NOKIA N900



Improving Literacy in Developing Countries Using Speech Games on Mobile Devices

Anuj Kumar¹, Pooja Reddy², Anuj Tewari³, Rajat Agrawal¹, Matthew Kam¹

- ¹ Carnegie Mellon University, USA
- ² American Institutes for Research, USA
- ³ University of California, Berkeley, USA

Q	w ²	E	R	T 5	Υ	U	1 "	0 9	P	1
A	S	D	F	G	H	J	K	L	.,	4
+	Z	X	C	V 5	В	N	М	<u></u>		Ctrl

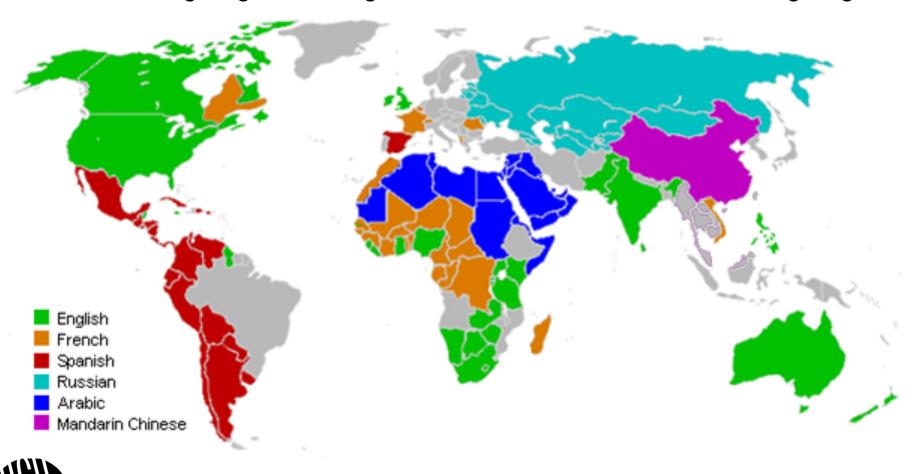




Over 50% of world is multilingual

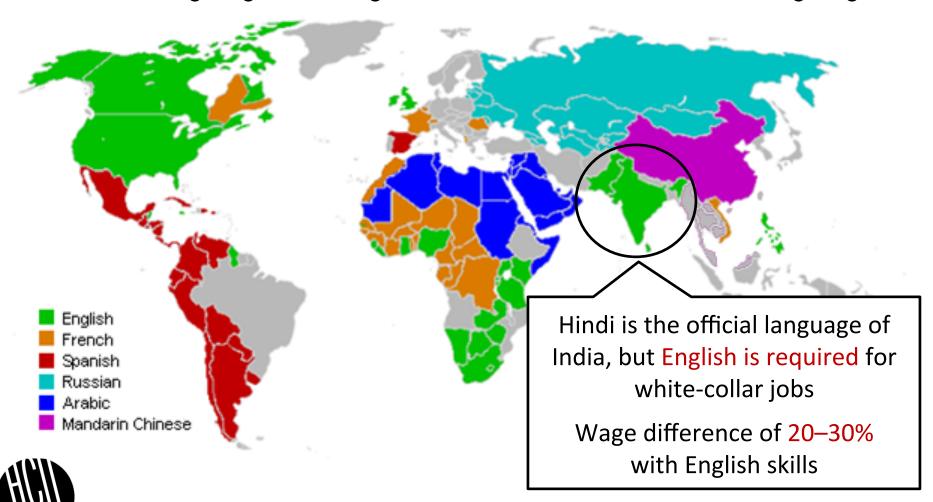
But... not all languages share equal status

"Power" languages emerge – different from official language



But... not all languages share equal status

"Power" languages emerge – different from official language



Challenges with 2nd Language Instruction

Complex host of problems in developing regions







Low teacher attendance

Children need to work; Poor quality of don't attend school instruction in series language

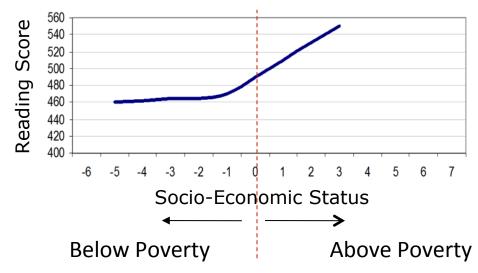
Poor quality of instruction in second language (i.e. English)



Over 80% of grade 5 children cannot read & understand simple grade 1 English sentences (ASER, 2009)

Word Reading

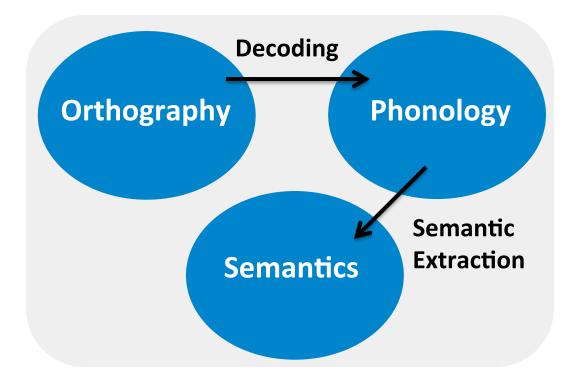
- Ability to read and understand a written word
- Need to understand 98% of words in a text to comprehend
 it (Hu & Nation, 2000; Carver, 1994)
- Word reading is a major bottleneck for L2 learners at grades 4-5 (Carlo et al., 2004; August et al., 2005)
 - Problem amplified for low-income learners:
 90% of poor children read two grade levels lower than children above poverty line (Willms, 2004)





Word Reading: Theoretical Framework

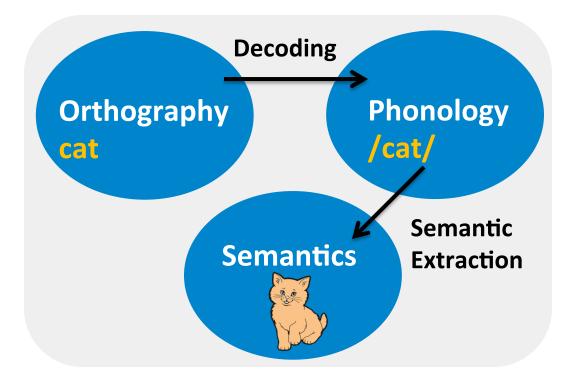
- Depends on quality of two sub-skills (Perfetti, 2001):
 - Decoding: maps written letters to its phonology i.e. sound
 - Semantic Extraction: maps sound to semantics i.e. meaning





Word Reading: Theoretical Framework

- Depends on quality of two sub-skills (Perfetti, 2001):
 - Decoding: maps written letters to its phonology i.e. sound
 - Semantic Extraction: maps sound to semantics i.e. meaning





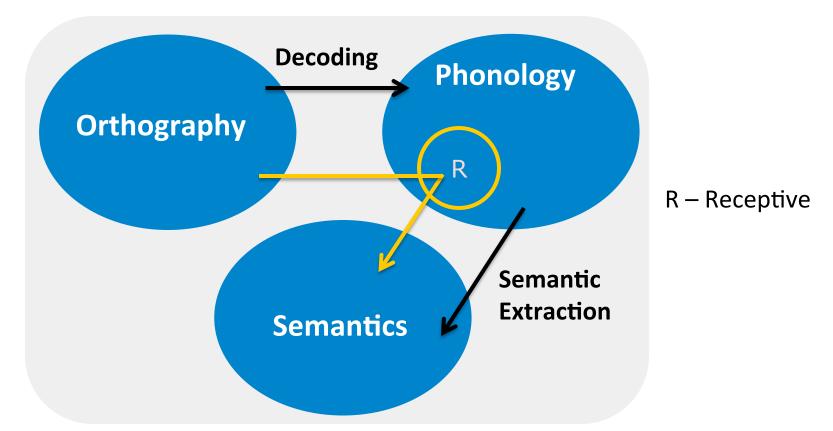
How can we improve word reading sub-skills?

- State-of-the-art vocabulary learning software is mostly receptive i.e. "recognizing what written words mean"
- Production i.e. "saying words aloud" is arguably more beneficial:
 - Draws attention to one's linguistic abilities, which generates new knowledge or consolidates existing knowledge (Swain & Lapkin, 1995; Ellis & He, 1999)
 - Self-generated input back to your mind strengthens word representation (De Bot, 1996)



First Hypothesis

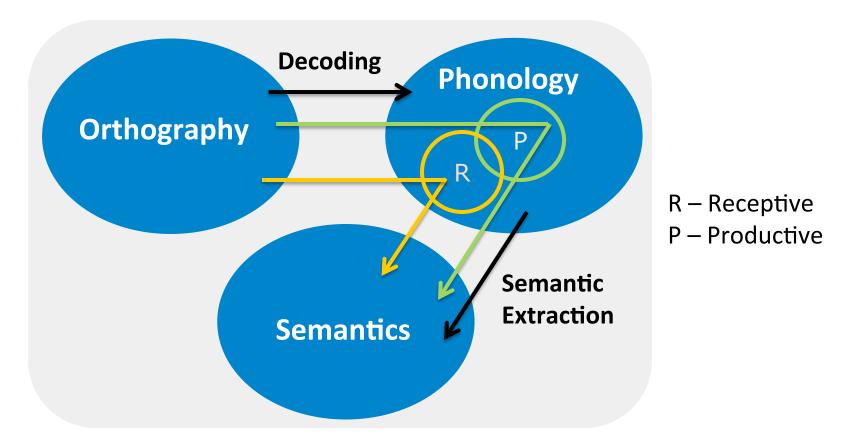
 H1: Productive training improves word reading more than receptive training





First Hypothesis

 H1: Productive training improves word reading more than receptive training





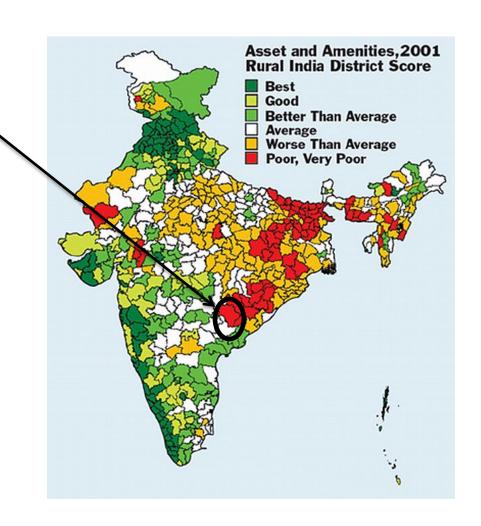
Study Context

Public schools in rural India

Economic baseline

\$200 to \$1,500 annual household income

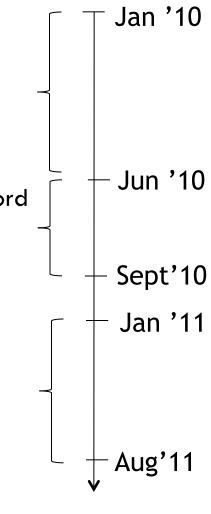
 Most of our participating households had at least one cellphone





Study Timeline

- Phase 1: System Development (21 weeks)
 - Games: Design, development & usability evaluation
 - Speech Recognizer: Context-based adaptations
- Phase 2: Experiment 1 (13 weeks)
 - H1: Understand the role of productive practice towards word reading (5 weeks)
 - Data Analysis (8 weeks)
- Phase 3: Experiment 2 (32 weeks)
 - Games: Redesign and development (16 weeks)
 - H2-H4: Understand the role of hints in enabling productive practice (6 weeks)
 - Data Analysis (10 weeks)



Study Timeline

- Phase 1: System Development (21 weeks)
 - Games: Design, development & usability evaluation
 - Speech Recognizer: Context-based adaptations
- Phase 2: Experiment 1 (13 weeks)
 - H1: Understand the role of productive practice towards word reading (5 weeks)
 - Data Analysis (8 weeks)
- Phase 3: Experiment 2 (32 weeks)
 - Games: Redesign and development (16 weeks)
 - H2-H4: Understand the role of hints in enabling productive practice (6 weeks)
 - Data Analysis (10 weeks)





Game Designs

- Two games, each covered 15 new words
 - Curriculum designed by an expert, local ESL teacher
 - Both games comprised of three phases



Game Designs: Teaching Phase

Introduced 5 new words at a time



Teaching of "Brinjal" in native language (i.e. Telegu)

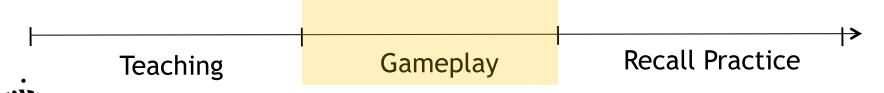
Teaching Gameplay Recall Practice

Game Designs: Gameplay

 Designs based on common actions from village games, e.g. catching or evading a player (Kam et al. '09)

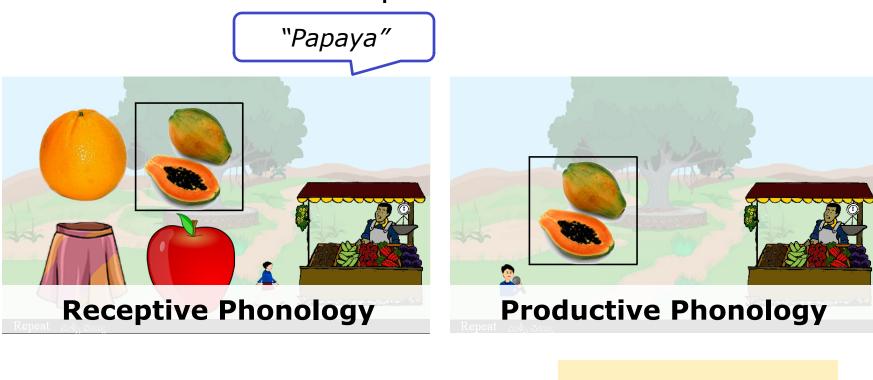






Game Designs: Practice Phase

Practice based on the experimental condition





Study Timeline

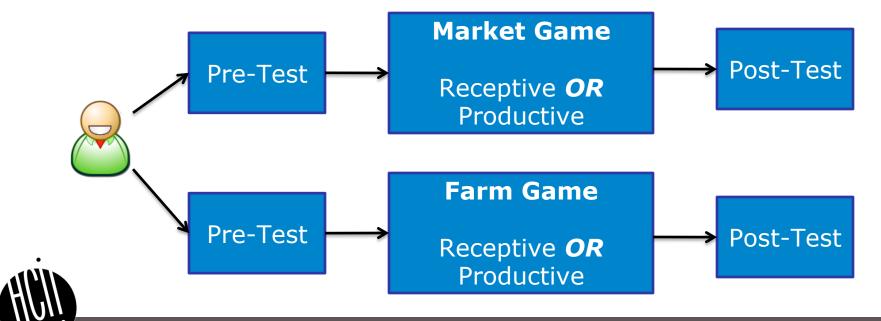
- Phase 1: System Development (21 weeks)
 - Games: Design, development & usability evaluation
 - Speech Recognizer: Context-based adaptations
- Phase 2: Experiment 1 (13 weeks)
 - H1: Understand the role of productive practice towards word reading (5 weeks)
 - Data Analysis (8 weeks)
- Phase 3: Experiment 2 (32 weeks)
 - Games: Redesign and development (16 weeks)
 - H2-H4: Understand the role of hints in enabling productive practice (6 weeks)
 - Data Analysis (10 weeks)





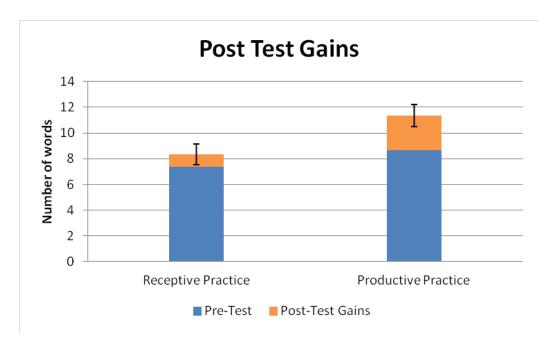
Experimental Design

- 1 hour sessions
- Between-subjects experiment w/ 21 participants
- Each participant played both games in separate sessions,
 but randomized for the game condition



Result: Productive practice outperforms

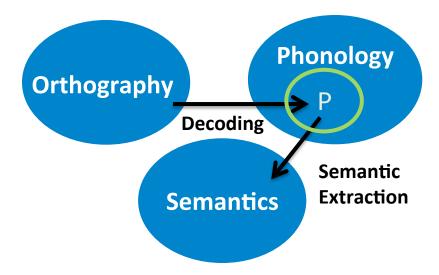
- Verbalizing the word aloud (production) led to better word reading scores than just silently voicing it in head (receptive)
 - Supports H1: F(1,40)=5.4, p=0.001*
 - Short-term gain of 3 new words after 30 min. of game play





Subsequent Questions

- Productive tasks are cognitively intensive (Nation, 2001)
 - At least 12 (out of 21) participants appeared to take greater time
 to recall and vocalize the words aloud
 - What instructional supports can facilitate productive recall?





Study Timeline

- Phase 1: System Development (21 weeks)
 - Games: Design, development & usability evaluation
 - Speech Recognizer: Context-based adaptations
- Phase 2: Experiment 1 (13 weeks)
 - H1: Understand the role of productive practice towards word reading (5 weeks)
 - Data Analysis (8 weeks)
- Phase 3: Experiment 2 (32 weeks)
 - Games: Redesign and development (16 weeks)
 - H2-H4: Understand the role of hints in enabling productive practice (6 weeks)
 - Data Analysis (10 weeks)



Two Different Hints

Orthographic Hint:

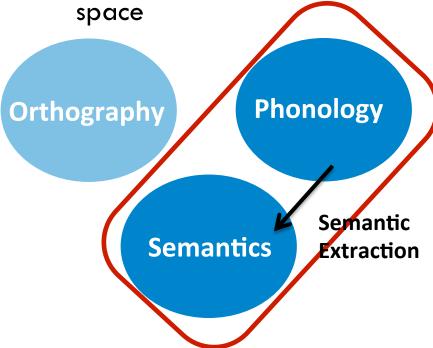
- First alphabet + word length
- Supports decoding by helping cognitive search using letterto-sound rules

Orthography Phonology

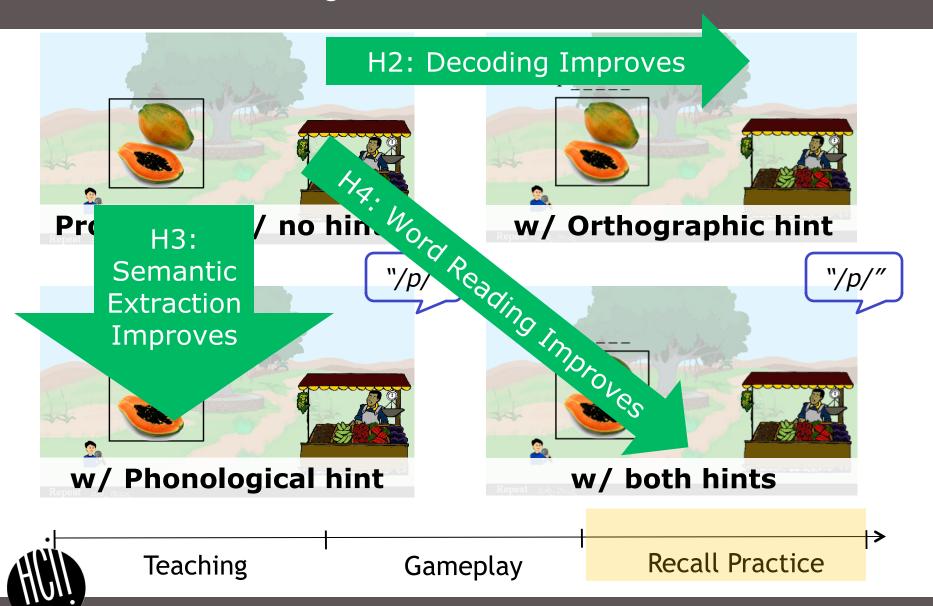
Semantics

Phonological Hint:

- First phoneme of word
- Supports semantic extraction
 by reducing semantic search
 space



Game Redesigns for Practice Phase



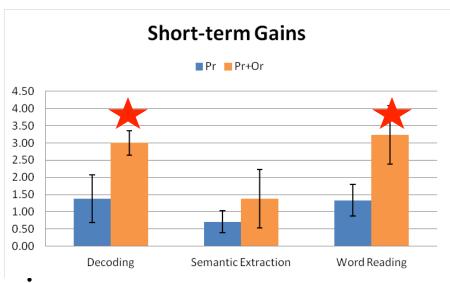
Experimental Setup

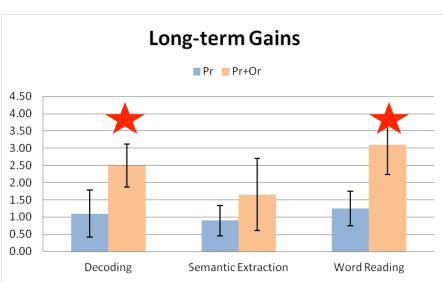
- In total, 40 participants (13 girls) took part in this study
- Similar setup as in experiment 1, except:
 - Learning outcome measured for decoding & semantic extraction, in addition to, word reading
 - Delayed post-test after 7 days to assess long-term gains



H2 Results: Orthographic hint only

- Led to decoding gains, both short- and long-term (but not semantic extraction)
- Eventually, led to word reading gains, both short- and long-term

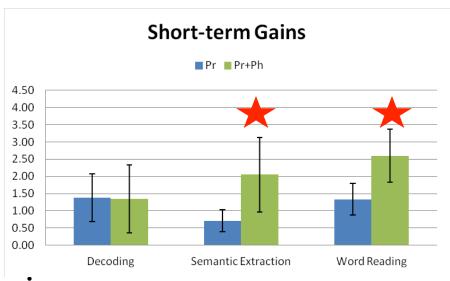


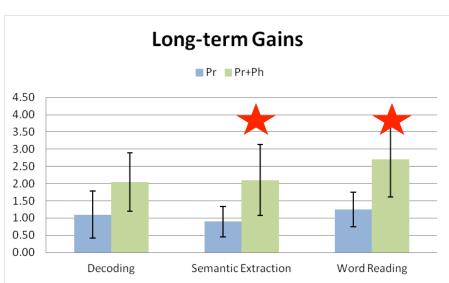




H3 Results: Phonological hint only

- Led to semantic extraction gains, both short- and long-term (but not decoding)
- Eventually, led to word reading gains, both short-term and long-term

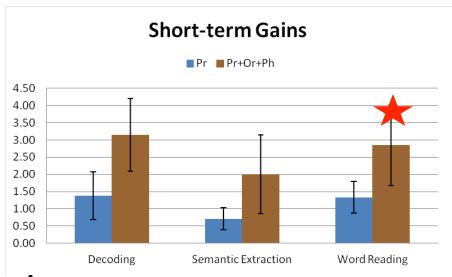


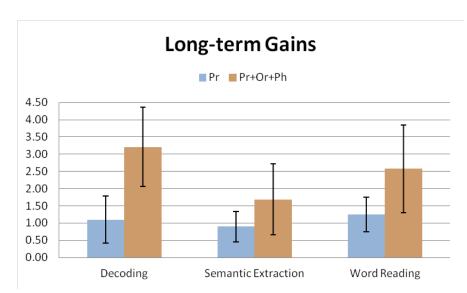




H4: Both Orthographic and Phonological hints

- Led to word reading gains, but only for short-term
 - Possibly because both hints made the recall task too easy







Conclusion

- Productive practice can significantly improve word reading, and potentially, text comprehension
- Productive vocabulary training can support specific sub-systems of word reading
 - Decoding sub-system can be strengthened by providing an orthographic hint
 - Semantic extraction sub-system can be strengthened by providing a phonological hint
 - Both hints together don't necessarily add up to a long-term improvement on word reading

Implications for Design

- Increases reach to low-income learners
- Explore design of non-visual interfaces for learning;
 cheaper than visual interfaces
 - Phonological hint & productive practice could be delivered over voice modality
 - Semantic extraction is harder to master than decoding skills in L2 learners (National Literacy Panel, 2006)





NOKIA N900

Thank You! Questions? ©

Acknowledgements:

Technical Assistance: Venkatesh Keri, Kaushik Shelat, Yash Soni, Gaurav Arora & Indrani Vedula

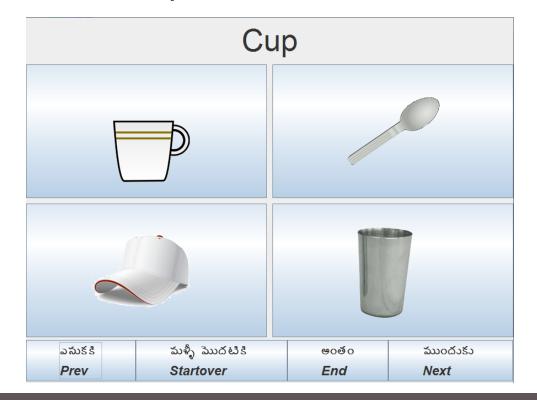
Funders: Human-Computer Interaction Institute at CMU, Nokia Research





Scoring Procedure: Word Reading

- Word Recognition test based on Peabody Picture Vocabulary Test (PPVT)
- +1 for each correct response, 0 for incorrect





Usability Improvement

 None of the typical icons used in speech applications that prompt the user to speak aloud seemed intuitive







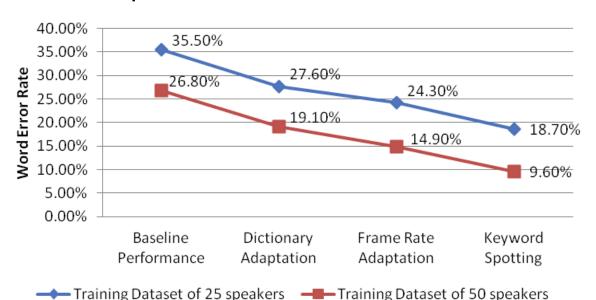
- Most intuitive icon: A boy holding a microphone
 - Designed after local elections where the minister would speak on the microphone





Speech Recognition Improvements

- Training set: 6250 utterances from 50 rural children
- Final misrecognition rate of 9.6% i.e. 1 out 10 times
 - Cases of participant speaking correctly, but system misrecognizing: the participant learned to repeat the word rather than switch the response.





Implications for Design

- Explore design of voice-command games for literacy
 - Many popular voice-command games (e.g. Nintendogs and Brain Age), but focus purely on entertainment
 - First study to explore the voice-command games for literacy



Related Work and Significance

- Several speech recognition systems that involve reading aloud have shown empirical benefits (Mostow, '94; Nix, '98; Williams, '00)
 - LISTEN project targets oral output at supra-lexical level
- Phonological representation at lexical level is important for long-term retention (Nation, 2001)
- More scalable solution as it is easier to build into applications

