

---

# Multimodal Literacy: Storytelling Across Senses



**Suana Sanchez**  
Keio University, Japan  
Su\_saps@hotmail.com

**Heng Gu**  
Keio University, Japan  
heng.gu@network.rca.ac.uk

**Kai Kunze**  
Keio University, Japan  
kai.kunze@gmail.com

**Masahiko Inami**  
Keio University, Japan  
inami@inami.info

---

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).  
*UbiComp/ISWC'15 Adjunct*, September 07-11, 2015, Osaka, Japan.  
Copyright is held by the owner/author(s). Publication rights licensed to ACM.  
ACM 978-1-4503-3575-1/15/09 . . . \$15.00  
DOI: <http://dx.doi.org/10.1145/2800835.2807940>

## Abstract

This paper describes a novel framework for literature that presents the reader with a multimodal experience (visual, haptic and auditory). We present an initial prototype, extending an iPad with a surface vibration transducer for haptic feedback to augment the reading of a short story. Text provides the reader with verbal information, whereas other senses perceive the context through non-verbal cues. We use our own framework to create this storytelling across senses, integrating text, sound, animation and tactile sensation. We believe with a multi-modal reading experience, users will enjoy reading more, will feel more immersed and remember the narrative better.

## Author Keywords

Dual-coding, valid co-occurrences, multimodal, reading.

## ACM Classification Keywords

H.5.m [Information interfaces and presentation (e.g., HCI)]:  
H.5.2 User Interfaces, Interaction Style.

## Introduction

With more and more interactive media, there is a constant decline people reading books over the recent years. Given the benefits of concentrated reading (e.g. higher vocabulary and critical thinking skills[3]), we want to use tech-

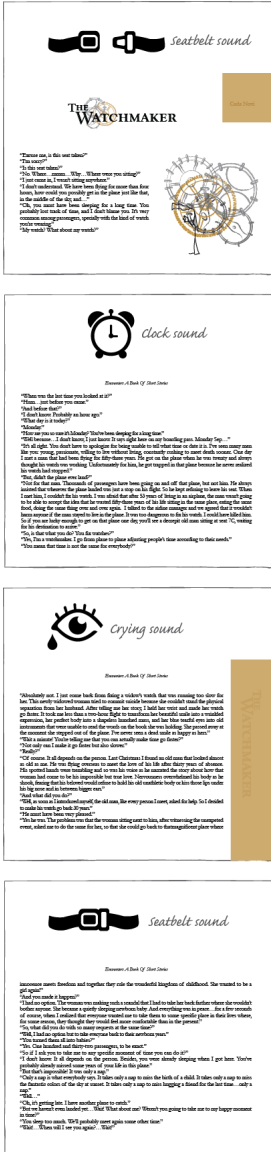


Figure 1: The short story of "The Watchmaker" augmented with sounds

nology to increase the users immersion, comprehension and recall especially when reading fiction. Thus, helping users to focus more on stories.

The perception of literature is purely visual, where the written word creates fictional narratives to describe past, future, and hypothetical environments. To do so, verbal expression is transformed into a visual coding of sounds, where meaning changes to admit a self-reflective subject. This change, allows literature to convey meaning in itself, leaving the narrative deprived of any stimulation other than the visual. However, perception is not limited to only one sense. Events in everyday life are registered by more than one modality, integrating the different information from various sensory systems in a unified perception.

The contributions of the paper are as follows: (1) we present a multimodal sensing approach to increase immersion while reading, (2) we describe a first prototype using visual, haptic and auditory clues aimed at increasing reader immersion, comprehension, and recall (3) we augment one story and performed an expert evaluation, (4) we describe a study design to assess immersion, comprehension and recall.

## Approach

We propose a multimodal framework for fictional narratives, that can be perceived in a mixture of senses: visual, auditory, and haptic. This multi-sensory literature provides with extra-textual content to create context. The simultaneous sensory information is designed to act as valid co-occurrences,[1] where the sensory information coming from one event is synchronization of light, noise and tactile sensation. These co-occurrences are set to produce a cross-modal interaction where the interpretation of read actions is influenced by adding information available in other senses; taking advantage of the intermodal redundancy.

The narrative offers a mix of verbal and non-verbal information that is processed in different channels, creating separate representations. In this way, we apply a dual-coding [7] across senses, simultaneously dealing with verbal and non-verbal cues to enhance perception. While language has a particular system to deal directly with writing, it serves at the same time, with a symbolic function of non-verbal cues. By delivering non-verbal cues to multiple senses we provide the reader with a storytelling that remembers it's oral past.

### Pre-Study: Sonic cues

In a first approach, to verify if multimodality is beneficial for reading, we did a study only using sonic cues. We evaluated fifty nine readers in an AB online survey: with and without non-verbal sonic cues. Both groups were directed to a link of the short story "The Watchmaker" by Carla Novi. In this prototype, we embedded the sound calculating the average time speed to release the sound. The study looked into the difference of engagement with the aid of cross-modal literacy, using an online immersion questionnaire, based on [4], to evaluate the levels of how challenging the text was; how boring; how distracted the participants were; and how interested they were in the short story. The immersion questionnaire showed the control group had a higher level of boredom compared to the study group who listen to the sound cues. However, as other comparisons (except the level of boredom) were not significant, we decided to redesign and adapt our approach adding haptic and visual interactions.

Question	With sound	No sound	Weighted Average
Challenging	17.86%	32.26%	1.82 >1.68
Boring	7.14%	19.35%	1.93 >1.81
Distraction	21.43%	29.03%	1.79 >1.71
Engagement	89.29%	90.32%	1.11 >1.10



## Initial Prototype

We created a multimodal storytelling that can be perceived by the eyes, ears and hands. We provide content through the text and create context from non-verbal cues in a visual, sonic and haptic interaction.

### Software

We use the Processing.org prototyping environment to create our prototype that displays the text, generates the vibration and audio signals. The reader's position in the story changes the multimodal cues to fit the part of the story being read. For instance the page that is distorted by the visual effect of rain drops changes from left to right page depending on which page is being read. We are designing a set of multimodal clues fit for a variety of short stories and working on an authoring tool that makes it easy to create multimodal reading experiences for new texts (defining positions and effects).

### Hardware

The multi-sensorial reading device uses the large surface vibration transducer to simulate haptic sensations and textures through vibration patterns' while the small speaker provides auditory stimuli. Both connect to an iPad mini through a stereo jack, which splits the input into vibration (left) and audio (right) channels. We use rudimentary eye-tracking with iPad camera to detect if the reader is focusing on the left or right page [5].

## Short Story

We augmented so far one short story, "John", with multi-sensory cues that communicate an atmosphere of a rainy day. John is having the worst day of his life, and this is reflected in the weather. Even though the text does not explicitly tell the reader it is raining, this can be inferred from the words "John's shirt was soaked". This aspect of the story is reinforced by the haptic visual and auditory cues of rain drops. The cross-modal cue of rain has an

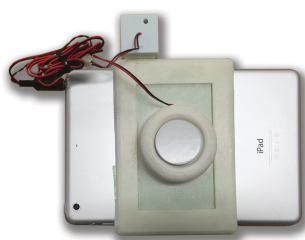
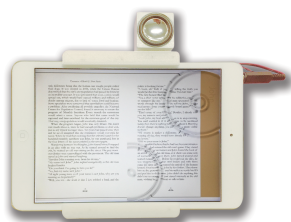
important emotional connotation in the story, since in an urban setting rain is often seen as irritant. The rain darkens the sky symbolizing John's misery. When the user starts reading the story, they get the increasing ambient sound of rain (first just a few drops later constant rain) and the haptic sensation of the rain drops on their hand.

## Expert Evaluation

We showed and discussed the augmented story on the initial prototype with 3 HCI experts (researchers in the field for over 10 years). One with background in haptic sensation. The feedback was in general positive. They believe the system can support immersion in the story. The expert with haptic feedback experience mentioned mounting the transducer in the back of the device limits the haptic sensation to some ambient haptic feelings. He recommends using "ambient" haptic feedback (e.g. rain, motor vibrations). We will curate more sensations suitable for feedback and augment more stories based on his input.

## Experimental Design

In an AB study similar to the sonic-cue modality, we want to look for the difference in immersion and recall between the two modalities. Our experimental setup includes the users either reading the short stories (so far 4 selected) without augmentation or with our multi-sensory framework. Afterwards they answer two questionnaires, one for immersion and one for comprehension. This is followed by an interview where the user retells the story; in order to measure the level of detail perceived and the complexity of the story's mental simulation. The Immersion Questionnaire looks into levels of engagement based in the immersion experience.[4] Here users rate from 1 to 5 different levels of immersion by answering if they agree with different statements, such as "I found the story to be boring". The Reading Comprehension Questionnaire looks for reading comprehension, using a sentence verification task (SVT) test of 16 sentences di-



**Figure 2:** Multimodal storytelling with sonic, visual and haptic cues.

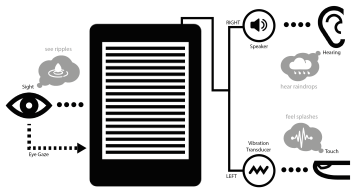


Figure 3: System's diagram.

vided in 4 groups. A SVT rests on the readers' memory retention preserving meaning during mental simulations of what is being read and not on exact words.[8]

Story Retell is an oral retelling from the subject. Using [9] methodology, the recorded session is compared to a prepared outline of the story and scored in six categories: Characters, Event details, Plot, Setting, Theme, Personal Connections.

### Discussion

We argue that by providing the reader with non-verbal extra-textual cues, we can enhance the reading experience. Covering a fuller range of senses will not only engage the reader more, calling for her attention, but will also provides aids for mental imagery, which has a connection to reading comprehension.[3]

### Related Work

There is a lot of work in multimedia and interactive e-books, especially for education [6]. Other researchers look into eye-tracking to trigger multi-modal information. For instance Text 2.0 [2] uses the eye-gaze to know where the reader is within the text, to provide with extra verbal-information when needed. If the reader spends more time with a word, the text will automatically provide the reader with the definition of that same word. As far as we are know, we are the first to explore the argumentation of a narrative with audio, haptic and visual stimuli to increase reader immersion, comprehension and recall.

### Conclusion and Future Work

In this paper we present our framework for literature that stimulates multiple senses to create a more immersive reading experience (see Figure 3). We introduced our results from sonic cues pre-study and present an enhanced prototype with haptic, sonic and visual cues to create a valid co-occurrence of context; and how to test it.

### Acknowledgements

This work is supported in part by the CREST project from JST and JSPS Grant-in-Aid Young Researchers B No. 26730095.

### References

- [1] Bertelson, P., and de Gelder, B. The psychology of multimodal perception.
- [2] Biedert, R., Buscher, G., Schwarz, S., Hees, J., and Dengel, A. Text 2.0. In *CHI'10 Extended Abstracts*, ACM (2010), 4003–4008.
- [3] Gambrell, L. B., and Bales, R. J. Mental imagery and the comprehension-monitoring performance of fourth- and fifth-grade poor readers. *Reading Research Quarterly* (1986).
- [4] Jennett, C., Cox, A. L., Cairns, P., Dhoparee, S., Epps, A., Tijs, T., and Walton, A. Measuring and defining the experience of immersion in games. *International journal of human-computer studies* 66, 9 (2008), 641–661.
- [5] Kunze, K., Ishimaru, S., Utsumi, Y., and Kise, K. My reading life: Towards utilizing eyetracking on unmodified tablets and phones. In *UbiComp '13 Adjunct*, ACM (New York, NY, USA, 2013).
- [6] Mana, N., Mich, O., De Angeli, A., and Druin, A. Interactive e-books for children. In *IDC '13* (2013).
- [7] Paivio, A. A dual coding approach to perception and cognition. *Modes of perceiving and processing information* (1978), 39–51.
- [8] Royer, J. M., Greene, B. A., and Sinatra, G. M. The sentence verification technique: A practical procedure for testing comprehension. *Journal of Reading* (1987), 414–422.
- [9] Sadoski, M. An exploratory study of the relationships between reported imagery and the comprehension and recall of a story. *Reading Research Quarterly* (1983).