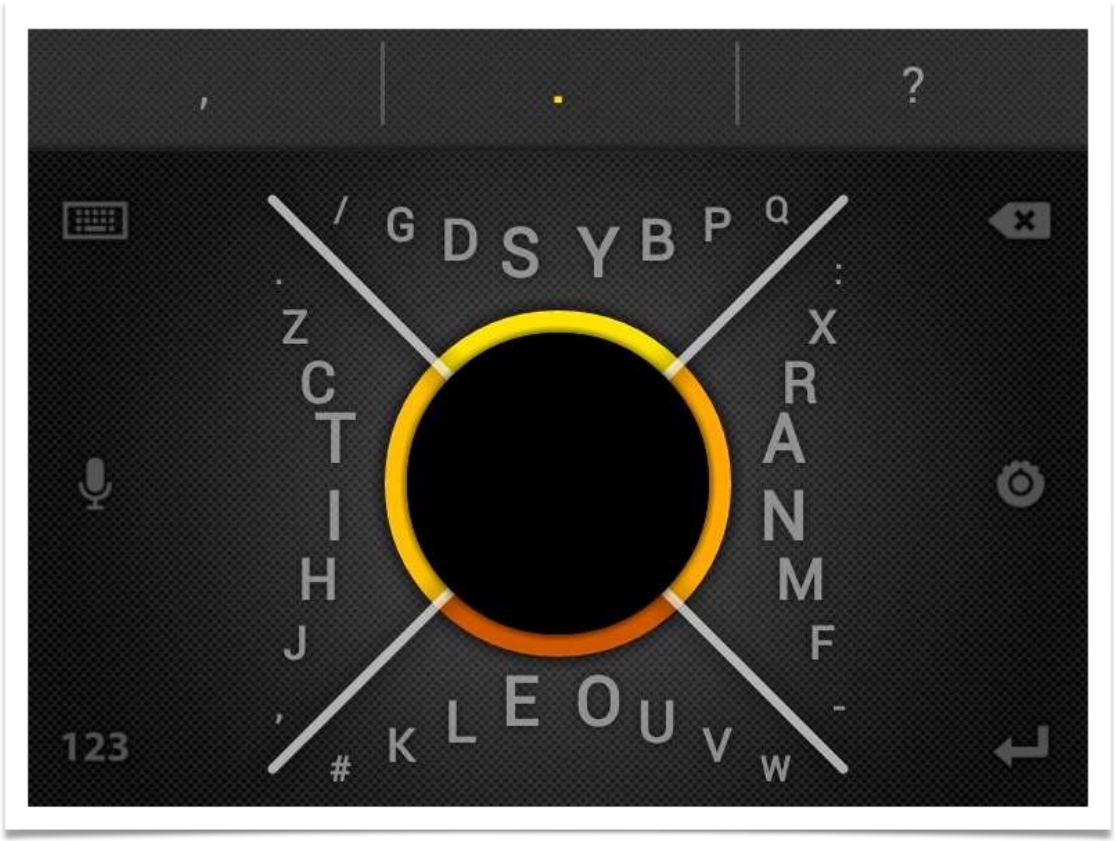
Current prediction mechanisms

Interfaces for fast typing





Interfaces for fast typing



Current prediction keyboards for Marathi

Sparsh

1. Logical keyboard
2. Auto completion
3. Problem of shift of attention
4. Frequency based prediction
5. Comparatively, less cognitive load as single word completion.

Swiftkey

1. Logical keyboard
2. Auto completion and next word prediction
3. Problem of shift of attention
4. N-gram frequency based prediction
5. Non-context based prediction.
   Ex: Rakhi Sawant
6. By swiping, users are held to keyboard and are not dodging between keyboard and prediction window till the word is complete.

User Group

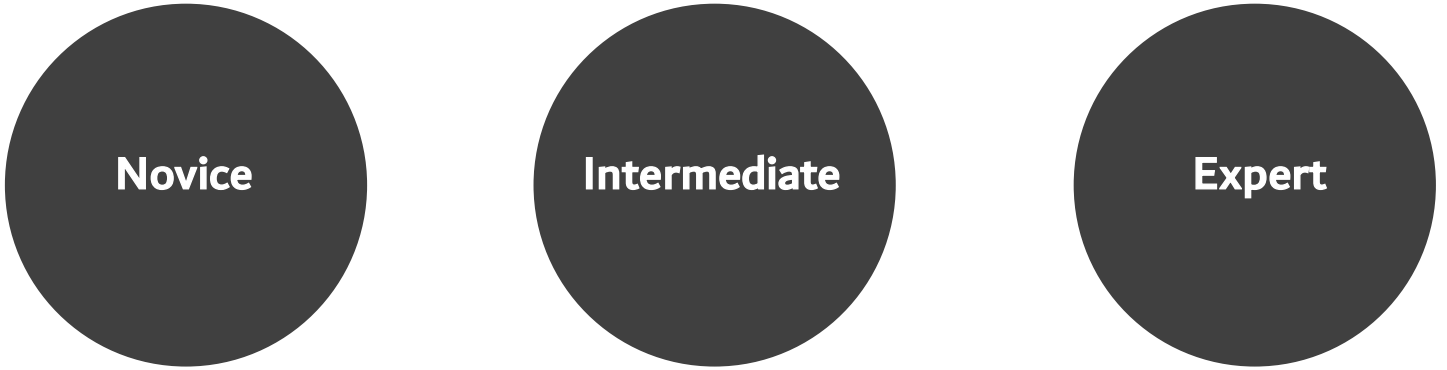
- Hunts and pick characters
- Stuggles with language rules

- Knows location of frequent characters on keyboard
- Knows language rules

- Knows location of frequent characters on keyboard
- Knows language rules
- Have speed in typing
- Doesn’t make typing errors at high speed

User Group

Prediction



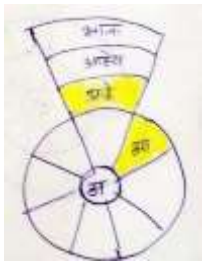
- Struggles with keyboard itself, less open to prediction

- Expected to be active in typing
- How predictive system is predicting? Creating prediction model
- Also, Needs to have keyboard knowledge and understand rules of text input
- Focus on the predictive model rather than message

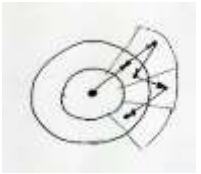
- Shift of attention reduces speed
- Has a conceptual model of the prediction system

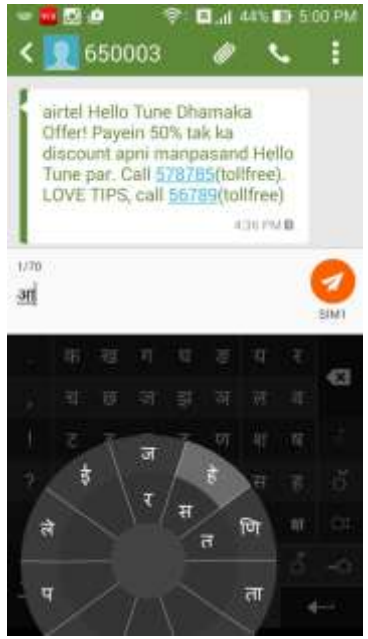
Design Ideas

Bringing the standard prediction window from top of the keyboard to chakra.



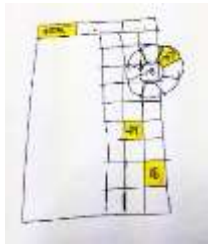
Gesture keyboard

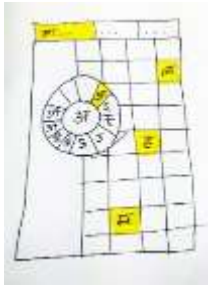




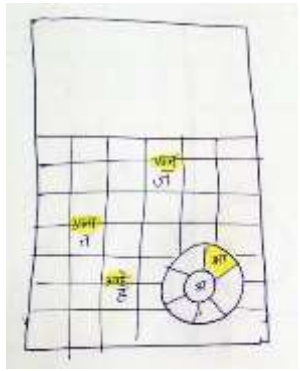
Design Ideas

Next word highlight

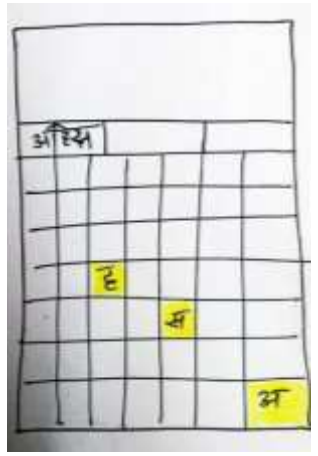




On key prediction



Prediction using multiple single keystrokes



Redefined Project Brief

The project aims to build a predictive interface method for Swarachakra Marathi. A novel interface is created to solve specific problems a user faces. The problems to solve and test for these project are:

1. Shift of attention problem.
2. Longer time taking and still uncertain conceptual model of predictive mechanisms.

Design Concepts

What to predict?

Where to predict?

How much to predict?

Why predict less?

- **Conceptual model** of prediction systems
- Enable them to **fasten the process of conceptual model** making
- Static predictions: Predictions will always appear in the same position in prediction chakra **(Muscle memory)**

Issues

- Non contextual and non-smart prediction
- Personal corpus won’t be included
- Regional variance of languages not accounted

Design Concepts

What to predict?	Where to predict?	How much to predict?
Complete word **घराच्याखाली**	Traditional on top of keyboard 	Entire corpus
N-gram **घरा + च्या + खाली**	Prediction Chakra 	Thresholded corpus

N gram + Traditional + Entire

- Number of keystrokes increases
- More predictability
- Traditional prediction window makes sense as the prediction will happen in sequence
- If the word not predicted completely, no spacebar should be added. Extra keystroke

N gram + Traditional + Less

- The problem of shift of attention may be solved
- Faster conceptual model making
- Having a muscle memory of most frequent words may be possible

N gram + Prediction chakra + Entire

- **Drag and see, drag and see, drag and see**
- **Prediction chakra does not make sense here**

N gram + Prediction chakra + Less

- **Drag and see, drag and see, drag and see**
- **Prediction chakra does not make sense here**
- **Low chances of remembering the position of the prediction predicted last time**

Word + Traditional + Entire

- **Default model**
- **All known problems**

Word + Traditional + Less

- **Shift of attention problem still exists**
- **Having a muscle memory of most frequent words may be possible**
- **Static prediction**

Word + Prediction chakra + Entire

- **Shift of attention may be solved**
- **No help in conceptual model making**
- **No muscle memory of prediction**

Word + Prediction chakra + Less

- **No shift of attention**
- **Helps in conceptual model making**
- **Prediction gestures in muscle memory**
- **Static prediction**

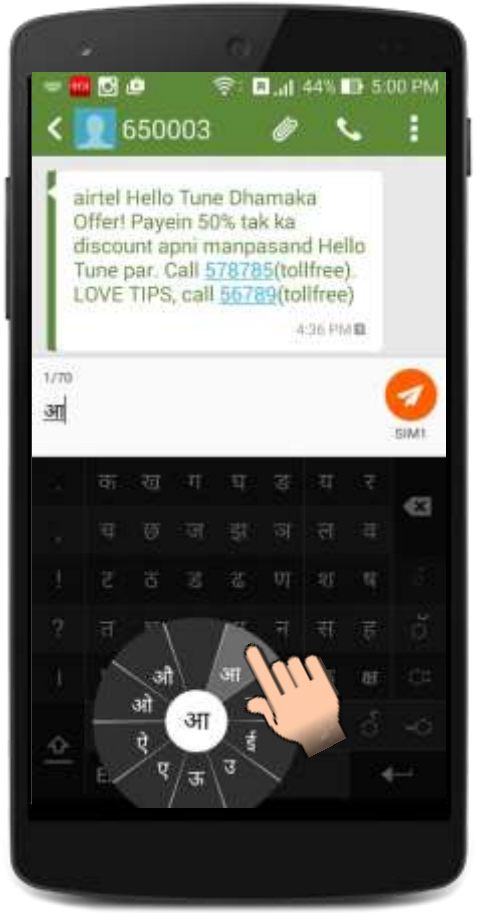
Design Concepts

What to predict?	Where to predict?	How much to predict?
Complete word **घराच्याखाली**	Traditional on top of keyboard 	Entire corpus
N-gram **घरा + च्या + खाली**	Prediction Chakra 	Thresholded corpus

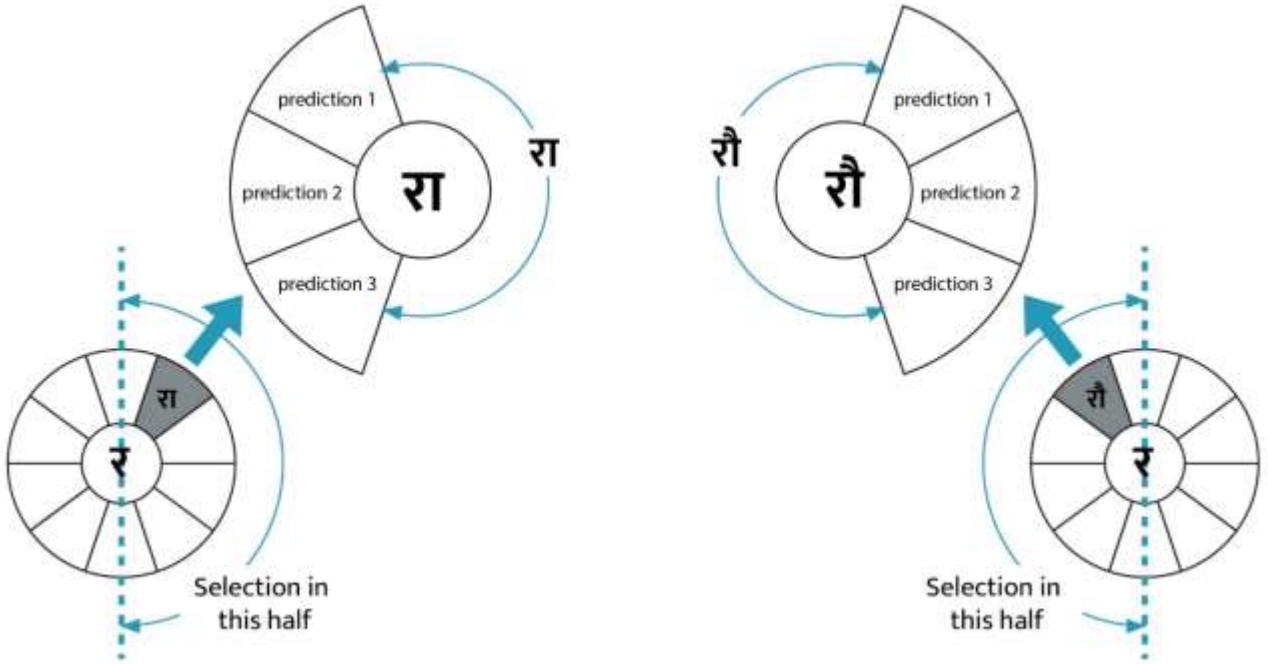
Selected Concepts

What to predict?	Where to predict?	How much to predict?
Complete word **घराच्याखाली**	Traditional on top of keyboard 	Entire corpus
N-gram **घरा + च्या + खाली**	Prediction Chakra 	Thresholded corpus

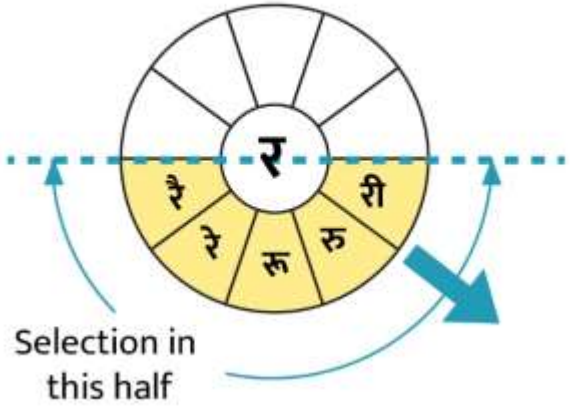
Final Concept

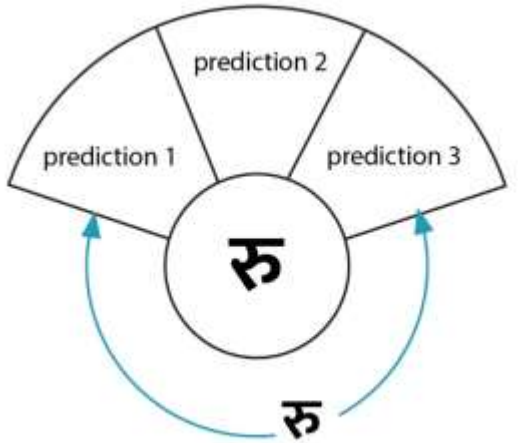


Final Concept

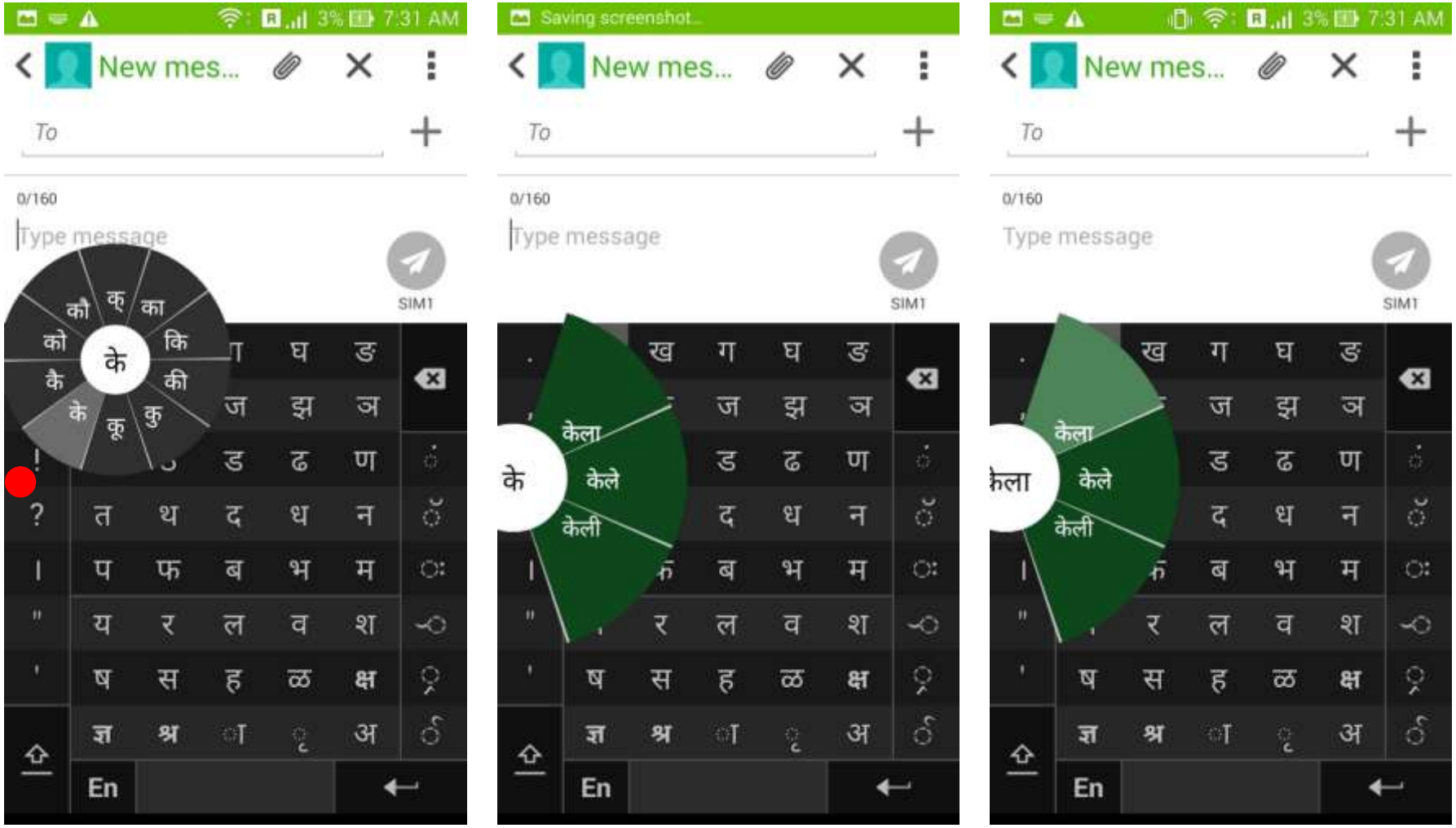


Final Concept





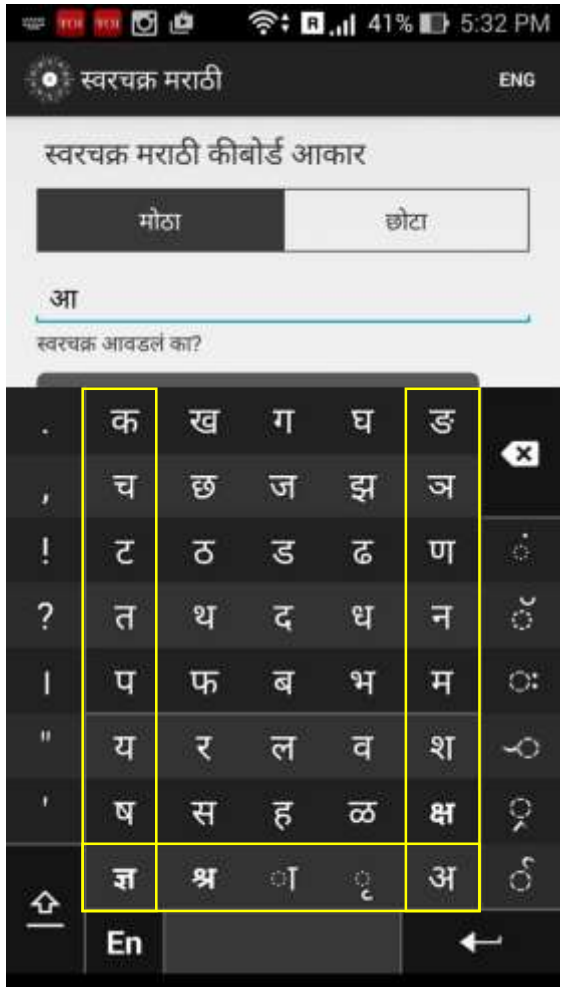
Final Concept

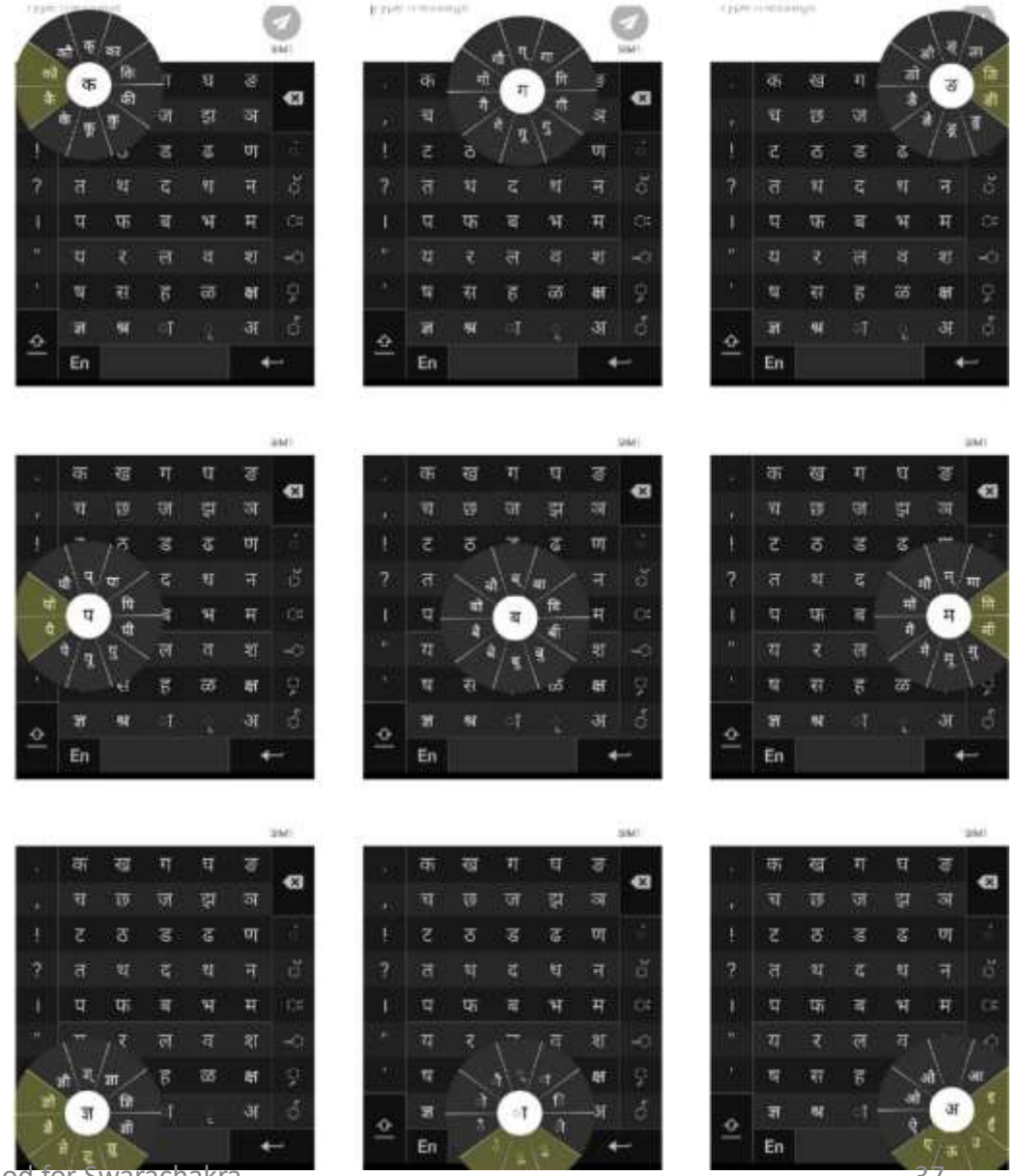


Final Prototype

Limitations

1. Words with no matras (vowel modifiers) won’t be predicted .
2. Problem of Extremities





Limitations

कै	को	ङि	ङी
चै	चो	ञि	ञी
टै	टो	णि	णी
तै	तो	नि	नी
पै	पो	मि	मी
यै	यो	शि	शी
षै	षो	क्षि	क्षी
ज्ञै	ज्ञो	इ	ई

ज्ञे	ज्ञू	ज्ञु
श्रे	श्रू	श्रु
ए	ऊ	उ

- **Covers 8.629 % of the entire corpus**
- **Value less than 8.629%**

  ex: कोकिळा

Evaluation

Experiment

Keyboards

1. Prediction Chakra + Less
2. Prediction Chakra + As Much
3. Swarachakra without prediction

- 6 users per keyboard for all 3 keyboards (within subject)
- 20 phrases per session
- 4 sessions per attempt
- 3 attempts per keyboard (same phrases)

Evaluation

Phrase set 1

झाली सकाळ सरली रात
ससा ससा दिसतो कसा
बरेच ढग दिसत आहेत
मला थोडे पाणी देता का
रमेश जेवण कर
बागेभोवती भिंत आहे
आठवण आहे ना तुला
घरी सगळे कसे आहेत
पावसाच्या रेघांत खेळ खेळू दोघांत
विंध्य हिमाचल यमुना गंगा
ते माझं कौतुक करू लागले
काखेत कळसा गावाला वळसा
एका छान अनुभवाला तो मुकला होता
नाच रे मोरा आंब्याच्या वनात
धुरांच्या रेघा हवेत काढी
डेबू मामाच्या शेतावर कष्ट करू लागला
पाखरे घरट्यांत जाऊन बसली आहेत
कुलस्त्री जसे हास्य ओठात शोभे
स्वतः मेल्याशिवाय स्वर्ग दिसत नाही
स्वराज्य हा माझा जन्मसिद्ध हक्क आहे

Phrase set 2

चिमणी करते चिव चिव
किती वेळ लागेल
जोवरी पैसा तोवरी बैसा
तू कसा आहेस
राव चढले पंत पडले
काळाकाळा कापूस पिंजला रे
भारतभाग्यविधाता
मामाची बायको सुगरण रोजरोज पोळी-शिकरण
नदीनाल्यांना आला पूर
कामापुरता मामा आणि ताकापुरती आजी
झोंबे अंगा वारे काया थरथरे
पळस गेलं कोकणात तीन पानं चुकेनात
खायला कोंडा नि निजेला धोंडा
पावसाच्या रेघांत खेळ खेळू दोघांत
सारे भारतीय माझे बांधव आहेत
पुढे मला काही कल्पना सुचू लागल्या
परहित आधी नंतर स्वहित साधावे
कराग्रे वसते लक्ष्मी करमध्ये सरस्वती
दैव देते आणि कर्म नेते
पचापचा शिव्या देई खाताखाता पान

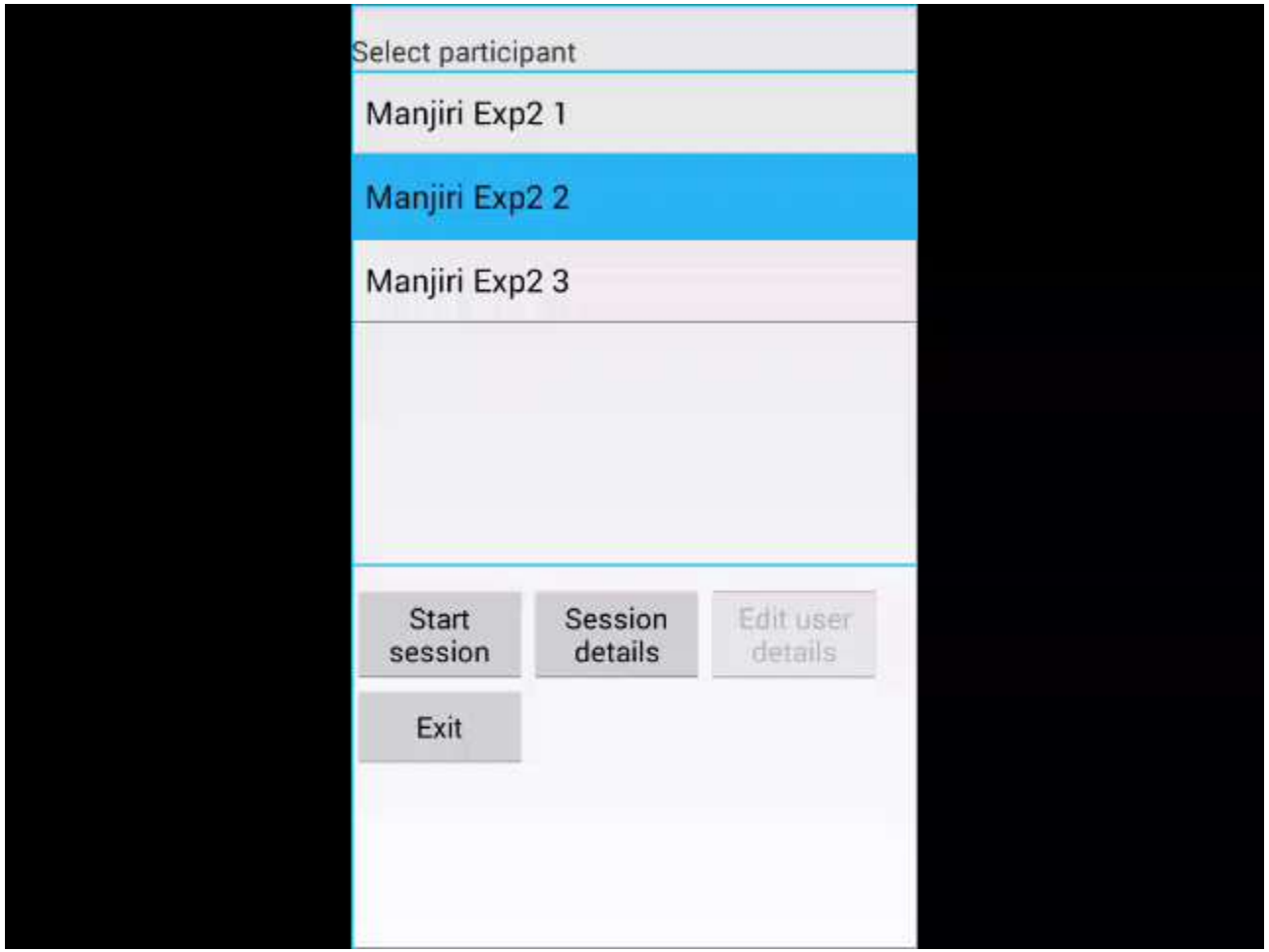
Phrase set 3

हवेत उडतो लाल लाल फुगा
तू कशी आहेस
घरी कशी मग सांगा जातिल
माकडाने रंगवले आपले तोंड
आपण कुठून आलात
टपटप पानांत वाजती रे
हात लावता पंख फाटतिल
झोंबे अंगा वारे काया थरथरे
मीना गोष्ट वाचत होती
दोघांनीही आपापले पैसे मोजले
आभाळात छानछान सातरंगी कमान
कुठूनही गेले तरी पोरांची नजर पडणारच
तुझंमाझं जमेना तुझ्यावाचून करमेना
सरड्याची धाव कुंपणापर्यंत
पाखरे घरट्यांत जाऊन बसली आहेत
सुधेसारखा साद स्वर्गीय गाणे
स्वतः मेल्याशिवाय स्वर्ग दिसत नाही
चार आण्याची कोंबडी अन बारा आण्याचा मसाला
दुपारी चारच्या सुमारास पाऊस सुरू झाला

Evaluation

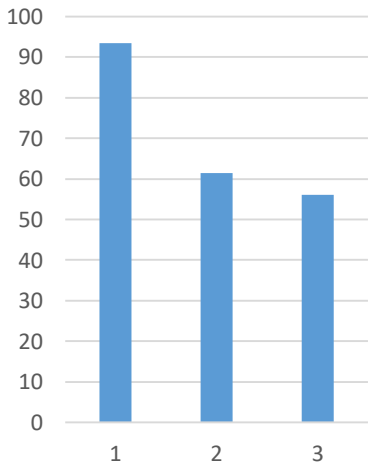
VKB tool was used for evaluation

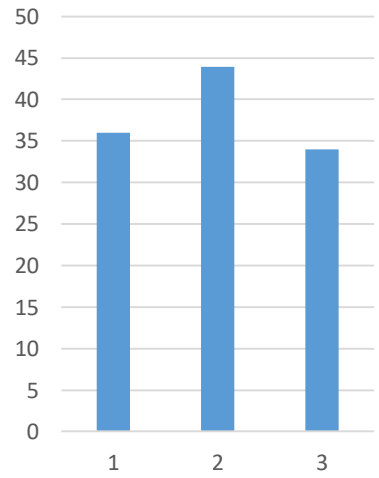
It calculates CPM, Accuracy, Edit distance, keystroke logs, time stamp, etc

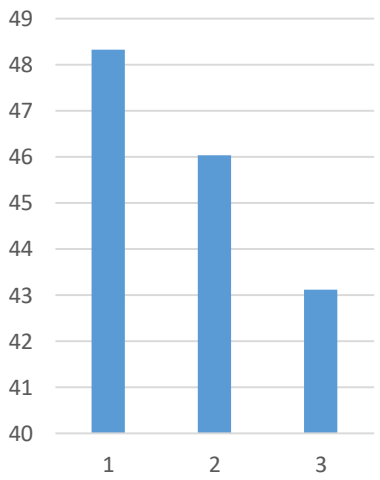


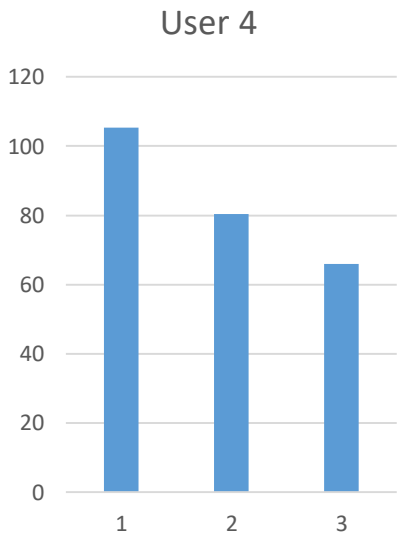
[Truth Table]

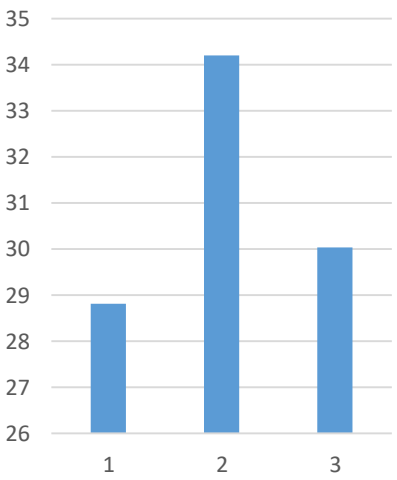
Descriptive Results









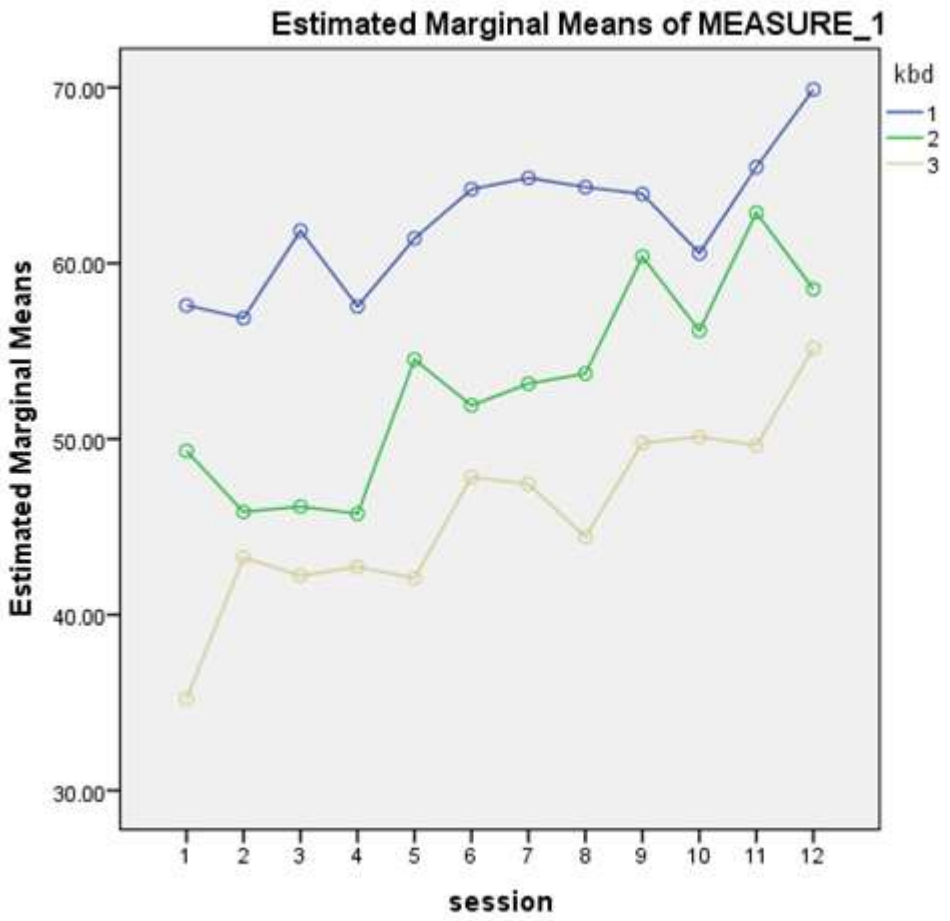


1 – Swarachakra without prediction

2 – Swarachakra with less prediction

3 – Swarachakra with entire corpus prediction

Results



1 – Swarachakra without prediction

2 – Swarachakra with less prediction

3 – Swarachakra with entire corpus prediction

Conclusions

- Swarachakra without prediction yields best results among all keyboards
- In prediction, Swarachakra with less corpus performs better than Swarachakra with entire corpus
- Bugs in the implementation can be a reason for poor performance of predictive keyboard
- Expert behaviour had an effect on results

Acknowledgement

I would like to express my sincere gratitude to Prof. Anirudha Joshi for his support and guidance throughout.
Thanks to Prof. Venkatesh Rajamanickam, Prof. Girish Dalvi, and Prof. Ravi Poovaiah for their valuable inputs during the course of the project.
A special thanks to Dileep Mohanan and Jayati Bandyopadhyay for taking me out of all the android development hurdles all the time. Thanks to Akshay Kore, Indrajeet Roy and Sagar Yende for all the motivation and support.
Thanks to the Manjiri and Shashank from Swarachakra Team, IDC for all the resources and support.
Most importantly, my family for all their love, patience and encouragement.

Thank you!