

SMART: Speech-enabled Mobile Assisted Reading Technology for word comprehension

Anuj Kumar^{1,1}, Pooja Reddy¹, Matthew Kam¹

¹ Carnegie Mellon University, Pittsburgh, PA, USA

Abstract. In this study, we designed and developed two educational games on mobile phones with support for speech-recognition to examine and train the cognitive underpinnings of word reading in English as a Second Language (ESL) learners in rural India. Specifically, we tested the hypothesis that articulating a word aloud will be more advantageous for strengthening the sub-lexical components required for word reading – orthographic, phonologic, and semantic – than silently practicing it. 31 children from grades 4 and 5 learning ESL in rural India participated in the study. The results corroborated the hypothesis, suggesting that production is important for second language word reading development.

Keywords. Word Reading, Mobile Learning, Speech Recognition, Literacy

1 Introduction

Reading and understanding written words is fundamental for literacy development. Yet it is one of the most challenging skills to acquire, especially in a second language, because it requires the integration of visual, sound, and meaning information [5], all in a new language. In order to address this need, we examined the possible benefits of practicing producing words aloud, rather than simply reading them receptively in one’s mind, for ESL word reading development in rural India students of grades 4-5.

Word reading – the ability to decode (sound out) and understand written real words – is a multi-faceted construct that depends on the quality of representation of three linguistic and cognitive sub-systems: orthographic (visual script), phonologic (sound), and semantic (meaning), according to the Lexical Quality Hypothesis (LQH) [5]. When a written word is encountered, each orthographic unit must be connected to its appropriate phonological unit, allowing a learner to assign sound information to a word and thus *decode* it accurately. For instance, the letter “c” must be mapped with the sound /k/, “a” with /a/, and “t” with /t/ and so forth. A connection must also be made between this phonological representation (the sound /cat/) and its appropriate meaning (small, furry animal), and thus *semantic extraction* must also occur. If the quality of any of these sub-systems is compromised, then word reading will be hampered, and thus, reading comprehension will be impaired [4].

¹ anujk1@cs.cmu.edu

At the same time, it has been argued that oral production is critical for new learners of a language as an oral output provides specific input back to the mind, which in turn assists a learner to transition from declarative knowledge (ability to declare that you know a word) to productive knowledge (ability to fluently use the word) [1].

Integrating these strands of research, our conceptual framework is that there are three – orthographic, phonologic, and semantic – representations that are required for word reading, and the connections between them can be processed productively or receptively. Based on this framework, we expect that saying the name of a picture out loud (productive processing) will reinforce semantic extraction more strongly than if the link is receptively processed, i.e. matching a picture and its name [1, 6]. Thus, our experiment consisted of two training conditions: 1) Receptive (Re); and 2) Productive (Pr), and we hypothesized that: (H1) Productive training will be more beneficial for word reading than Receptive training.

2 System and Game Design

Due to prior success of using games for education [2, 3], our intervention consisted of two English literacy learning games: Market Game, and Farm Game. These were prototyped using ActionScript 3.0 for Nokia N810. Our game designs drew on our experiences from prior field studies of traditional Indian games that rural children enjoyed most, including physical tag-like games that have actions such as *catching* or *evading* a player [3]. The speech recognizer was fine-tuned for the accent, noise and speaking rate of the participants, and the final recognition accuracy was 91%.

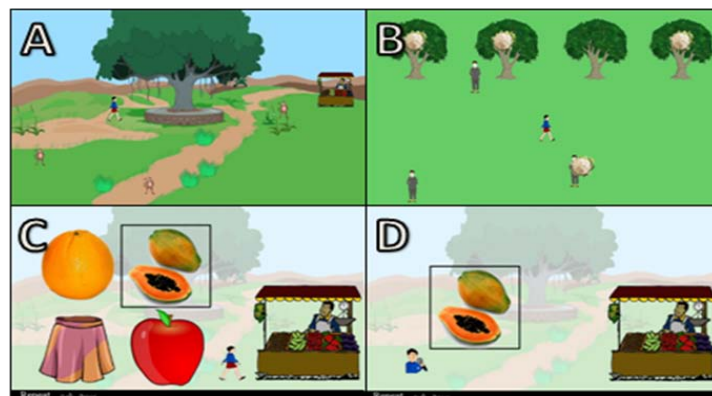


Figure 1. (A) shows the boy moving towards the shop in the market game, while the monkeys attempt to catch him; (B) shows the boy catching the thief in the farm game; (C) Re training condition in market game; (D) Pr training condition in market game.

In the Market Game (Figure 1A), the boy character had to travel from home to the market to buy items while avoiding monkeys en-route. At the market, depending on the experimental condition that the user had been assigned to, she/he purchased items by either *selecting* the correct item that corresponded to the said word (Figure 1C, Re

condition), or by *saying the word aloud* that corresponded to the image displayed (Figure 1D, Pr condition). Similarly, in the Farm Game (Figure 1B), the objective was to save a farm by catching all the thieves and retrieve the items that they had stolen in one of the two ways described above for the Market Game.

3 Experiment and Early Results

31 participants (18 boys) participated in this study. They were 9-13 year olds ($M=10.5$) and were in grades 4-5. All participants were attending a public school in a rural part of South India. The experimental design was a pre-post test block design, and the intervention comprised of the above described games. The outcome variable, word reading, was tested before and after each game. 25 words were selected from grade 4-5 level government-issued textbooks for the intervention. Each child played both games, but was randomly assigned to one of the two experimental conditions. The findings from this study demonstrated that with even 30 minutes of targeted practice of words and their meanings, it is possible to increase word reading scores, regardless of whether it is productive or receptive; however, as predicted, we found that productive training is significantly more beneficial for word reading than receptive training, $t = -3.01, p < .05$. This is in line with SLA theories, which stress that output of linguistic forms consolidates knowledge [1, 6]. In this case, when a learner is forced to make a link between a word's meaning and its pronunciation productively, the bridge between a word's meaning and its name (sound) are more strongly reinforced, making it more deeply embedded in memory.

Acknowledgments. We thank Anuj Tewari, Venkatesh Keri, Rajat Agarwal and all others who helped in developing and deploying the software.

References

1. De Bot, K. The psycholinguistics of the Output Hypothesis. *Language Learning*, 46, 529-555, 1996.
2. Gee, J.P. *What Video Games Have to Teach Us About Learning and Literacy*. P. Macmillan, 2004.
3. Kam, M., Mathur, A., Kumar, A., and Canny, J. Designing Digital Games for Rural Children: A Study of Traditional Village Games in India. In Proc. of *ACM Conference on Human Factors in Computing Systems (CHI '09)*, Boston, Massachusetts, April 4-9, 2009.
4. Perfetti, C. A. Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11, 357-383, 2007.
5. Perfetti, C. A., & Hart. L. The lexical quality hypothesis. In L. Verhoeven, C. Elbro & P. Reitsma (Eds.), *Precursors of Functional Literacy* (Vol. 11, pp. 67-86). Amsterdam: John Benjamins, 2001.
6. Swain, M., & Lapkin, S. Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16, 371-391, 1995.