Figure 1: Vocabulary drill game based on Tetris
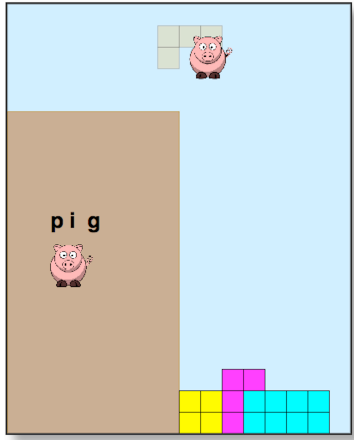
Adapting Arcade Games for Learning

Carrie J. Cai
MIT CSAIL
32 Vassar St.
Cambridge, MA 02139
cjcai@mit.edu

Abstract

Although memory exercises and arcade-style games are alike in their repetitive nature, memorization tasks like vocabulary drills tend to be mundane and tedious while arcade-style games are popular, intense and broadly addictive. This suggests an opportunity to modify well-known, existing arcade games for the purpose of memorization and learning. A key challenge to modifying any existing arcade-style game is the incorporation of learning on top of an already fast-paced, mentally demanding game. This paper presents a web-based vocabulary-drill game, based on Tetris and augmented with speech recognition (Figure 1). To help adult learners acquire new vocabulary, we embed an in-game mechanism that presents vocabulary items to be learned via word-picture associations. Using our working speech recognition prototype, we investigate the extent to which *retrieval practice*, an educationally effective but cognitively demanding strategy, impacts learning and engagement in this fast-paced game environment.

Author Keywords

Design; Education

ACM Classification Keywords

H.5.2. [Information interfaces and presentation]: User Interfaces – Interaction styles, Voice I/O; K.3.1. [Computers and Education] Computer Uses in Education – Computer-assisted Instruction (CAI)

Introduction

The pervasive spread of computer games has made a significant impact on game-based learning as a serious topic in the field of education. Research evidence has shown that fun and enjoyment are central to the process of learning because they increase learner motivation [2]. Games embody core elements of intrinsic motivation such as challenge, fantasy, competition, and recognition [9] that make game-play a potentially powerful means for learning. Good games often induce a state of flow, a psychological setting in which the player is so absorbed in the game that the activity persists purely by virtue of intrinsic motivation [11]. Games also offer learners opportunities to practice a new skill and consolidate mastery through "cycles of expertise" [4]. In striking a balance between mastery and challenge, games can motivate players to learn through repeatedly doing the game itself until they have virtually automatized their solutions.

To date, research on educational games has focused primarily on the design of custom arcade-style learning games or elaborate extensions of virtual environments. Less research has explored the potential of combining the addictive effect of existing arcade-style games with the potent learning gains of *retrieval practice*. Retrieval practice, the act of repeatedly attempting recall from memory, not only benefits learning as much as non-recall studying, but also improves long-term retention over study alternatives [8]. Retrieval practice is posited to be powerful because it offers opportunities to strengthen memory encodings through multiple exposures to memory cues. In some cases, retrieval practice has demonstrated an advantage even over more complex active learning strategies, such as elaborative studying with concept mapping [8].

Although the highly repetitive nature of existing arcade games makes them natural settings for embedding retrieval practice, most adaptations of existing games emerge from turn-based frameworks like card games [10] or from complex virtual environments [12], perhaps due to less time pressure on learners and greater amenability to structural changes. However, arcade-style games such as Tetris and Pacman are advantageous in that they are logically much simpler to manipulate, have open source code bases, and allow a wider range of time commitment from players. Just as flashcards enable students to review vocabulary on the run, arcade games allow players to either indulge in short spurts or stay indefinitely.

In this paper, we focus on investigating useful techniques to modify existing arcade games for the purpose of learning, using Tetris as a prototypical example. There are two challenges in this research. First, it is unknown what constitutes a good design for learning amidst an already cognitively intensive and fast-paced game. Second, although previous studies evaluated the effectiveness of retrieval practice, to the best of our knowledge there have been no studies measuring the impact of retrieval practice within a time sensitive, arcade-style game. Because retrieval practice exerts greater cognitive effort than non-recall studying, and the interactivity of gaming introduces an additional cognitive load [3], it is unclear whether retrieval practice will remain advantageous in an intense game setting. Cognitive research suggesting somewhat independent working memory channels for visuo-spatial and phonological processing [1] gives some hope to the possibility of uniting vocabulary learning with visually demanding arcade games.
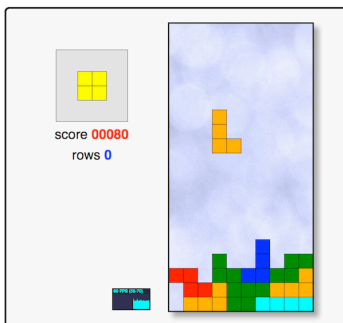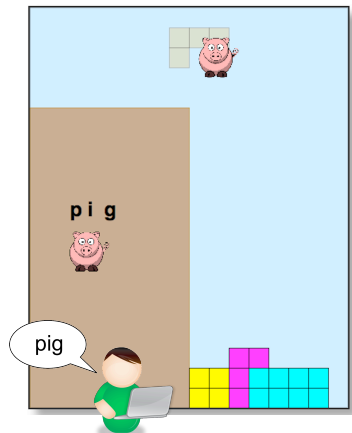

Figure 2: Open source Tetris implementation.

Figure 3: Saying the word associated with the picture unlocks block rotation.
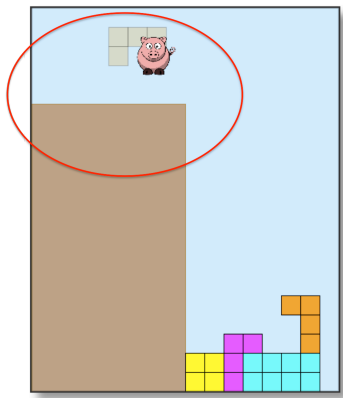

Figure 4: Tetris piece moves horizontally through a chute to give the learner more time.

To explore these questions, we built a modified version of Tetris, augmented with educational features and speech recognition. Through insight from initial user studies, we propose two techniques to make learning more natural in a fast-paced game setting: 1) making up for lost time, and 2) increasing efficiency of input and feedback. With our Tetris prototype, we investigate the extent to which retrieval practice impacts learning and engagement through an online user study.

Game Design

Building on an existing open source web implementation of Tetris[1] (Figure 2), we modified traditional Tetris rules to offer an incentive for learning word-picture associations. Players see each Tetris block attached to a picture and must correctly speak the word associated with the picture before block rotation can be unlocked for the trial (Figure 3). We selected this particular rule modification because it most closely aligns word-learning incentives with the core means to success in the original game.

We enhanced traditional Tetris with multimodal speech interaction for two reasons. First, prior research in speech-based literacy games has indicated that productive speech practice strengthens word knowledge by providing learners with phonological input back to the mind [7]. Secondly, because arcade games like Tetris are already fast-paced and involve multiple keystrokes in controlling game pieces, speech interaction allows players to keep up with the pace of the game more so than textual input.

---

[1] http://codeincomplete.com/posts/2011/10/10/javascript_tetris
© Jake Gordon

Interface Improvements

Our system was iteratively prototyped and refined over a period of six weeks. In total, eight MIT students and 20 workers on Amazon Mechanical Turk play-tested successive versions of the game.

From initial usability tests, we observed that users needed additional time to speak the target word before block rotation. Players also instinctively started maneuvering blocks as soon as they appeared, even though this multitasking distracted them from word learning. To address these issues, we redesigned the blocks to move horizontally for a few seconds before dropping (Figure 4). The horizontal chute provided extra time for the learner to recall and speak words without undue pressure to maneuver the block.

Due to the additional time allocated to speaking, users cleared rows less frequently than in regular Tetris games. Recognizing that the pace of feedback on progress is crucial to the experience of flow [11], we shortened each row from 10 to 8 columns wide so that users could experience progress at a rate more similar to that of regular Tetris. Users indicated that easier row clearing did not negatively impact the game in light of the additional stimulus gained from word learning.

An important requirement in a fast-paced interactive experience is the efficiency of user input and system feedback. Because the user's gaze is on the block at the time he says the word, we added multisensory feedback to be delivered at the locus of attention. Whenever users pronounce the correct word, they not only hear a "success" game sound but also witness the picture on the block disappear as the block transforms from translucent to opaque. To increase the efficiency
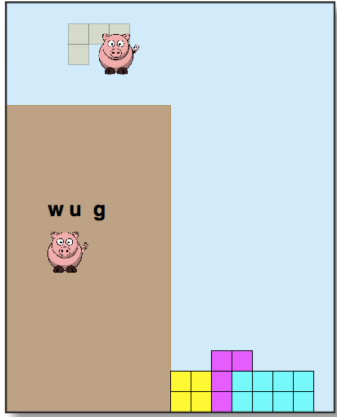
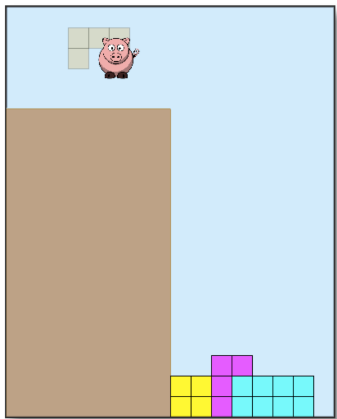**Figure 5a**: First appearance of the word-picture pair.

Figure 5b: In the retrieval condition, learners must recall the associated word every subsequent time.

of user input, our final interface uses a spring-loaded hold-to-talk spacebar, rather than the more traditional two-step process of push-to-record followed by push-to-stop.

Speech Recognition Improvements

To recognize speech input, we used the WAMI (Web-Accessible Multimodal Interface) software [6], a framework that allows audio to be captured at the web page and transmitted to the SUMMIT speech recognizer [5] running remotely.

During initial user testing, we observed that false negative speech recognition errors (i.e. user spoke correctly but speech recognizer misrecognized) were most common and significantly inhibited the flow and enjoyment of the game. False positive errors were rarer because users tended to speak only when they were fairly confident in their answer. Because the target answer was revealed whenever the user succeeded, users appeared amused rather than misdirected by false positives. To decrease the likelihood of false negatives, we add words to the recognizer's vocabulary dynamically as words are introduced in the game, and furthermore expand the set of accepted responses to include the top three speech recognition hypotheses once the vocabulary size reaches a sufficiently large threshold. Thus, if the correct word appears anywhere in the top few words, it is selected as the user's utterance.

Preliminary Study

Although retrieval practice has been studied extensively in memory and cognition communities, there is limited work exploring the dual effects of retrieval practice on both learning and entertainment within the context of a fast-paced game. We conducted a preliminary study to understand this issue.

In order to evaluate learning gains without giving any user an advantage due to prior linguistic exposure, we generated an artificial vocabulary using a probabilistic model on English phonemes. The vocabulary consisted of 14 novel names for familiar animals.

Thirteen users on Amazon Mechanical Turk between the ages of 21 and 51 (M=31.8) took part in our 15-minute within-subjects study. Subjects played two games of the modified Tetris, one with retrieval practice and one with only study practice. In the retrieval condition, learners saw the word-picture pair the first time it appeared (Figure 5a), and subsequently only saw the picture displayed (Figure 5b). The word was revealed if the learner had said nothing after four seconds, or as soon as the learner recorded a response regardless of correctness. In the study condition, learners instead saw the word-picture pair displayed every trial (Figure 5a). Order of presentation and word sets were fully counterbalanced, and participants were randomly assigned to conditions.

In each condition, users learned the semantic meaning of seven artificial words during gameplay. The words were split into two groups for initial introduction and three rounds of practice within each group, followed by two repetitions of all seven words. After each condition, participants completed a 45-second distractor task consisting of simple arithmetic questions, followed by a two-part quiz. The quiz consisted of a production component, in which participants saw pictures one at a time and filled in the associated word, and a multiple-choice component, in which participants selected the answer from the word list given a displayed picture. For
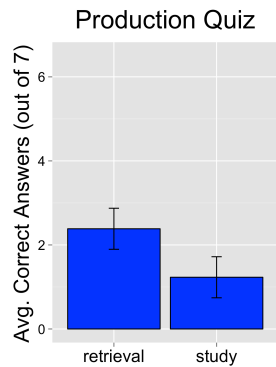
Production Quiz

Figure 6a: Production quiz scores with 95% confidence interval, 3-5 days after exposure.
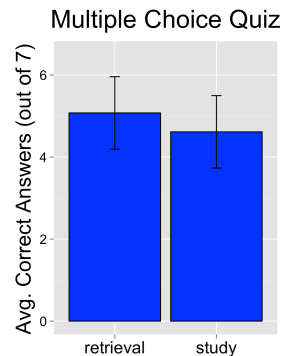
Multiple Choice Quiz

Figure 6b: Multiple choice quiz scores with 95% confidence interval, 3-5 days after exposure.

the purpose of assessing long-term retention, a delayed post-test was administered to the Amazon Mechanical Turk workers between 3 and 5 days after gameplay.

Informed by previous work on retrieval practice, we have the following hypotheses for this study:

H1: Retrieval practice will be advantageous for long-term retention, an advantage that will be more salient in the production evaluation due to practice in active recall.

H2: Those who struggle with word learning will likely find the study condition more enjoyable, due to the high cognitive burden of learning while playing a game.

Results

Multiple-choice quiz results suggest that, with even as little as 15 minutes of arcade-style game play, the average learner can retain at least half of the 14 total total word-picture associations 3-5 days after initial exposure. This gain is particularly promising given that, unlike many other educational games studies, the words are introduced during the game rather than in a pre-game training period.

We ran a 2x2x2 (practice type x presentation order x word set) repeated measures ANOVA on the dependent variable quiz score. Although no significant differences were found in multiple choice evaluations (Figure 6b), results revealed a minor interaction effect of practice type and order on the production quiz score ($F(1,9) = 4.71$, $p = 0.058$) immediately after the game. In the follow-up production quiz 3 to 5 days later, the retrieval condition demonstrated a significant advantage over the study condition ($F(1,9)=19.47$, $p=0.002$), as illustrated in Figure 6a. Thus, there is critical evidence

supporting a significant long-term production benefit of retrieval practice over study practice, even when the user is actively producing speech utterances in both conditions. This advantage manifests more strongly in the setting where players experience the retrieval condition after the study condition, suggesting that acclimation to the game may be a prerequisite for benefitting from retrieval practice.

However, the question remains as to whether the potent benefits of retrieval practice are worth the potential risk of reduced enjoyment, particularly for learners who struggle with word learning. To discover whether engagement differed between low and high performers, we divided participants in half, separating them into two groups based on quiz scores. High performers rated their enjoyment equally between the two conditions, with only one person preferring the study condition. In contrast, only half of the low performers rated them equally. The other half preferred the study condition, noting that it was less stressful and gave them more opportunities to learn the associations due to greater visual exposure. In line with these trends, low performers also reported having learned better in the study condition, even though this was not always consistent with actual learning gains.

Such findings suggest the importance of balancing between potent learning strategies and user engagement, particularly when the act of retrieval may impose a cognitive burden on slower learners. Because the experience of flow hinges more on perceived than on actual skills and challenges [11], an in-game educational feature that gives players more opportunities to succeed mid-game may sustain engagement for a longer time despite slower learning.

Conclusion and Future Work

In this work, we have proposed several design techniques for leveraging existing arcade-style games in a learning context, such as adjusting for lost time and optimizing efficiency of input and feedback. In our controlled study, we have also demonstrated that, even when embedded in a fast-paced game, retrieval practice offers a potent production vocabulary gain over studying, but that this benefit may come at the cost of reduced engagement, particularly for slower learners.

While existing arcade-style games have experienced limited adoption in the realm of education compared to custom-made arcade games or virtual environment games, our promising results suggest that this domain deserves more attention. In particular, greater emphasis should be placed on understanding potential trade-offs between learning benefits and user enjoyment in the educational adaptation of fast-paced games. The results from this study also open a gateway for exploring other forms of retrieval practice that may be less cognitively demanding and include more instructional scaffolding, such as displaying a few word choices. Gradual adaptation of learning challenges, perhaps by offering a combination of study and retrieval trials, should also be explored in future work.

Acknowledgements

References

[1]   Baddeley, A.D., & Hitch, G. *Working memory.* In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* 8 (1974), 47–89. New York: Academic Press.

[2]   Bisson, C., & Lunckner, J. Fun in learning: The pedagogical role of fun in adventure education. Journal of Experimental Education 9, 2 (1996), 109-110.

[3]   deHaan, J., Reed, W.M., and Kuwada, K. The Effect of Interactivity with a Music Video Game on Second Language Vocabulary Recall. Language Learning & Technology. 14.2 (2010): 74-94.

[4]   Gee, J. What video games have to teach us about learning and literacy. Palgrave Macmillan, New York, NY, USA, 2003.

[5]   Glass, J. A probabilistic framework for segment-based speech recognition. *Computer Speech and Language* (2003).

[6]   Gruenstein, A., McGraw, I., and Badr, I. The WAMI Toolkit for Developing, Deploying, and Evaluating Web-Accessible Multimodal Interfaces. *In Proc. ICMI* (2008).

[7]   Kumar, A., Reddy P., Tewari, A., Agrawal, R., and Kam A. Improving literacy in developing countries using speech recognition-supported games on mobile devices. *Proc. CHI. 2012*, ACM Press 2012, 1149-1158.

[8]   Karpicke, J. D. Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21* (2012), 157-163.

[9]   Malone, T. (1980). What makes things fun to learn? A study of intrinsically motivating computer games. Palo Alto, CA: Xerox.

[10] McGraw, I., Yoshimoto, B. and Seneff, S. Speech-enabled Card Games for Incidental Vocabulary Acquisition in a Foreign Language. Speech Communication, 2008.

[11] Nakamura, J., & Csikszentmihalyi, M. Flow theory and research. In C. R. Snyder & S. J. Lopez (Eds.), Handbook of positive psychology, 195-206. Oxford: Oxford University Press, 2009.

[12] Van der Spek, E.D. (2011). Experiments in serious game design: a cognitive approach. Doctoral dissertation