Augmented Narrative: Amending the Dialog Between Writer and Reader in an Interaction with Sound.

Susana Sánchez Pérez

Graduate School of Media Design

Keio University

Category: Design

Doctor of Media Design

This thesis is submitted to the
Graduate School of Media Design
for the degree of:

DOCTOR OF MEDIA DESIGN
at
KEIO UNIVERSITY
in the category of Design

Abstract

Augmented Narrative creates an interaction for literary storytelling by transforming the book into a cognitive agent capable to assist its reader. The system consists of smart glasses worn by the reader to detect her engagement. This glasses allow the literary storyteller to collect data of the reader's mental workload in real time. A higher mental workload is a signal that the reader is being transported into the story, and thus tells the book that the reader is engaged. However, if the system finds that the mental workload is low, it infers that the reader has not been able to immersed herself into the narrative text. This is when the book, now a cognitive agent, can intervene. The intervention is design to support and complement the text rather than change the written word, or take the reader's attention away from it. Thus, the system uses sonic interventions of non-verbal information that do not compete with the reading task, but give the reader extra-textual cues in a second channel to support imagery. For instance a soundscape that depicts the landscape of the story's setting, which the reader can integrate in a dual coding across senses.

The biofeedback, made possible by the smart glasses, allows a natural dialog that resembles oral communications. Literature changes into a storyteller that preserves the text in the authorial omniscience of the written word, mixed with the spontaneity of sound, that addresses a particular situation that vanishes with the situation itself. In this way Augmented Narrative is set to redefine literature as an empathetic storyteller that adapts to each individual giving an improved reading experience that brings the ear into an activity that had only been for the eyes.

Acknowledgements

I would like to express my special appreciation and thanks to my principal advisor professor Naohito Okude, whose support and guidance allowed me to grow as a researcher. I have enjoyed every book and every subject he had recommended, which in turn, allowed me to discover a world of knowledge that became my passion. Your advice has been invaluable. I would also like to express my gratitude to my second advisor, professor Masahiko Inami, who never found my lack of knowledge in electronics and technology as an obstacle. Only a challenge that, he believed I could easily overcome. The challenge was not that easy, but his confidence in me made me find a way. I want to give a special recognition to professor Kai Kunze, who arrived when I most needed him. He opened the door to an academic community to which I had not been able to gain access to. A big thanks for always treating me as your equal, as part of your team. It definitely made me give my best every time, to prove I was up to your level. Finally I would like to thank professor Masashi Nakatani, who was the cherry on the cake. Thank you for taking the time to teach and guide me in the last stages of my thesis. To the four of them a final big thank you for being part of this extraordinary journey. I am deeply grateful to the National Council of Science and Technology in Mexico (CONACYT) for their trust and financial support. I am also grateful to the Keio University Doctorate Grant-in-Aid Program for their one-year financial support.

Last but not least, I want to thank my family. My husband who has always been there to listen and help. I must admit, he gave me many good ideas, but above all, he gave me unconditional support. My parents, for whom I want to be the best daughter. I owe them my strength and confidence. To my aunts, whom I can always trust will be there for me, no matter what. And my uncle, who has always looked after me, unconditionally. This thesis is dedicated to them, but specially dedicate to my grandmother, because she is the most amazing woman.

Contents

List of Figures and Tables

FIGURES

TABLES

Chapter 1

Introduction

Synopsis. In oral communication, senders are receivers and vice versa; we are capable to speak but also to listen. However, in print, the written discourse detaches the author form the narrative, making writer and reader unreachable to each other. Until now, the analytical reflection that spontaneity should come as a natural dialogue has been in discordance with the essence of print, as the medium is unresponsive. For this reason, the present research looks into the cognitive process of reading, and how they relate to involuntary changes of the body to create a framework that transform literature into a truly interactive medium. Unbound from the restrictions of print, this new intelligent agent (IA) uses the spontaneity of oral communications to better communicate what has been set on paper. The IA consists of a system called Augmented Narrative, where the author not only codes his narrative into written language, but also spontaneous sonic interventions that create context. The reader then uses the system's biofeedback to interact with the multimodal narrative, allowing for an embodied experience of the literary text.

Reading is our most important achievement, since unlike spoken language, decoding the artificial symbols of the written word is not one of our natural abilities. We are born with an innate ability to tell apart sounds that conform a language, and seem to be pre-programmed to talk without formal instruction. However, this is not the case with reading. This artificial accomplishment comes from the brain's capacity to call upon neural regions designed for other tasks, and putting them together to decode a language that expresses our thoughts and emotions on paper.[1]

1. Sousa, David A., ed. *How the brain learns to read*. Corwin Press, 2005.

Walter Ong, in his exploration on the transition from orality to literacy, asserts that oral speech became literature and not the other way around.[2] Societies were conceived with the aid of oral speech, where human communication and thought were strongly related to sound. The transition came with the Greeks, who transformed verbal expression into a visual coding of sounds, where meaning changed to admit a self-reflective subject.[3] This change, allowed literature to convey meaning in itself, leaving the narrative deprived of any stimulation other than the visual. However, it was Milman Parry who by researching South Slavic heroic poetry found that the Homeric epic, unlike Plato, used mnemonic aids. These are metrical patterns that fit into the composition of a recited poem, such as "swift footed Achilles" to help the sonneteer remember his poem.[4] Parry's argument has been used to define the transition from an oral culture in Homeric Greece to a literary culture in Platonic Greece.

The exploration of Ong into this subject was influenced by his professor Marshall McLuhan. In turn McLuhan's own work in oral culture parts from studying the Irish novelist James Joyce. In fact, McLuhan's work is so connected to Joyce's, that to be able to study him one needs to be familiar with all the works of Joyce, who before McLuhan and Ong, explored the shift form speech to writing and print.[5] For instance, in Joyce's *Finnegans Wake* there is an acknowledgement of the importance of the sensory system, where understanding emerges from sound. The language developed by Joyce forces the readers to become aware that *Finnegans Wake* has to be pronounced, preferably out loud by using aural reading, where the reader listens to the story. In examples like this, McLuhan found that Joyce's work is a clear precedent in the study of the reconciliation between orality and literacy, where speech and text become both: the medium and the message.

---

2. Walter, Ong. "Orality and Literacy." *The technologizing of the word, TJPress, London* (1982).

3. Havelock, Eric A. *Preface to Plato*. Belknap Press, Harvard University Press, 1963

4. Parry, Milman, and Adam M. Parry, eds. *The making of Homeric verse: The collected papers of Milman Parry*. Oxford University Press, 1987.

5. McLuhan, Marshall, and W. Terrence Gordon. *The classical trivium: The place of Thomas Nashe in the learning of his time*. Gingko Pr Inc., 2006.

As Joyce, Toni Morrison, Nobel Prize in 1993, drawn by the tradition of oral storytelling of African American cultures, makes use of aural reading. In her fiction, she encourages a non-literary experience, urging the reader to respond as an illiterate or preliterate reader would. For Morrison, as for Joyce, orality is crucial for understanding. Morrison privileges oral memory in her characters as much as she does in her readers. For example, in *Song of Solomon* there is a battle between orality and literacy, where Morrison uses different spelling for the name of Solomon, with the literate form Shaliman, and oral form of Charlemagne. In this way she uses literacy as an impediment for discovery, not only for the main character, but for the reader as well. The false orthography and resilience of the reader to focus on the written word, rather than the sound, prevents the perception of the similarities between names. As Joyce, Morrison builds on an oral experience, using sound to provide the reader with the answer to the written plot. Even though we experience reading as a silent activity, sound has remained as an important component. In this way Morrison exercises Ong's call to reconsider the eighteen-century literary conventions of exact spelling to reconcile orality and literacy. In *Song of Solomon,* she uses the storyline to submerge the reader into the irony of Plato's arguments against writing, where the orality of African American cultures is preserved because it has been set on paper. [6]

In his *Seventh Letter* epistle, Plato disregards writing as a legitimate way to convey knowledge in any significant way. Written words are unresponsive and indiscriminate to whom they address themselves.[7] His criticism comes from the resistance of the enormous change in Greek culture, from a dominantly oral tradition to a literate one. However, the change from orality to literacy is a more gradual one, going through what Ong calls residual oral cultures, where literature has not been fully interiorized.

---

6. Middleton, Joyce Irene. "Orality, Literacy, and Memory in Toni Morrison's Song of Solomon." *College English* (1993): 64-75.

7. Dobra, Susan. "The gift of Theuth: Plato on writing." Retrieved from: www.csuchico.edu/phil/sdobra _mat/platopaper.html. September 2nd, 2014.

The acclaimed playwright William Shakespeare is part of these residual oral cultures. Because of it he was able to make a profound impact in the English language, refashioning oral expression through pen and paper. Shakespeare has been credited for over 1700 original words by changing nouns into verbs, verbs into adjectives, and by connecting words never before used together. As part of a residual oral culture, it is unlikely that Shakespeare was involved in publishing his own plays, since his writing was meant to be spoken, and not read.[8] His plays entail an interaction that is directly addressed to a particular audience. As John Marston, a contemporary playwright claimed: scenes invented to be spoken should not be enforced to the public for reading. It is even suggested that Shakespeare's literary debut is a consequence of piracy, where publishers copied his plays in Quartos. Or as his editors found, corrupted memory-reconstructions of the play.[9] Be that as it may, it is in these unlawful Quartos that we can find the orality that comes from writing and reshapes the text to improve the storytelling. For instance, in print, punctuation is the only extra-contextual clue for intonation; as in *Hamlet* "To be or not to be, that is the question." the opening remains in the same textual intonation with each reading. However, for the play, extra-textual context is missing, for Shakespeare, his audience was unknown, leaving the actors to determine how to deliver the text on stage. In the 1603 Quarto of *Hamlet*, thought to be a corrupted memory reconstruction, the opening reads: "To be or not to be, Aye. There's the point. To die, to sleep, is that all? Aye all". This difference in words appears like a spontaneous modification of the text by the actor, calling for the audience's attention, in an interaction unable in print. This change of lines by the actor provides evidence that there is an interaction with the audience. In this example of *Hamlet*, the actor makes spontaneous changes to stimulate his audience, trying to keeping their attention, and hopefully drifting them away from boredom.

8. Pollard, Alfred W. *Shakespeare's Fight with the Pirates and the Problems of the Transmission of His Text*. Cambridge University Press, 2010.

9. Mibillard, Amanda. Shakespeare in Print. *Shakespeareonline*. 20 August 2004. Retrieved from: www.shakespeare-online.com/biography/shakespeareinprint.html.

Boredom comes from the ineffectiveness of the external world, in this case the actors on stage or the storyline of the play in itself, to provide helpful stimulation. This lack of stimulation implies a refusal from the audience to keep paying attention to the play. Nevertheless, theatre similar to books as storytelling mediums, are supposed to resist boredom, since performance and literature are cultural advancements meant to withstand it. For instance, at the moment of picking up a book, the reader expects to have her interest engaged.[10] However, boredom can come from the lack of understanding, creating a feeling of anxiety or frustration, emotions that Sean D. Healy [11] argues are just a different physical form of boredom. In this sense, the word 'boring' serves as an emotional connotation to describe literary works that cause frustration and antagonisms to their reader. For example, such connotation could be applied to some period literature. Set in a historical present, period literature turns out to be problematic for subsequent readers, since it confers the readers with a simplified lessons from the past that becomes meaningless with time, lacking any form of stimulation. Narratives that use moral and theology doctrine in their plot, such as old fashion Victorian didacticism, are set to provoke permanent yawns in contemporary readers; books like *The History of Sir Charles Grandison*, *Coelebs* or *In Search of a Wife* or *Robert Elsemere*. These novels may arouse frustration to today's readers with its unfamiliar vocabulary and complicated sentence structure.

To avoid boredom, books, with their authorial omniscience, place non-natural characters into a natural frame, allowing authors to stimulate the reader with the artificial effect of experiencing and viewing outside the natural world. In turn, readers attribute a mental stance to these non-natural characters -in the same way they do with people in everyday life- by building mental reconstructions of read emotions and actions.[12]

---

10. Spacks, Patricia Meyer. *Boredom: The literary history of a state of mind*. University of Chicago Press, 1995.

11. Healy, Sean Desmond. *Boredom, self, and culture*. Rutherford, NJ: Fairleigh Dickinson University Press, 1984.

12. Ashby, J., and Rayner, K. (2012). Reading in alphabetic writing systems: evidence from cognitive neuroscience. *Neuroscience in Education: The good, the bad, and the ugly*, 61

This simply means that to understand other, we place ourselves in "their shoes".[13] Empathy is the mind-reading ability that, as done in the natural world, allocates mental states to fictional characters.[14] This is because literature capitalizes on, and simulates mind-reading mechanisms used on a daily basis; even when narratives give minimum textual cues to construct a mental representation.[15] Here, the reader understands the emotional stance of the fictional character by finding behavior patterns based on narrow slices of self-experience for recognition.[16] To prove this, Phelan and Rabinowitz studied the case of Temple Grandin, a woman with a low degree of autism, who unable to interpret emotions in facial expressions. [17] For Grandin, reading fiction is just like reading the named facial expressions, since both call for mind reading; much needed in human communication. Thus, she lacks desire in reading narratives, since making sense of the text is linked to the verbal constructions of characters with potential for thoughts that act as scaffolding for empathy by predicting emotions through actions.

Malcolm Gladwell calls this "rapid cognition", where a person feels something is not right but cannot tell why. [18] In Gladwell's idea of rapid cognition, decisions are made about someone in a split-second, based on the body's ability to react before the mind, by using small cues of information. For example, he tells about a study that predicted which doctors were more likely to get sued by their patients. Based not on their medical abilities, but in small slices taken from the doctor's linguistic tones.

---

13. Goldman, A. (2005). Imitation, mind reading, and simulation. In S. Hurley, and N. Chater, Perspectives on Imitation II (pp. 80-81). Cambridge, MA: MIT Press

14. Herman, David. "Narratology as a cognitive science." *Image and Narrative* 1 (2000): 1

15. Rabinowitz, Peter J. *Before Reading: Narrative Conventions and the Politics of Interpretation. The Theory and Interpretation of Narrative Series*. Ohio State University Press, 1070 Carmack Road, Columbus, OH 43210-1002, 1997

16. Gladwell, Malcolm. *Blink: The power of thinking without thinking*. Back Bay Books, 2007

17. Phelan, James, and Peter J. Rabinowitz. *Understanding narrative*. The Ohio State University Press, 1994

18. Gladwell, Malcolm, and Madelon E. Ruiter. "Blink: the power of thinking without thinking." *Gedragstherapie* 41, no. 2 : 199. 2008

The study found the doctor's voice produced a physiological arousal contracting muscles, elevating the heart rate and accelerating respiration that was then translated by patients into specific emotion of fondness or aversion; if the voice was warmth and understanding, or firm and cold. More importantly rapid cognition is a mind-reading ability that can be translated to literature, as it promotes the construction of characters with potential for thought. Furthermore, rapid cognition can be seen in theatre, where audiences have a remarkable ability to understand rapidly and efficiently the settings of the play based on the behavior of sound-producing objects and sounds sources. Even when the source is not in their field of vision sound effects are a powerful tool for stimulation. Audiences are able to recognize the sounds by linking what is being heard to sounds heard in the past and therefore realizing the identity of the source, and the significance it has to the listener. Even when the sensory information is insufficient for the listener, the perceptual system still analyses the situation taking into consideration previous knowledge acquired from the surrounding sound world.[19] In a cognitive point of view, the sonic information must be interpreted to give a coherent understanding. Furthermore, Ecological Psychology, [20] tells us that the physical nature of the sounding object, the means by which it has been set into vibration and the function it serves to the listener, are perceived directly without any intermediate processing. Even though this approach seems to contradict Information Processing, [21] where recognition is a multi-stage process between perceptual qualities of sound source, abstract representation in memory, meanings, and associations, the ecological approach is concerned with the invariant properties of the sound. For instance, one is able to recognize someone's voice even if this person has a terrible cold. Thus the ecological approach might involve an understanding on why sound effects, even when they are synthetic can produce the same understanding as natural sounds do to the same listener.

---

19. McAdams, Stephen Ed, and Emmanuel Ed Bigand. "Thinking in sound: The cognitive psychology of human audition." In *Based on the fourth workshop in the Tutorial Workshop series organized by the Hearing Group of the French Acoustical Society.* Clarendon Press/Oxford University Press, 1993.

20. Michaels, Claire, and Claudia Carello. *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall, 1981.

21. Lindsay, Peter H., and Donald A. Norman. *Human information processing: An introduction to psychology*. Academic Press, 2013.

As described, theatre offers a mix of visual and sonic verbal and non-verbal sonic information, which the audience processes in a different channel, creating separate representations. Here, sound effects and music are mixed with stage props -that carry non-verbal visual information- and dialog to represent the characters and their relation with the environment.[22] It was Allan Pavio [23] who looked into this unique skill of human cognition specialized for dealing simultaneously with verbal and non-verbal objects and events. He acknowledges that any representational theory must accommodate a functional duality. Thus, he sees language as a peculiar system that deals directly with speech and writing, while at the same time serves with a symbolic function with respect of nonverbal objects, events and behaviors. In his theory of Dual Coding, the general assumption is that there are two classes of phenomena handled by separate cognitive subsystems: one for representation and processing of information; and the other for dealing with language. But unlike theatre, books offer only one input, the visual, neglecting any other kind of stimuli for representation; conveying all information in only one channel. That leaves the written world to only be able to relay in pictures or illustrations.

Along the same vein of cognitive processes, we discuss the perception of literature as the means to transport this last one into the narrative world. These cognitive processes also serve as a method to measure levels of involvement and attention. For example, the study of the empathetic relations that readers build with fictional characters in novels has focused in verbal feedback to measure involvement and attention. In one specific study, the self-involvement of readers has been found in oral expressions such as: “I felt I was there right with Phineas” in John Knowles, book *A Separate Peace* of. [24]

22. Hug, Daniel. “Investigating narrative and performative sound design strategies for interactive commodities.” *In Auditory Display*, pp. 12-40. Springer Berlin Heidelberg, (2010).

23. Paivio, Allan. Mental representations: A dual coding approach. Oxford University Press, 1990.

24. Purves, Alan C., and Victoria Rippere. "Elements of Writing About a Literary Work: A Study of Response to Literature.” (1968).

This involvement, or transportation into the narrative text, is attention provoking, and not an emotional response. Furthermore, this cognitive process can be related to a physical change. The segment of the story affects the degree of involvement, begin more prominent during the climax of the story where the reader is expected to be paying more attention. [25] This deeper involvement takes part in the Flow Theory of Csikszentmihalyi [26] of positive experiences of flow, which come when engaged in task demands where the person is in deep sense of control and the activity feels rewarding. Flow can be considered to involve straining tension and mental load. However, the stress found in flow is a positive experience referred as to eustress. By following the reader's physical changes caused by stress and eustress, we can determine if the reader is in a state of flow or frustrated. For example, Johannes Keller studies the psychological mechanisms of flow-experience, and in his studies he reviewed the demanding character of flow activities in the heart rate variability of overloaded tasks. Keller's measurements of the heart comprise the components of James-Lange theory: arousal and valence. Here there are a physical reactions experienced when, for example, dealing with a predator animal or a potential mate. These physical changes are the same in the beginning for both experiences, like a rise in the heart rate, a change in breathing, and perspiration; this is called arousal. However, the emotional direction between attraction and aversiveness, or in this case fear and love, is measured in the valence. In the James-Lange theory as in Gladwell's rapid cognition the physical reaction comes before the psychological reaction, meaning that the body reacts before the mind. Consequently, emotional and cognitive processes, which do not arise through free, can be followed by monitoring physiological changes, since these are sensorial and intellectual responses to what the body perceives. The physical and psychological are very much part of one another, and if the physical comes first, then the physical changes instigated by the nervous system give a first insight of cognitive processes and emotional states.

---

25. Angelotti, Michael, Ralph R. Behnke, and Larry W. Carlile. "Heart rate: a measure of reading involvement." *Research in the Teaching of English* 9, no. 2 (1975): 192-199.

26. Czikszentmihalyi, Mihaly. "Flow: The psychology of optimal experience."*Praha: Lidové Noviny. Cited on page* (1990).

Our nervous system is divided in peripheral and central. The peripheral nervous system deals with sensing information, where the Autonomic Nervous System's components regulate the involuntary responses from information within the body. To do so, the body uses a system of parasympathetic and sympathetic components to relax and stimulate organs and glands. For instance, involuntary responses controlled by parasympathetic components slow the heart rate (HR) and sympathetic components raise it. In the James-Lange Theory, an emotional response is caused by an arousing experience, where the sympathetic nervous system increases the number of heartbeats per minute in the HR. In the absence of arousal, the valence can be detected in the variations between heartbeats, in the heart rate variability (HRV).[27] On the other hand, engagement can be associated with a decrease in the HR; contrary to the increment found in an emotional response, engagement is attention provoking. Thus, information of engagement can be determined by a positive valence, detected in the ratio of low frequency energy to high frequency energy that represents the extent of sympathetic and parasympathetic influence in the HRV. On the other hand the highest frequency indicates boredom, in negative valence, reflecting the lowest mental load.

Similarly, when exposed to stressful situations the sympathetic nervous system is aroused making pupils dilate.[28] Along with pupil dilation, eye activity considers other classes of eye information that contribute as an indication of emotional and cognitive states, for instance blinking. Eye blinks occur between two to four times per minute for functional purposes, however endogenous blinks, or non-functional blinks, are link to behavioral activity and therefore are associated to cognitive tasks. For example, blinking is avoided to maximize perception of information in high content tasks demands, such as reading.[29]

---

27. Keller, Johannes, Herbert Bless, Frederik Blomann, and Dieter Kleinböhl. "Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology* 47, no. 4 : 849-852. (2011)

28. Knapp, B., Jonghwa Kim, and Elisabeth André. "Physiological signals and their use in augmenting emotion recognition for human–machine interaction." Emotion-Oriented Systems (2011): 133-159

29. Irwin, D. E., and Thomas L. E. "Eyeblinks and cognition." *Tutorials in visual cognition* (2010): 121-141.9

During this time, the blinks that do occur coincide with a decrease in attention. They also occur at points when the blink will not interfere with the processing flow of information intake; at the end of a sentence or line. Involuntary blinking can also be evidence of higher nervous processes of information, like emotional states, with an increasing rate in induced high arousal. Even though a decreased arousal reduces blink rate, long period tasks that might lead to boredom or fatigue, increase blink rate.[30]

In addition to eye activity, facial temperature has been found to reliably discriminate between positive and negative emotions, and cognitive tasks. Israel Waynbaum [31] in his Vascular Theory of Emotional Expression attributes the experience of emotions to follow facial expressions rather than preceding them. His hypothesis bases on the fact that the supply of blood to the brain and face comes from the same source: the carotid artery. Therefore, the reactions to circulatory perturbations in the facial artery produce disequilibrium in the cerebral blood flow. The facial muscles contract and push against the skull bone structure and act as a tourniquet on arteries and veins regulating the blood flow; affecting the cerebral blood flow, reducing or complementing it. These alterations cause temperature changes that modify the neurochemistry of the brain. The thermoregulatory action influences peptides and neurotransmitters -temperature dependent- found to produce emotional changes, with cooling associated to pleasant and warming to unpleasant feelings.[32] Moreover, these temperature changes affect the nose, which has been considered a reliable region to detect cognitive tasks. A decrease in nose temperature has been found in an increase of mental workload.[33]

---

30. Stern, J., Walrath, and Goldstein. "The endogenous eyeblink." *Psychophysiology*21, 1 (1984): 22-33

31. Zajonc, R. "Emotion and facial efference: A theory reclaimed." *Science*228, no. 4695 (1985):15-21.

31. Adelmann, P., and R. Zajonc. "Facial efference and the experience of emotion." *Annual review of Psychology* 40, no. 1 (1989): 249-280.

32. I. Stephanos, Ebisch, Aureli, Bafunno, Ioannides, Cardone, Manini, Romani, Gallese, and Merla. "The Autonomic Signature of Guilt in Children: A Thermal Infrared Imaging Study." *PloS one* 8, no. 11 (2013)

33. Or, Calvin KL, and Vincent G. Duffy. "Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement."*Occupational Ergonomics* 7, no. 2 (2007): 83-94.

Mental workload increases with memory work, -used in reading comprehension, where previous knowledge is required- and correlates to engagement tasks of sustained attention.[34] In these three examples the physical activity of the eyes, skin and heart reflect the cognitive process of engagement and boredom giving information of the reading experience without the need of verbal feedback from the reader.

To sum up, the cognitive processes required for perceiving and engaging with fictional literature, along with the potential of sensors to detect these processes in the reactions of the body, provide us with the foundations to create an Augmented Narrative. The system of this Augmented Narrative consists in an intelligent book that provides 1) spontaneous stimulations using 2) sonic interventions when the 3) smart glasses detect a lack of engagement in the physiological metrics.

1.1 Spontaneous stimulation

Shakespeare and his plays hold the clue on how to design an interaction between reader and book: spontaneity. Shakespeare's verbing, changing nouns into verbs, has a deeper impact into human conscience than just transforming oral expression.[35] In his play *Coriolanus,* the main character returns to Rome for revenge. Menenius, well regarded by Coriolanus, is sent to persuade him to halt his crusade for vengeance. After sending him back with no truce, Coriolanus recognizes: "This last old man, whom with a crack'd heart I have sent to Rome, loved me above the measure of a father; nay, **godded** me, indeed." In this dialog, Shakespeare stimulates the reader by turning the noun God into a verb, where the new word godded causes a grammatical disruptions that triggers a P600 wave in the brain.

---

34. Berka, C., Levendowski, Lumicao, Yau, Davis, Zivkovic, Olmstead, Tremoulet, and Craven. "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks." *Aviation, space, and environmental medicine* 78, no. Supplement 1 (2007): B231-B244.

35. Thierry, Guillaume, Clara D. Martin, Victorina Gonzalez-Diaz, Roozbeh Rezaie, Neil Roberts, and Philip M. Davis. "Event-related potential characterisation of the Shakespearean functional shift in narrative sentence structure."*Neuroimage* 40, no. 2: 923-931. (2008)

This wave is a positive voltage variation peak in an electroencephalogram recording, a response found in the posterior part of the scalp that starts about 500 milliseconds after hearing or reading an ungrammatical word. The wave reaches its maximum amplitude around 600 milliseconds, giving the reason for the name P600.[36] More importantly, P600 is related to violations in the probability of occurrence of a stimulus, placing it in the P300 family.[37] This makes Shakespeare's new verbs unexpected stimulation rather than just grammatical disruptions, bringing spontaneity into play. Thus, even though every author has his own literary tools to stimulate his readers, we put forward the idea of translating Shakespeare's grammatical disruptions into disruptions of sound. Sound has many properties that can be exploited to create spontaneous stimulation for attention and perception. The spontaneity of sound can allow literature to reconnect with its oral past.

1.2 Sonic Interventions

Authors like Joyce and Morrison who used aural reading in their literary works, calling for a reconnection between oral storytelling and print, are not the rule. Neither do we live in a residual oral culture as Shakespeare did, ours is a literary one, and a digitalized culture for that matter. Continuing to use Shakespeare and his theatre as a source of inspiration, the proposed sonic spontaneous stimulations can emulate theatrical sonic agency. This could bring orality and literacy closer to assist each other. Since language comprehension is already strongly related to sensorimotor activation, the mental simulation of described events can be multimodal, covering in this way a fuller range of our sensory capabilities.[38] A different incentive from the visual stimulation of print would call for attentional resources to update neural representations of the written word by adding sensory evoked potentials.

---

36. Gouvea, Ana Cristina. How to examine the P600 using language theory: what are the syntactic processes reflected in the component. *Revista do Programa Pos-Graduaçåo en Linguística da Universidad Federal do Río de Janeiro* Vol 7 (2): 88-97 (2011)

37. Coulson, Seana, Jonathan W. King, and Marta Kutas. "Expect the unexpected: Event-related brain response to morphosyntactic violations." *Language and cognitive processes* 13, no. 1 (1998): 21-58.

38. Polich, John, ed. *Detection of change: event-related potential and fMRI findings*. Springer, 2003.

Sound would provide the reader with extra-textual information to help boost her imaginative capacity. A narrative that involves both eye and ear could reinforce the connection between verbal and non-verbal representations that assist mental simulation. These sonic interventions could add extra-textual information to the written word as sound effects do during a play. In this way, the work of individual authors is not affected but enhanced by the spontaneity of sound. This unconscious interactivity with sound leaves readers to concentrate only in the reading activity, promoting the development of reading skills.

1.3 Smart glasses

Physiological metrics applied in Human Computer Interaction, using sensors to detect the emotional and attentional processes caused by the different stimuli of the text, allow the design of an IA. Using the circumflex model of James Russell [39] to divide valence in a positive-negative horizontal axis and a vertical axis of high-low arousal, the reader's psychophysiological metrics can determine not only emotions, but also perception and attention by measuring mental workload. The interaction assists readers to better perceive the meaning of the narrative and keep their attention. This can be done by giving the written word, which has been unable to dissociate itself from the restrictions of print, the ability to create a biofeedback with the readers. To do so, the paradigm we proposes the use of smart glasses that would allow digital books to follow the involvement of their readers. This interaction between the IA and the reader is private, as it follows the evolution of reading from group reading to a solitary activity. The intelligent interface of the book responds to a particular reader in a particular situation to stimulate imagination increase engagement. As a result Augmented Narrative is set to compete with other forms of mass media, withholding the written word as its channel, but through an empathetic medium that better delivers its message using sonic spontaneous stimulations.

39. Russell, James " A circumflex model of affect". Journal of personality and Social Psychology 39: 1161-1178 (1980)

1.4 Viewpoint

In a time of extraordinary advancements for human communication, there is a place for a new digital form of literature. Literature that does not fail to remember its strengths, but embraces its power to make its readers feel and imagine unreachable worlds, holding the most complex meaning in a single sentence. But literature that is not detached from its reader. The proposed interaction of text and sound is thought to supports the readers' imagination, assisting them to better simulate the story, and ultimately keeping them in a engaged. In this way we redefine literature as responsive, combining the sonic information of orality with the enclosed meaning of written language in order to bring back the face to face like-interaction -lost with print, but now available with the advancements of digital communications. This proposal of an interaction between text and reader is still private, as it follows and retains the evolutionary changes of reading from group reading to a solitary activity. The conceptualization of an Augmented Narrative considers literature as part of the human cognitive evolution, and regards reading as the brains most important achievement that needs to be encouraged.

1.5 Contribution

The contribution is to give literature, which has been unable to dissociate itself from the restrictions of print, the ability to communicate and assist tis reader using spontaneous stimulations of sonic information. Accordingly, this work puts forward the concept of an Augmented Narrative, which consists of: 1) smart glasses to create a biofeedback using the metrics of nose temperature to detect engagement; 2) the design of sonic interventions to assist the reader overcome the lack of prior knowledge; 3) prototyping the system with the help of off-the-shelf eye-tracking technology; 4) the implementation and data analysis of the prototype; 5) and finally the evaluation and discussion of the concept.

Chapter 2

Literature Review

Literature can provide a better storytelling bringing the qualities of sound, and using it for an interaction with its reader. This is the core our concept of an Augmented Narrative. On that account in this review of the literature we present an exploration on three different fields that support this concept (See Figure 2.1): 1) Cognitive Neuroscience; 2) Orality & Literacy; and 3) Interactive Storytelling.

Figure 2.1: Overview of the Literature Review

2.1 Cognitive Neuroscience

More than forty years ago the ecological approach started to be applied to psychology. As an illustration, the anthropologist Gregory Bateson coined the term Ecology of Mind (1972). He argued that the mind is defined through and outside the body, in the surroundings of a cognitive ecosystem. He described culture where individual relationships shape socially shared meanings. This collective meaning then informs individuals the understanding of their actions. More recently Hutchins,[1] cognitive psychologist, has followed Bateson's definition when applying the ecological approach to Cognitive Science. Hutchins defines Cognitive Ecology as the study of context, in particular, the mutual dependence between elements in a cognitive ecosystem. Here, the mind arises within a physical system distributed over space and time. Hutchins uses Andy Clark's [i] definition of the brain to suggest sensory and motor processes are not peripheral; making the relations of brain-and-body-interactions with the environment an important unit of analysis. Thus, neuroscientists have become aware of the need to expand the boundaries of 'unit of analysis' to consider a wider cognitive ecology. They now understand one modality may lie partly in another modality. For example, Town et.al [2] studied the visual influence on auditory cortical neurons and their connection with other areas including: visual cortex, parietal cortex and the thalamus. They found evidence to support the involvement of the visual cortex and outside of the visual cortex with the auditory cortex - in a multi-sensory integration. Furthermore, this communication between visual and auditory cortexes involved excitatory as much as inhibitory mechanisms.

---

1. Hutchins, Edwin. "Cognitive ecology." *Topics in Cognitive Science* 2, no. 4 (2010): 705-715.

2. Town, Stephen, Katherine Wood, and Jennifer Bizley. "Multisensory integration in ferret auditory cortex: Effects of inactivating visual cortex." *Multisensory Research* 26 (2013): 221-221.

[i] Clark's Extended Mind hypothesis describes the cognition as a cycle that runs from the brain through the body. World and back. The mind extends into the organism's environment.

Before defining Cognitive Ecologies, Hutchins had previously developed the Distributed Cognition theory (1995) as a framework to describe, in computational terms, human work systems where knowledge lies not only within the individual but the individual's social and physical environment. The goal of the theory was to describe how distributed units are coordinated by analyzing the interactions between individuals, the media used, and the environment within which the activity takes place. Sutton and Tribble [3] apply this theory along with Andy Clark's Extended Mind to create their own framework for Shakespearean Studies. They too use the cognitive ecology approach to analyze theatre across a system of neural and psychological mechanisms, bodily and gestural norms, physical environments, cognitive artifacts (playbook), and technologies of sound and light; elements, that in their view, affect and modulate each other. In a theatrical cognitive ecology, attention and perception are not simply biological, but social and material systems of cognitive mechanisms shaped by social knowledge. In times of Shakespeare, audiences had an intuitive understanding of social conventions and hierarchy that would have been used in the economy of attention on stage. In addition, attention in theatre is also technologically mediated, not only by the vocal expertise of actors, but by sound effects and music. In Shakespeare's plays for example, voice, sound effects, and music are extensively used to command attention during the play. For instance, trumpets will announce the entrance of an important character to the scene. This not only calls for the audience's attention, but also gives them a cue as to the importance of the role of the character. All these sounds carry information that could be verbal or non-verbal.

One could say that Sutton and Tribble's cognitive ecology of attention and perception could be completed using Gibsons' Ecological Psychology. This theory bases perception on information rather than sensation.

---

3. Sutton, John, and Evelyn Tribble. "Cognitive Ecology as a Framework for Shakespearean Studies."Shakespeare Studies 39:94-103 (2011)

Here, information is detected as perceivable opportunities for action in the environment. For example, Pauletto [4] in her research designed a methodology to use sonic interaction to convey information. She borrows techniques from theatre to explore the different perspectives of the interacting subject, observer and sound designer, making them all part of a cognitive ecology, where one affects the others. In her methodology she describes the same kind of ecological approach from Sutton and Tribble, where sound psychoacoustics and perception allow the audience to distinguish between a variety of sounds and how to interact with them.

In literature the closest theory to a cognitive ecology is the Reader-Response Criticism. The idea of a literary cognitive ecology is not to be confused with Ecocritisim, a revisionist movement that examines literature through its treatment on the subject of nature. Being that clarified, the reader-response criticism is a literary theory that began in the 1960's, and considers that the reader completes the meaning of a text through interpretation. It views literature as a performing art in which the reader creates her own text-related performance. Amongst the different views of reader-response theorists, Individualists focus only on the reader's experience, while Uniformists assume that text and reader have a shared responsibility to convey the meaning. This makes reading both, subjective and objective. One must look into reading processes to create meaning, and experience to understand the text. Here there are two levels of understanding: the information explicitly presented in the text, and the integration of the different pieces of information from that same text. The Reader-Response Criticism is based on Kintsch's[5] construction-integration model. This general theory contends that comprehension arises from an interaction between the to be comprehended text, and the general knowledge and experience the reader brings to the text.

4. Pauletto, Sandra. "Film and theatre-based approaches for sonic interaction design." *Digital Creativity* 25, no. 1 (2014): 15-26.

5. Kintsch, Walter. "The role of knowledge in discourse comprehension: a construction-integration model." *Psychological review* 95, no. 2 (1988): 163.

Contrary to the Reader-Response Criticism are the text-oriented schools. For example Formalism, which claims readers can understand texts while remaining immune to their own culture. However, Kunze et al. [6] in their psychophysiological study measured the level of engagement, through readings of the nose temperature, heart rate, eye gaze and blinking rate. They used boring texts that included Victorian novels and appealing texts of coming of age novels to measure such engagement. They found that in fact, the novels that embody old fashion Victorian didacticism proved to be frustrating to contemporary readers. Set in a historical present, period literature turned out to be problematic for participants in Kunze's study since this literary genre conferred them with a simplified lesson from the past that has become meaningless with time. [7] In fact, studies based on the construction-integration model that follow Kintsch, stress the importance of background knowledge and the ability to generate inferences to fill in the gaps in the text, especially when the text is less cohesive. One of these studies is from Eason et. al [8] who looked into the predictors of specific text types: narrative and expository. Their study revealed that expository passages required a higher level of cognitive skills of inference and organizing, in addition to basic language skills needed in narrative texts. Their findings support a multidimensional scale of text complexity, and the interactions between these texts and the reader's cognitive skills. For example, Graesser and McNamara [9] find that in a multilevel comprehension framework the different levels of meaning constrain each other asymmetrically. The components at each level are successfully built if the text is considerate and the reader has a combination of enough background knowledge and reading skills.

---

6. Kunze, Kai, Susana Sanchez, Tilman Dingler, Olivier Augereau, Koichi Kise, Masahiko Inami, and Terada Tsutomu. "The augmented narrative: toward estimating reader engagement." In *Proceedings of the 6th Augmented Human International Conference*, pp. 163-164. ACM, 2015.

7. Spacks, Patricia Meyer. *Boredom: The literary history of a state of mind*. University of Chicago Press, 1995.

8. Eason, Sarah, Lindsay Goldberg, Katherine Young, Megan Geist, and Laurie Cutting. "Reader-Text Interactions: How Differential Text and Question Types Influence Cognitive Skills Needed for Reading Comprehension." *Journal of Educational Psychology* 104, no. 3 (2012): 515-528.

9. Graesser, Arthur C., and Danielle S. McNamara. "Computational analyses of multilevel discourse comprehension." *Topics in Cognitive Science* 3, no. 2 (2011): 371-398.

We come to the synthesis that all these happen in ecology of cognitive artifacts, neural and psychological mechanisms. In this way, in our view the reader-response criticism was constructed, maybe not intentionally, as a cognitive ecology. This literary criticism sees the text affects the reader, as much as the reader affects the texts; with inference based on background knowledge. This might be why the reading experience can become subjective, since the reader ads meaning to the written word in a cognitive ecology that where the economy of attention has not been technologically mediated.

Neuroscience in Literature

Cognitive neuroscience uses theories from cognitive psychology, to deal with linguistics, and findings in neuroscience as evidence of the brain's activity. These mental processes allow the written world to become the storyteller, and readers to understand fictional characters and their emotions, by grounding them to lived experiences. Cleary et. al.[10] found that these lived experiences are part of an evolved adaptive memory, in which familiarity helps for a rapid cognition, even when minimal information is given. This emotional capacity that allows readers to perceive the feelings of fictional characters, also known as empathy, has been proposed to have two dimensions: mirroring and self projection. Corradini and Antonietti [11] suggest that mirroring, allows to physically resonate what others are experiencing, whereas self projection implies imagining what should be felt, and then attributing those imagined experiences to others. Mirroring allows an intuitive and immediate comprehension of actions, and self-projection uses an inferential process to reason intentions into emotions.

---

10. Cleary, Anne M., Anthony J. Ryals, and Jason S. Nomi. "Intuitively detecting what is hidden within a visual mask: Familiar–novel discrimination and threat detection for unidentified stimuli." *Memory and cognition* 41, no. 7: 989-999 (2013)

11.Corradini, A., and Antonietti, A. "Mirror neurons and their function in cognitively understood empathy." *Consciousness and cognition* (2013).

Philosophy and Literature 36, no. 1 : 122-135. (2012)

In accordance, Bal and Martjin [12] consider the first dimension, mirroring, as a simulation exercise where the reader physically replicates actions and feelings of characters by emotional resonance. As an illustration, Helms [13] found that in Shakespeare's *Twelfth Night*, to gain knowledge of Orsino's mind, the emotional resonance of love and rage puts the reader in the "shoes of" Orsino, making it possible to understand the character's temperamental and volatile mind.

Moreover, Perlovsky and Roman [14] found that the mirror neurons, located in the premotor cortex of the brain's frontal lobe, create neural representations that are mathematically model as a pair of language representations by connecting perceptual-cognitive brain areas, oro-laryngeal movement control, and the Broca's area, to construct mental representations of what is being read. This emotional resonance is embodied even by fictional characters who by feeling make readers experience their emotions. For instance, in *Twelfth Night,* the reader understand Orsino's suffering for Olivia when he pays the Clown for his song about un-responded love saying: "There's for thy pains". Orsino mirrors the song lyrics, which in turn making the reader feel in a lower degree, the affliction of not being loved back. On the other hand, in Corradini and Antonietti's second dimension of empathy, imagining what the character is experiencing also allows the reader to attribute emotions to that same character. Orsino is regarded as a real person in the reader's imagination, making it possible to build concepts about his conflicted fictional mind: thoughts, feelings and desires. This means that to understand Orsino's position, the reader creates her own process of Orsino's mental state such as longing and fear of rejection from Olivia, by holding all the competing emotional states that reveal his conflicted human nature.

12.Bal, P. Matthijs, and Martijn Veltkamp. "How Does Fiction Reading Influence Empathy? An Experimental Investigation on the Role of Emotional Transportation." *PLOS ONE* 8, no. 1: e55341. (2013)

13. Helms, N.R. "Conceiving Ambiguity: Dynamic Mindreading in Shakespeare's Twelfth Night."

14. Perlovsky, Leonid I., and Roman Ilin. "Mirror neurons, language, and embodied cognition." *Neural networks.* (2013)

In this process of understanding, the reader uses a representation of herself in that same situation to perceive Orsino's complex mind, allowing the character to go from melancholic lover to a threatening character capable of murdering Viola. This dimension allows Orsino's character to dramatically turn his temper and actions, yet maintaining coherence in his personality, allowing readers to understand the reasons for Orsino's mood changes. Overall, gaining an understanding of Orsino's emotions, depends on cognitive processes that make use of mirroring and self projections of read actions -in a first person perspective- in what Sturber first called re-enactive empathy.[15] Sturber, found that empathy simulates the capacities of rational agents by replicating their actions and feelings in a lower degree, to understand and be immerse into a literary text.

2.2 Orality & Literacy

Long before a reader-response criticism theory could even materialize, storytelling had to transition from orality to literacy. However, this transition turned out to be long and painful. It has involved a physiological adjustment, since reading is not an inherited ability, but a new and very complex acquired skill. So much so that Ashby and Rayner[16] went on to investigate on the complexity of reading tasks for comprehension by studying the brain. They found the left hemisphere of the brain forms a network of circuits for word recognition in a triangle model of orthographic, phonological and semantic processes. But this written narrative did not go straight form speech to print. The transition has been through a gradual, but continuous physiological adjustment. Havelock [ii] thinks this transition might have started in Attic theatre, which in itself arose from the need of Athenians in the sixth century, to discover their own identity as a single city-state.

15.Stueber, Karsten R. "Varieties of Empathy, Neuroscience and the Narrativist Challenge to the Contemporary Theory of Mind Debate." *Emotion Review* 4, no. 1: 55-63. (2012)

16. Ashby, J., and Rayner, K. . Reading in alphabetic writing systems: evidence from cognitive neuroscience. *Neuroscience in Education: The good, the bad, and the ugly*, 61. (2012)

[ii] Havelock's thesis is that Western thought, and the ideas available to the human kind come from the point when Greek philosophy transitioned from orality to literacy. Along with Walter Ong, Havelock founded a new field to study this transition, which in turn has influenced media theorists such as Marshall McLuhan.

Attic theatre was a supplement for Homer, whose epics furnished the Greek identity, moral, politics and history. But Homer's tales were Ionic and Pan-Hellenic, and not part of the native dialect with which to address the new Athenians. However, Attic theatre gives birth to authors who are more producers than writers. They composed their vision, in the tension of oral and written communication, by dictating to a literary assistant and hearing it back from this assistant to edit, retaining only the ear to compose. Havelock suggests the secret for the brilliance of Attic theatre is due to the tension caused by the transition from orality and literacy, which has not been repeated in history. Until this moment narratives had been shaped only through the mouth, but with a new alphabet the Greek language transferred the inherited habits of mouth and ear to the acquired habits of hand and eye. Writing and reading required physiological adjustments to master the new skills. An adjustment that was slow and painful. [17]

The differences between oral cultures and literate cultures affect on how meaning is shared. Written language conveys a monothetic meaning which semantic component makes it time transcendent, sign-oriented and conceptual.[18] In contrast, theatrical performances are a form of communication that do not completely rely on idealisable or objectifiable semantic content. Folkerth [19] finds that in plays meaning is time-immanent, perceived in fullness in an embodied experience of the moment. Attic theatre was composed in the tension of oral and written communications but retained the ear to edit what had been but down in paper. In this way theatre allows to experience in a process within the embodied consciousness, and not as an object of consciousness. Furthermore, plays offer us a monothetic meaning but still ask for a polythetic experience. This is only possible because of the gradual and continuous transition of residual oral cultures into literacy.

---

17. Havelock, Eric A. "The oral composition of Greek drama." *Quaderni Urbinati di Cultura Classica* (1980): 61-113.

18. Schütz, A. "Making music together: A study in social relationship." *Social research* (1951): 76-97.

19. Folkerth, Wes. *The sound of Shakespeare*. Routledge, 2014.

This transition can be seen in Shakespeare's work. His plays were meant to be heard, but print has made it necessarily to edit these texts causing them to lose its aural nature.[20] Generations of editors have added layers of silent emendations, holding meaning through grammatical punctuation, with commas and periods that set off clauses for the eye. These editorial revisions make Shakespeare more readable, since print is sophisticated and precise, but takes away the rhetorical and auditory punctuation, disconnecting our embodied experience. [21]

Radio drama is, in essence, a theatrical representation supported only by sound to convey polythetic meaning. Without any visual component radio drama holds its meaning in the auditory dimension depending only in a written dialog, music, and sound effects, to deliver enough information for the listeners imagination. To create a more potent listening atmosphere radio drama uses sound effects to round out the dialog, filling the absence of visual cues that convey meaning. Sound effects allow listeners to become part of the intricate moments of the play. For example, a character that delivers a slap across the face. Weaver [22] calls this medium Theatre of the Mind where imagery has no limitation. In this auditory experience there is an absence of reason and logic. Listeners can lose structured thoughts and invite ideals that cannot be explained. Weaver exemplifies with the events taken in 1938, when the Theatre of the Mind caused hysteria amongst listeners who tuned in Orson Welle's radio play *War of the Worlds*. Thousands of people ran screaming into the streets, made frantic phone calls, took their belongings and fled; the Martians were invading. Listeners forgot they were tuning in to hear their regular program, in their minds the radio play adaptation of Welle's novel was a real event. This imagery is also referred as listening in the mind, an aural visualization the brain builds when there is no picture to stimulate the mind It is more appropriately to call it the Theatre of the Ear.

---

20. McLuhan, Marshall. "The Gutenberg Galaxy." *Toronto, UTP* (1963).

21. MacLuhan, Marshall. *The Effect of the Printed Book on Language in the 16th Century*. Gingko Press, 2005.

22. Weaver, Pascha. "RADIO DRAMA: A "VISUAL SOUND" ANALYSIS OF JOHN, GEORGE AND DREW BABY." PhD diss., University of Central Florida Orlando, Florida, 2012.8

Regardless of the advantages of sonic non-verbal cues in radio dramas, Audiobooks have only retained the printed word, in a monothetic oral storytelling. But not all components of orality are lost, a thoughtful lively performance of the narrator can sometimes work better than the printed word. As an example, Shokoff [23] found Ken Follett's *Pillars of the Earth* tedious and heavy-handed in reading. But listening to the audiobook abridgement narrated by the British stage and TV actor Tim Pigott-Smith revealed the fullness of the literary text. In fact, listening to what is being read can help comprehension. Turkylmaz et al. [24] looked into the differences of oral, silent and retell reading fluency. Fluency refers to the ability to process the meaning of the text for a deeper lever of understanding, since it is a reflection of reading comprehension. In their results they found that oral reading fluency was the strongest predictor of reading comprehension compared to the other two. Young-Suk et al. [25] also looked into fluency between oral and silent reading and went further to find how they relate to reading comprehension. In their results they found that oral and silent reading fluency tasks measure different but highly related underlying skills. Between the two, oral reading fluency was a better predictor of reading comprehension. Yet, their most striking finding was that while decoding fluency was a more important predictor for reading comprehension in average readers, in both silent and oral, for skilled readers was listening comprehension. Even though decoding fluency constrains performance in average readers, reduces its constraining role in skilled readers with superior decoding fluency. This allows skilled readers to use oral language skills -listening comprehension- to play a greater role in reading comprehension.

---

23. Shokoff, James. "What is an audiobook?." *The journal of popular culture* 34, no. 4 (2001): 171-181.

24. Turkyılmaz, Mustafa, Remzi Can, Kasim Yildirim, and Seyit Ateş. "Relations among Oral Reading Fluency, Silent Reading Fluency, Retell Fluency, and Reading Comprehension." *Procedia-Social and Behavioral Sciences* 116 (2014): 4030-4034.

25. Kim, Young-Suk, Richard K. Wagner, and Elizabeth Foster. "Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers." *Scientific Studies of Reading* 15, no. 4 (2011): 338-362.

Literature makes use of visual aids for comprehension, just as radio dramas use sound effects and music to fill the absence of visual cues of information, print uses illustrations. For instance Jalilehvand [26] applied the Dual Coding Theory as a rationale to enhance information processing through two different visual channels: pictures and text. She looked into the effects of text length and images on reading comprehension. In a sample of seventy-nine first grade female students she applied four conditions: long text, long-text and images, short text, and short text and images. Jalilehvand found that in a reading comprehension test students performed better on the reading comprehension test in both condition of mixed text and image, making these images a key variable for comprehension.

In this way we argue that literature uses visual extra-textual information in the forms of illustrations, while theatre and radio dramas use sound to deliver these same extra-textual information. While in orality the dual coding has used verbal and non-verbal sonic cues, such as speech and sound effects, literacy has focused on visual cues of writing and illustrations. Even thought in our view both forms of storytelling, oral and literate, have made use of the dual coding, this has been done utilizing in only one sense: the auditory or the visual. Thus, there seems to be an unexplored application to use the potential of sonic information mixed with the intrinsic meaning of writing, which is to say the meaning that the text conveys itself. In this application, it is our presupposition that the sonic cues will fill in information gaps in the text. Here, sound can become part of a new cognitive ecology of literary texts to create meaning by stimulating the reader's imagination, as done in theatre. In our view it will allow for a technologically mediated sonic economy of attention where the text interacts with its reader through the spontaneity of sound. In this ecology there is an interaction where all agents understand and affect each other.

26. Jalilehvand, M.(2012) "The Effects of Text Length and Picture on Reading Comprehension." *Asian Social Science* 8 (3): 329.

The Quality of Sound

The perception of a literary text is purely visual, where the reader depicts the immediate external stimuli to the organism as neural representations, and relates it to internal portrayals encoded in our memory. This allows for a descriptions of past, future, and hypothetical environments, such as fictional worlds, in a top-down process.[27] 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. For example, in oral cultures, sound has been present even in the earliest human theatrical events such as primitive harvest festivals that were accompanied by drums, to represent and help relate with the environment.[28] During the Bronze Age theatre was also accompanied by sound, much more so than scenery or props. Later on Commedia dell'arte applied sound effects to actions, which in Shakespeare's plays became an indispensable element to the production.

In Elizabethan plays not all sights of the action could be portrayed on stage, and even though some actions could not be seen by the audience they all could be heard, making these sounds an important part of the storytelling. These off stage sounds were so important that packs of hounds and soldiers with their full arsenal were hired to produce authentic noises on cue, allowing for a multimodal perception of the play being portrayed. Specific cues, such as sound effects or music fall into different dimensions, all of which need to be informational in nature to support of-the-moment events.[29] For example, Houix et al.[30] investigated sonic information in the perceptual relevance of the state of mater: liquid solid or aerodynamic.

---

27. McNorgan, Chris. "A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery." *Frontiers in human neuroscience* 6 (2012).

28. Hug, Daniel. "Investigating narrative and performative sound design strategies for interactive commodities." *In Auditory Display*, pp. 12-40. Springer Berlin Heidelberg, (2010).

29. Kaye, Deena, and James LeBrecht. *Sound and Music for the Theatre: The Art and Technique of Design*. Taylor and Francis, 2013.

30. Houix H, Lemaitre G, Misdariis N, Susini P, Urdapilleta I (2012) A lexical analysis of environmental sound categories. J Exp Psychol Appl 18:52

Houix experimented with sixty sounds encountered in a kitchen context. In the study all participants appeared to formulate the same cognitive organization based on the mechanics of the sound source: machine and electric device sounds, liquid sounds and aerodynamic sounds. Moreover, Lemaitre and Heller[31] looked into the perception of state of matter in sound-generated human-made illusions; just like the sound effects found in theatre. They used friction sounds to represent solid matter or dripping sounds to denote liquid matter and discovered listeners could recognize the set of matter above-than-chance levels with an accuracy of seventy five percent. Subjects in Houix, Lemaitre and Heller studies associated sound to its source; just as audiences do with auditory cues in theatre. This is because sound has the ability to carry on non-verbal information. Sonic cues are able to structure perception by associating themselves to images and meanings, contributing for a significant experience.[32] Here sound is at the heart of interpretation, the combined stimuli across senses leads to improved detection and faster reaction as a result of the combination of these sonic complementary cues.[33] In fact, Giordano and Avanzini [34] looked on how to apply the informational value carried by sound. They believe Interactive Sound Design should be based in the study of ecological perception, where sound-synthesis techniques render the perceptual material properties and mechanical actions. These qualities of sound to convey meaning indicate that this channel is as effective as words are in a page, if not better. The deep relationship between sound and text is a clear indication that they can, and should, be used together in an interaction between author, text, sound, and reader.

---

31. Lemaitre G, Heller LM (2013) Evidence for a basic level in a taxonomy of everyday action sounds. Exp Brain Res 226(2):253–264

32. Verstraete, Pieter. "The frequency of imagination: auditory distress and aurality in contemporary music theatre." (2009).

33. King, Andrew J., and Gemma A. Calvert. "Multisensory integration: perceptual grouping by eye and ear." *Current Biology* 11, no. 8 (2001): R322-R325.

34. Giordano, Bruno L., and Federico Avanzini. “Perception and Synthesis of Sound-Generationg Materials”. In *Multisensory Softness*, pp. 49-84. Springer London, 2014.

2.3 Interactive Storytelling

Based on the relation between mind and body, computers are capable to understanding emotional states. Consequently these computers can take actions by recognizing likes and dislikes of users. In this way they evolve from mere tools to personal companions. Human emotions improve communication between people and computers, in the same way speech and gestures are used in verbal communications. Sensors give logic and reason to mechanical devises making them able to empathize with their users, by looking for positive and negative reactions to a task. For example, Kuo and Chih-Chi,[35] using theories in affect-sensitive cognitive machines,[36] developed an intelligent agent (IA) that recognizes the desire to learn. Based on the limited performance of learners that cause negative emotions, this IA uses empathy to interact with Chinese students. The system assists these students, who want to improve their English listening comprehension skills, by adapting listening comprehension exercises according to the individual performance. It uses a galvanic-skin-response (GSR) sensor to recognize satisfaction, disappointment, frustration and hopefulness. At the beginning of the listening exercise, the learner is in quadrant one: satisfaction; high arousal positive, valence meaning he is learning. If the system detects that his emotions pass to quadrant two: disappointment; high arousal, negative valence, it takes that the learner needs assistance to overcome the problem and releases a transcript of the listening exercise in English subtitles. If this does not help the student, and he passes to quadrant four: frustration: low arousal, negative valence, it then provides a transcript of the listening comprehension exercise in Chinese subtitles in addition to readjusting the level in the next listening exercise to fit the student. However, it is important to point out that Kuo and Chih-Chi created an interaction by detecting emotions, and thus the GSR data was useful.

---

35. Kuo, Yu-Chen, and Chih-Chi Tseng. "Recognizing the emotion of learners by physiological sensors to improve English learning performance." In *Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on*, vol. 4, pp. 2152-2156. IEEE, (2011).

36. Kort, Barry, and Rob Reilly. "Theories for deep change in affect-sensitive cognitive machines: A constructivist model." *Journal of Educational Technology and Society* 5, no. 4 (2002): 56-63.

Moreover, the use of emotional cues to change content has also been applied in digital entertainment to allow interaction in narratives. For instance, Interactive Storytelling captures the user's emotions to create an affect-based interactive narrative. In his research Zhao [37] indicates that in general, affective gaming creates new experiences by adapting the game to the player's emotions in two modalities: managing the story-line to achieve the author's emotional goal or adapting the narrative only to generate positive emotions. However this is done without considering the author's work, since the narrative is left to evolve as the user sees fit. For example, PINTER[38] creates an affective and empathetic experience while watching an animation of the medical drama *Arztserie*. In PINTER, the viewer's emotions act as input telling the system how to unfold the story. By balancing the potential empathic relationship between the user and the characters PINTER changes the narrative in accordance to those emotions.

IS can also be dependent on engagement levels. For example FOCUS is an EEG augmented reading system that uses a commercial Brain Computer Interface (BCI)- Emotiv- and a projector to interact with the book being read. By using the BCI, FOCUS can tell if the reader has a low level of engagement while reading. If so, it triggers the BCI training mode, where the projector displays a related animation within the designated white area of the book. This causes the reader to concentrate on the animation increasing the EEG-engagement level, helping the student with his attention deficit. Overall these three examples: FOCUS, PINTER, and the listening comprehension learning system, serve to show the potential of different sensors in creating empathetic systems capable to adapt to their users by delivering personalized experiences in texts, animations and audio.

37. Zhao, Huiwen. "Emotion in Interactive Storytelling." PhD diss., Communication University of China, 2013.

38. Gilroy, Stephen, Julie Porteous, Fred Charles, and Marc Cavazza. "Exploring passive user interaction for adaptive narratives." In Proceedings of the 2012 ACM international conference on Intelligent User Interfaces, pp. 119-128. ACM, (2012)

However we find print is only comfortable with finality, with it's tight plotting that extends into lengthy narratives with a sense of closure; and its fixed point of view to provide a shared understanding between author and reader. Thus, rather than adapting the narrative to balance the reader's emotions, as done in Interactive Storytelling, the interaction can focus into keeping the reader engaged. But this engagement should be about the reading experience, in an interaction that respects the essence of the novel, and at the same time allow the medium to improve as a storyteller. To do this we believe literary criticism should be brought to the field of human computer interaction.

Cognitive and Emotional Metrics

Cognitive processes and emotional states are linked to a physical response caused by the Autonomic Nervous System. For instance, eye blinking is a psychophysiological response of the central nervous system that has been found to be in parallel with both visual and auditory emotional stimuli, such as reading and listening. [39] Guo et. al.[40] used an eye-tracker to measure emotional reactions and their connection to attention. Guo's study was based on Mahlke and Minge [41] findings that suggest subjective feelings significantly relate to cognitive assessment. Guo's results indicate that a moderate arousal is related to positive valence, while a lower arousal would be related to boredom. Moreover the results indicate that the increase in blinking when task were unpleasant correspond to an approach-avoidance tendency.

In our daily, life different actions call for different amounts of attentional investment. For example, brushing your teeth only requires a small amount of attention, whereas reading a book will demand much more attention.

39. Klingner, Jeff, Barbara Tversky, and Pat Hanrahan. "Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks."*Psychophysiology* 48, no. 3 (2011): 323-332.

40. Guo, Fu, Yaqin Cao, Yi Ding, Weilin Liu, and Xuefeng Zhang. "A Multimodal Measurement Method of Users' Emotional Experiences Shopping Online."*Human Factors and Ergonomics in Manufacturing and Service Industries* (2014).

41. Mahlke, Sascha, and Michael Minge. "Consideration of multiple components of emotions in human-technology interaction." In *Affect and emotion in human-computer interaction*, pp. 51-62. Springer Berlin Heidelberg, 2008.

To detect attention levels, Chang and Huang[42] investigated the relationship between with psychophysiological responses, such as heart rate, heart rate variability and eye blinks. They based their study in the Feature Integration Theory, a psychological model of human visual attention that proposes the features of a stimulus are registered early automatically and in parallel, while objects are identified separately at a later stage. The increase of task demand has been found to decrease in the metrics of heart rate variability, eye blink and blink duration.[43] However Chang and Huang could not find any significant change in eye blinking to discriminate attentional levels. On the other hand, they found that an increase in attention levels decreased the strength of low frequency components, influenced by the sympathetic system, decreasing HRV.

The sympathetic system is also responsible of lowering the temperature of the nose in mental load tasks that can drop it up to 0.55 ^{o}C below baseline. The temperature decrease also signed of distress and shifts in attention. [44] In fact, task engagement and distress have been found to take part of affective, cognitive and motivational domains. [45] On the other hand frustration increased the blood volume into supraorbital vessels, which in turn increases the skin temperature.[46] These attentional and affective states, such as immersion when reading a story, are characteristics of the Csíkzenmihályi's flow experience.[47]

---

42. Chang, Yu-Chieh, and Shwu-Lih Huang. "The influence of attention levels on psychophysiological responses." International Journal of Psychophysiology 86, no. 1 (2012): 39-47.

43. Fairclough, Stephen H., Louise Venables, and Andrew Tattersall. "The influence of task demand and learning on the psychophysiological response."International Journal of Psychophysiology 56, no. 2 (2005): 171-184.

44. Fournier, Lisa R., Glenn F. Wilson, and Carolyne R. Swain. "Electrophysiological, behavioral, and subjective indexes of workload when performing multiple tasks: manipulations of task difficulty and training."International Journal of Psychophysiology 31, no. 2 (1999): 129-145.

45. Matthews, Gerald, Sian E. Campbell, Shona Falconer, Lucy A. Joyner, Jane Huggins, Kirby Gilliland, Rebecca Grier, and Joel S. Warm. "Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry." *Emotion* 2, no. 4 (2002): 315.

46. Ioannou, Stephanos, Vittorio Gallese, and Arcangelo Merla. "Thermal infrared imaging in psychophysiology: potentialities and limits." *Psychophysiology* 51, no. 10 (2014): 951-963.

47. Csikszentmihalyi, Mihaly, and Mihaly Csikzentmihaly. *Flow: The psychology of optimal experience*. Vol. 41. New York: HarperPerennial, 1991.

For instance, the state of flow produced by listening to the story of *The Ugly Duckling* by Hans Christian Andersen relates to the absence of demanding tasks in emotional processing, where the reader is immersed in the story. In *The Ugly Duckling* study Wallentin[48] looked into the continuous intensity and emotional variation form positive to negative, comparing measurements of the heart with those of the brain. He found an increased response in the heart rate variability (HRV) during the intense parts of the narrative. The responses had a correspondence with a bold response in the brain regions that contain emotional reactions from the auditory stimuli. In his study he looked for an emotional response but unveiled a mix of emotional and attentional response.

In Johannes Keller's study the mental load was affected by involvement, the core of the flow theory, during which information-processing increments the workload.[49] Keller looked into the impact of flow in conditions of boredom, fit, and overload while completing a knowledge task using questions from the game based on the TV show "Who wants to be a millionaire?" In his results, he reveals that the highest HRV was found in boredom, reflecting the lowest mental load, while a decreased reflected involvement in the fit condition showing that higher levels of flow can be associated with a decreased HRV, contrary to the increment found in emotional responses, since it's attention provoking and not an emotional response. His hypothesis is that the physiological response of coherence is only a correlation of flow. Unfortunately, due to their methodology which included three different tasks between two games, one for flow, one for coherence, with a final questionnaire, they were unable to find any correlation between flow and coherence in a percentage of time, even though their literature, which included Keller's study, for them suggested otherwise.

48.Wallentin, Mikkel, Andreas Højlund Nielsen, Peter Vuust, Anders Dohn, Andreas Roepstorff, and Torben Ellegaard Lund. "Amygdala and heart rate variability responses from listening to emotionally intense parts of a story."*Neuroimage* 58, no. 3 : 963-973. (2011)

49. Keller, Johannes, Herbert Bless, Frederik Blomann, and Dieter Kleinböhl. "Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology* 47, no. 4 : 849-852. (2011)

However, in the James-Lange Theory emotional response is caused by an arousing experience where the sympathetic nervous system increases the number of heartbeats per minute in the HR, whereas as Keller found, in the absence of arousal the valence of different conditions of involvement, such as engagement boredom and overflow, can be detected in the variations between these heartbeats, in the HRV time interval referred as the QRS complex. Here, information of involvement, determined by a positive valence, can be detected in the ratio of low frequency energy to high frequency energy that represents the extent of sympathetic and parasympathetic influence on the heart. To analyze the heart data a Short Time Fast Fourier Transform (STFT) [50] can be applied in a floating window of N beats, applied to every new Interheart Beat Interval (IBI) value, using Power Spectral Density (PSD) to represent the amount of power per unit of frequency to extract three different frequencies to obtain a usable HRV signal. The HRV has three frequency peaks: low under 0.1Hz, medium near 0.1Hz and high between 0.15Hz-0.6Hz sampled at a rate between 250Hz to 500-1000Hz to capture the R wave peaks. In these frequencies the ratio of low frequency energy to high frequency energy represents the extent of sympathetic and parasympathetic influence on the heart. The HRV signal can then map the psychological measures with the Psycho Physiological Emotional Map (PPEM) by linking the stimuli-provoked data from a first and third person approach. McCraty [51] proved the used of PPEM defining the emotions of anger and appreciation, with which Pramila [52] implemented it in a Support Vector Machine (SVM), pioneered by Vapnik in 1998, to identify boredom and engagement (See Table 2.1).

---

50. Villon, Olivier, and Christine Lisetti. "A user-modeling approach to build user's psycho-physiological maps of emotions using bio-sensors." In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, pp. 269-276. IEEE, 2006.

51. McCraty, Rollin, Mike Atkinson, William A. Tiller, Glen Rein, and Alan D. Watkins. "The effects of emotions on short-term power spectrum analysis of heart rate variability." *The American journal of cardiology* 76, no. 14 (1995): 1089-1093

52. Rani, Pramila, Changchun Liu, Nilanjan Sarkar, and Eric Vanman. "An empirical study of machine learning techniques for affect recognition in human–robot interaction." *Pattern Analysis and Applications* 9, no. 1 (2006): 58-69.

SVM allows the use of linear algebra and geometry to separate data like with nonlinear rules in the input space. This is important for systems that require non-linear methods to effectively analyze the continuous signal of the HRV captured in frequencies using the Radial Basis Function kernel (RBF) to satisfy Mercer conditions. Here the systems learning task is reduced to a dual quadratic programming problem with Lagrange multipliers to give efficient approach to optimization. The downside is that the SVM separates the data only onto two classes. The efficiency of SVM in real-time applications for the learning algorithms employed in speed can be expected to go faster when using a large data sets and high number of features and is space efficient.

Table 2.1: (Rani Pramila et al. 2006) Sentence Verification Task Results

Anxiety	**Boredom**	**Engagement**	**Frustration**	**Anger**
88.86%	84.23%	84.41%	82.82%	88.74%

Furthermore McCraty[53] also developed a typology of six distinct patterns of the HRV to denote the different modes of psychophysiological interaction. They used a circumflex model as a conceptual map to show the relationship between emotional experiences and the different HRV wave patterns related to each one of them. In their graphic they only describe the wave forms as emotional states, divided into normal every day life emotional experiences and hyper-states of emotional experiences. Unlike Keller McCraty et al. looked for emotional rather than attentional metrics of the HRV. However if we overlay the Kellers findings using the axis of arousal and balance with McCraty's circumflex model we can find that attention corresponds to what McCraty calls 'serenity' and boredom with apathy. Even though both scholars look for different cognitive processes in the HRV, their results show correspondence with wave patterns that link emotional and attentional metrics of the heart (See Figure 2.1).

---

53. McCraty, Rollin, Mike Atkinson, Dana Tomasino, and Raymond Trevor Bradley. "The coherent heart: heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order." *Integral review* 5, no. 2 (2009): 10-115.

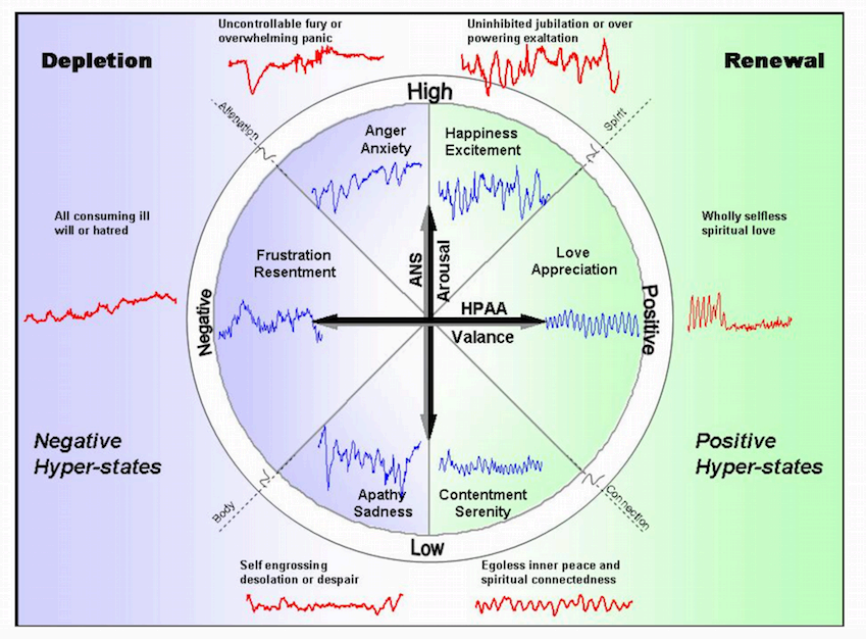


Figure 2.2: (Rolling McCraty et al. 2009, p.29) Typology of the HRV

2.6 Overview

As seen in this literature review, storytelling has its origins in oral communications, becoming literature with the invention of the alphabet and technologies of pen and paper. However, writing and reading have been seen as natural skills, ignoring the achievements of the brain that allow us to code and decode language. Literature has also made narratives to be experienced as a solitary and silent. Editorial processes have left out any stimulation other than the visual, that is, with few exceptions of authors that in their writing intended for an aural reading. This is why we pinpoint theatre as a middle ground between orality and literacy. In theatre we perceive with eyes and ears, and are encouraged to imagine places and actions that cannot be portrayed in stage, by using sound effects. On that account, this literature review presents sound as an important component of literature, even when we have silenced this medium. Sound can bring extra information to create a personal interpretation and performance of the text.

As readers empathize with fictional characters in the story through mirror neurons and self-projection, print has made it impossible for it to be the other way around, as print cannot empathize with the reader. Interactive Storytelling has looked into interactions between narratives and their audience using data from the heart, skin temperature and eye-activity. However, the use of sensors to promote these interactions have not yet been applied in a classic approach to literature; where the linearity of the story and author's set point of view cannot be changed, due to its origins in print. Print sets itself apart from episodic structures of oral storytelling by developing into linear texts that represent the word of the author in final form. This linearity is the core of the novel, and thus essential to sustain this literary genre.

In order to challenge the silence of print and its lack of interactivity, we introduced works that make use of psychophysiological metrics that allow narratives to interact with their users. In these metrics there is different information: emotional, cognitive and physical. However this review concentrates in the distinction of cognitive and emotional physiological metrics, which is somehow blurred. Engagement in literature can be regarded as emotional, where readers understand and share the feelings of fictional characters such as Orsino. However, attached to this emotional immersion there is a cognitive process of attention, that even though is very closely associated to an emotional immersion is a cognitive process. We discuss these differences with the works of Johannes Keller in the metrics of the HR to detect emotion; and HRV to detect mental workload.

Our synthesis of the presented literature review is that, while works in Interactive Storytelling have looked into emotional data to create interaction, the biofeedback for a reading-interaction should focus on cognitive processes. This is because the reader crates meaning not just from what she extracts from the text but by adding her own experience. Therefore controlling the emotional reaction of every reader is not feasible; since readers should be free to interpret the text according to their own experience. Looking into measurements of attention and engagement through mental workload means the reader is free to

create her own interpretation of the text, without modifying it. In reciprocity the text can interact with the reader without forcing an emotional reaction, creating a balance in the shared meaning. Sound, with its many properties, can be used to create these interactions using it as an intervention to assist when there is not enough shared understanding between author and reader. In this way we put forward the idea that the medium can be re-deigned to be an intelligent cognitive agent capable to understand the reading experience. Create a book with an intelligent interface that looks for attention, and assists the reader to engage in the narrative. This interaction depends on the capacity of the text to measure attention and comprehension through physiological metrics that do not

Chapter 4

Sonic Interventions

We are born with an innate ability to tell apart sounds that conform a language from other sounds, and seem to be pre-programmed to talk without formal instruction. However, this is not the case with reading. This artificial accomplishment comes from the brains capacity to call upon neural regions designed for other tasks, and putting them together in order to decode written texts that express our thoughts and emotions. Reading is no easy task, but it allows us to extend the mind into texts, where we have achieved higher levels of though construction. In reading, there is a reflective thought, an authorial omniscience that places the non-natural mind of fictional characters into a natural frame, allowing authors like James Joyce and Toni Morrison to provide the reader with the artificial effect of experiencing and viewing outside the natural world. However, in narratological studies there is a consensus that orality, which is in a close connection with the natural world, has nothing in common with literature. This field of study, considers the experience of the natural world is detached from the omniscience of the written word.[1] By doing so, we contend narratology has failed as a comprehensive discipline. On that account, it is of our interest to bring forward this discussion for examination, by presenting the literature that integrates sounds. With this multi-sensory literature it is our purpose to express the importance of the connection between orality and literacy, which in our view, could bring better literary storytelling to a society where the interaction with gadgets is fundamental. Seeing that if the medium influences how the message is perceived, then the interaction with that medium can define how that same message is understood.

---

1. Fludernik, Monika. *Towards a'natural'narratology*. London: Routledge, 1996.

Shakespeare, who made a profound impact in the English language through pen and paper by creating new words, and refashioning oral expression, inspired this new concept. In the same way that his new words stimulate the reader, calling for her attention, the proposed system translates this grammatical stimulation into interventions of sound, leaving the work of the author intact, but creating the same boost into the reader's mind. Similar to actors in a play who change the written word to captivate their audiences, the proposed literary storytelling stimulates the reader by adding concepts of orality that allow for an interaction. The goal of this interaction is to bring the spontaneity of oral communications into literature. Thus, we put forward the idea that literature can be reshaped by orality, as much as orality has been reshaped by literature.

This empathetic literature interacts with the reader by delivering sonic cues. For instance in theatre not all can be portrayed on stage, especially if the budget is tight. This leaves the task to the audiences' imagination to fill in the context and setting where the actor is delivering his performance. But not all is left to the audience, directors and actors will work with sound designers to provide them with sound effects that can trigger the imagination and fill in gaps of a limited stage design. The sound designer can create the sound of a car and it's horn to create an image that there is a street on the other side of the stage, an imaginary street with which the actor can interact. This ability of sound to create the notion of something that cannot be seen, is then reinterpreted to fit literary narratives where readers -that perceive the verbal information of the story, not by the actor's performance, but by their own performance- can use this non-verbal sonic information of sound effects and soundscapes to precipitate and increase mental imagery. On the other hand, the dramatization of the actor also creates context to the play, using intonation and rhythm, a change in the voice tone to express emotions and situations. This performance is done by the reader in literature, without the direction or aid from a director to correctly present the story. Thus, theatre gives a deep insight on how sound is used to create a meaningful storytelling, helping audiences to immerse into the story. This is why we see the potential of sound as beneficial to literature, and propose to use this sonic agency as interventions to assist the reader in her own performance of the text.

4.1 Prior Knowledge

The proposed sonic interventions are based in our hypothesis that when the reader cannot imagine the story being read, this creates frustration, drifting attention and ultimately negatively affecting reading comprehension. In this sense we propose sound designed to assist the average reader. The sonic agency is proposed to be use by the general readers, who we believe can benefit from the interventions of sound to undergo an enhanced reading experience that will increase the number of books an individual reads per year. We come forward with this statement based on our own data, where we studied the difference in immersion between Victorian novels and contemporary teen fiction.

From our study in estimating the reader's engagement, discussed in chapter three we take the data analysis further and found the applied immersion questionnaire were useful, not only to find correspondence between the physiological data and the subjects' feedback, but also served to show a correlation between frustration and mental imagery (See Table 4.1). In the immersion questionnaire each subject gave a grade from one to five when asked if the text was frustrating and if they were able to imagine each of the stories.

Table 4.1: Immersion Questionnaire (Refer to study in 3.4)

	Frustration	**Mental Imagery**	**Comprehension**
Silas Marner	16 pts	19 pts	15 pts
Sir Charles Grandison	23 pts	14 pts	13 pts
Celebs in Search of a Wife	23 pts	10 puts	11 pts

We argue that most readers would not have much prior knowledge of Victorian life, and therefore these three novels would be harder to mentally simulate. However, while in all three books, religion is a constant, in *Silas Marner,* George Elliot gives us more detail of context. Elliot's books have been praised for the detail and importance she gives to her readers about the ordinary country life; drawn from her own experience. Jonah Lehrer, in his book *Proust was a Neuroscientist,* tries to describe Elliot for what and who she was.

Elliot considered her novels as a set of experiments in life, a form of research on the capabilities of human emotions and thought. Her writing process was in fact a scientific process, with a blend of empiricism and imagination. In her novels she looked to answer epic questions. In *Silas Marner* for instance, her concerns on the existence of free will. For Elliot writing was an endeavor to give a vision of ourselves. She acknowledged the human malleability and the freedom to change, and believed that the ability of the mind to alter itself was the source of freedom. Elliot herself changed her mind about religion and positivism.

As Elliot, Hanna More was an extraordinary woman of her time, a prominent campaigner against slavery, and most importantly a social reformer. However, her only novel is used as a platform for her political work, and not as a quest of life itself. In, *Coelebs, in Search of a Wife,* she is addressing the women of her time, educating them to understand their function in society, and instructing them to carry out a revolution in manners. In More's eyes, women are different from men -whose manners are coarse in a commercial humanism- and have the responsibility to balanced society with their female virtue. More and Elliot are important female writer that should be praised for their work, but the lack of prior knowledge amongst the subjects in our study prevented them to have a shared understanding with both writers.; with More's work showing least engagement in all subjects. We attribute this, to the lack of context within More's text. Contemporary readers are not experiencing life as readers in the early nineteen century were, the context is not obvious nor relevant; it has to be previously studied and researched.

We do not only base the assumption that prior knowledge or the lack of it influence mental imagery based only on our own data. Melanie Green [1] finds that factors that increase fluency of mental imagery are key for engagement. One of those factors is repetition, as repetition makes the reader familiar with the narrative increasing the understanding of the cognitive components. Readers that have at least a mental model constructed, due to memories or prior exposure to the text, have a better chance to be transported into the story.

Green describes transportation as the reader being in a state of flow in a combination of cognitive engagement, emotional engagement and mental imagery. In her study to measure the difference in transportation across media in a repeated exposure Melanie Green et. al. found that reading before watching the film for the same book increased transportation into the film, while watching the movie twice decreased transportation. Unfortunately she found no clear difference of transportation in repeated reading, but found transportation had no correlation with imagery . However, she did find that need for cognition was a factor in individuals that were more transported when reading. This has to do with the effort they put on imagery as print demands more mental simulation than film.

4.2 Dual-Coding

As visually done in printed books with text and illustrations, we find that the dual-coding theory can also be applied to the sense of hearing. For instance, radio drama uses verbal and non- verbal cues to tell a story. Without any visual component, radio drama holds meaning in the auditory channel depending only in a written dialog dramatized by actors, music, and sound effects, to deliver enough information for the listeners' imagination. To create a more potent listening atmosphere radio plays make use of non-verbal cues such as sound effects to round out the dialog, filling the absence of visual cues to convey the message. Sound effects allow listeners to become part of the intricate moments of the play. For example when one of the characters decides to deliver a slap across the face, sound can provide information about the intensity of the slap. For these reasons, this medium is also called the Theatre of the Mind, where imagery has no limitation.

---

1. Green, Melanie C., Sheryl Kass, Jana Carrey, Benjamin Herzig, Ryan Feeney, and John Sabini. "Transportation across media: Repeated exposure to print and film." Media Psychology 11:4 (2008): 512-539.

More importantly in this auditory experience, there is an absence of reason and logic, listeners can lose structured thoughts and invite ideals that cannot be explained. This can be perfectly exemplify with the events of 1938 where the Theatre of the Mind caused hysteria amongst listeners who tuned in Orson Welles radio play War of the Worlds. Thousands of people ran screaming into the streets, made frantic phone calls, took their belongings and fled as they thought the Martians were invading. Listeners forgot they were tuning in to hear their regular program, in their minds the radio play adaptation of Welles' novel was a real event. This imagery is also referred as listening in the mind, an aural visualization for the brain to build upon, when there is no picture to stimulate the mind; it is more appropriately the Theatre of the Ear.

Regardless of the advantages of non-verbal cues we find the dual coding has in radio dramas, Audiobooks have only retained language to provide sonic information. Without any extra-textual information, Audiobooks with a thoughtful lively performance of the narrator, can sometimes work better than the printed word. Voice can give a sense of context when listening the emotion the reader has in the voice intonation. However, not all audiobooks use dramatization, and without any other sonic cues, depend solely on the verbal performance provided by the narrator. Thus, we proposed an application where the mix of the visual and auditory senses, facilitate the literary work with extra-textual content that brings simultaneous information about what is occurring in the narrative text in what Paul Bertelson [2] calls valid co-occurrences, where in his example, the sensory information coming from one event such as an explosion is a synchronization of light, noise, heat and pressure.

---

2. Bertelson, Paul, and Béatrice de Gelder. "The psychology of multimodal perception." *Crossmodal space and crossmodal attention* (2004): 141-177.

These co-occurrences applied in literature, using the embedded sonic cues, are thought to produce a cross-modal interaction where the interpretation of read actions is influenced by adding information available in another sense; taking advantage of an intermodal redundancy. During these co-occurrences, the sonic cues could operate in what Evelyn Tribble calls a cognitive ecology of multidimensional content, where the reader interacts with the text to remember, feel, think and imagine through the use of extra-textual content.

We find the use of verbal and non-verbal cues work within both, visual and auditory, in a dual-coding where the eyes can retrieve the text as much as the illustrations in the page, and where the ear combines the dialog with sound effects, to understand the story. For this reason we argue for the use of both channels: the visual to extract verbal information from the text and the auditory to extract information about the context, providing non-verbal information. The aim is to engage the reader without distracting her from the reading task, and find that instead of distracting the vision of the reader toward an illustration or a photo we could provide some of the story's information with sonic cues. In this way we aim to create a dual-coding across senses, where the components of language are extracted from the text, and the sonic components support that information benefiting the reader to extract meaning trough content and context. In this way, the reader makes use of different resources in a process of physical and social distributed cognition.

4.2.1 Dual-Coding Across Senses

In a first approach, to apply our idea of a dual coding across senses, we evaluated sixty-four readers in an AB online survey equally divided in a control group and a test group. Both groups were directed to an online survey containing a PDF file, of the short story *The Watchmaker* by Carla Novi (See Figure 4.1). The difference between the control group and the test group is that the PDF file of the test group was embedded with sound triggered within certain time, calculating the average time speed to release the sound.

Figure 4.1: Sonic Storyboard *The Watchmaker*.

The short story was augmented with four sound effects that complemented the story. It starts with the sound that signals when flight passengers need to put or take off their seat belts. The text never mentions this happening, however the watchmaker presents himself halfway a flight. In this way the sonic cue marks the place: a plane; and a time: to be seated and put on your seatbelt. The next sound supports the main character that is a watchmaker that changes time. The sound of the clock represents the time passed, the time that has been lost. On the fourth page we can hear a woman crying, supporting the negative emotion she has about time passing by her, and not being able to change this on her own. The last sound is the same as the first sound, the seat belt signal that passengers can take off their seat belt, even though they are still flying, but gives the cue for the watchmaker to disappear.

Ground truth Motivation to Read Profile

As ground truth all subjects started the study by answering a Motivation to Read Profile (MRP)[3] to asses the reading motivation of self-concept as a reader and the value each subject gives to reading (See Table 4.2). This gave the study a clear profile of each subject regarding reading habits and capabilities. The survey was designed to elicit information about self-perceived competence in reading and self-performance. The MRP is based on research and theories related between literary learning and motivation where perceived competence and task value are determinant for engagement.[4,5,6]

---

3. Gambrell, Linda B., Barbara Martin Palmer, Rose Marie Codling, and Susan Anders Mazzoni. "Assessing motivation to read." *The Reading Teacher* (1996): 518-533.

4. Wigfield, A. (1994) Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*(1), 48-78.

5. McCombs B.L. (1991). Unraveling motivation: New perspectives from research and practice. *Journal of Experimental Education, 60,* 3-88.

6. Kuhl, J. (1986). Introduction. In J. Kuhl and J.W. Atkinson (eds.) *Motivation, thought, and action* (pp. 1-6) New York: Praeger.

Table 4.2: (Linda Gambrell 1996) Motivation to Read Questionnaire

Motivation to Read Profile

1. My friends think I am ____
[ ] a very good reader [ ] a good reader
[ ] an OK reader [ ] a poor reader

2. I read____
[ ] not as well as my friends [ ] about the same as my friends
[ ] a little better than my friends [ ] a lot better than my friends

3. When I don't understand what is happening in the book I ____
[ ] keep reading and almost always figure it out [ ] keep reading and sometimes figure it out
[ ] go back to find what I've missed [ ] stop reading

4. I tell my friends about good books I read.
[ ] I never do this [ ] I almost never do this
[ ] I do this some times [ ] I do this all the time

5. When I am reading by myself, I understand____
[ ] almost everything I read [ ] some of what I read
[ ] almost none of what I read [ ] none of what I read

6. I am____
[ ] a poor reader [ ] an OK reader
[ ] a good reader [ ] a very good reader

7. I think reading is____
[ ] a boring way to spend time [ ] an OK way to spend time
[ ] a great way to spend time

8. During the week I spend
[ ] none of my time reading [ ] very little time reading
[ ] some of my time reading [ ] a lot of my time reading

9. I like to read from 1 to 4
[ ] Magazines [ ] Books [ ] Newspapers [ ] Comics

10. When I read out loud I am a____
[ ] poor reader [ ] OK reader
[ ] good reader [ ] very good reader

Immersion and Reading Comprehension Questionnaires

In order to verify and compare the difference of engagement and reading comprehension with and without the aid of multimodal literacy, we applied an immersion questionnaire [7] (See Table 4.3) and a sentence verification task (SVT) [8] (See Table 4.4) respectively. These two questionnaires were be filed out by every participant at the end of the reading exercise.

---

7. Jennett, Charlene, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. "Measuring and defining the experience of immersion in games." *International journal of human-computer studies* 66, no. 9 (2008): 641-661.

8. Royer, James M., C. Nicholas Hastings, and Colin Hook. "A sentence verification technique for measuring reading comprehension." *Journal of Literacy Research* 11, no. 4 (1979): 355-363.

In the immersion questionnaire subjects evaluated each story rating the level of engagement, boredom, challenge and distraction by answering if they agree "Yes" or not "No" with the four sentences provided.

Table 4.3: Immersion Questionnaire based on Charlene Jennett, 2008

Immersion Questionnaire

A. Please rate how far you would agree with the statements.
SD strongly disagree D disagree N neutral; A agree; SA strongly agree

1. I felt that I really empathized with the story.
[SD] [D] [N] [A] [SA]

2. I was interested in seeing how the story would progress.
[SD] [D] [N] [A] [SA]

3. I could imagine what the places and character in the story would look like.
[SD] [D] [N] [A] [SA]

4. Reading the story was frustrating.
[SD] [D] [N] [A] [SA]

5. I was unaware of what was happening around me while reading.
[SD] [D] [N] [A] [SA]

6. When reading time appeared to go by very slowly.
[SD] [D] [N] [A] [SA]

B. Please answer the following questions by circling the relevant number.
1 being not at all to 5 being very much so

1. To what extent did you find reading the text challenging?
[1] [2] [3] [4] [5]

2. Were there any times during reading in which you just wanted to stop?
[1] [2] [3] [4] [5]

3. To what extent did you find reading the story easy to understand?
[1] [2] [3] [4] [5]

4. How much would you say you enjoyed reading the story?
[1] [2] [3] [4] [5]

5. To what extent did you find the story boring:
[1] [2] [3] [4] [5]

6. Would you like to read the whole book?
[1] [2] [3] [4] [5]

For reading comprehension we used a sentence verification task (SVT) test consisting of sixteen sentences, four groups of sentences that appear in the original text with three variations: paraphrased without changing meaning, changing meaning altering the main word of the sentence and a distractor sentence consistent with the short story but unrelated to any original sentence.

Table 4.4: Sentence Verification Task Questionnaire based on James Royer, 1997

Sentence Verification Task

1. My friends think I am ____
[ ] a very good reader [ ] a good reader
[ ] an OK reader [ ] a poor reader

2. I read____
[ ] not as well as my friends [ ] about the same as my friends
[ ] a little better than my friends [ ] a lot better than my friends

3. When I don't understand what is happening in the book I ____
[ ] keep reading and almost always figure it out [ ] keep reading and sometimes figure it out
[ ] go back to find what I've missed [ ] stop reading

4. I tell my friends about good books I read.
[ ] I never do this [ ] I almost never do this
[ ] I do this some times [ ] I do this all the time

5. When I am reading by myself, I understand____
[ ] almost everything I read [ ] some of what I read
[ ] almost none of what I read [ ] none of what I read

6. I am____
[ ] a poor reader [ ] an OK reader
[ ] a good reader [ ] a very good reader

7. I think reading is____
[ ] a boring way to spend time [ ] an OK way to spend time
[ ] a great way to spend time

8. During the week I spend
[ ] none of my time reading [ ] very little time reading
[ ] some of my time reading [ ] a lot of my time reading

9. I like to read _____ (Please rate from 1 to 4, being 1 what you most like to read)
[ ] Magazines [ ] Books [ ] Newspapers [ ] Comics

Readers are presented with four groups of four sentences where two have the same meaning as in the short story -original and paraphrased- and other two that do not relate to the story - changed meaning and distracter, were the reader needs to write two "yes" and two "no" in each group, setting apart the related and unrelated sentences. A sentence verification task rests on the readers' memory retention preserving meaning during mental simulations of what is being read and not on exact words. The performance of the SVT falls at 75% of the readers being tested scoring an average of 80% and above for proficient reading comprehension and under 70% for struggling readers.[9]

---

9. Leslie, Lauren, and Jo Anne Caldwell. "19 Formal and Informal Measures of Reading Comprehension." *Handbook of research on reading comprehension* (2009): 403.

Both the immersion questionnaire and the SVT are analyzed to identify significant correlations between engagement and reading comprehension. For example, if in the immersion questionnaire the reader declares she empathizes with the story and rates it as engaging then the SVT should have a score above 80%. On the contrary if the reader finds the story boring and frustrating the STV score are expected to be low, under 70%, reflecting a lack of engagement and understanding. Base on these data our presupposition was that the test group would have better scores in the STV questionnaire than the control group, because of the valid co-occurrences of sound and words. Also that by including sound effects into the literary text, using sound design based in what has been done in theatre and radio drama, these sonic aids will not disrupt but enhance the reading experience. This should been seen in the date of the immersion questionnaire, where we expected a higher percentage of engagement than in the control group.

Table 4.5: Immersion Questionnaire Unpaired Test

	With Sound	**No Sound**	***t*-test**
Challenging	17.9%	32.26%	$t(58)=1.282$, $p=.205$
Boring	7.1%	19.35%	$t(58)=1.385$, $p=.171$
Distraction	21.4%	29.03%	$t(58)=0.676$, $p=.502$
Engagement	89.3%	90.32%	$t(58)=0.132$, $p=.896$

However the immersion questionnaire only showed significant differences within each modality except for the Challenging parameter, which showed no significant difference without sound. The control group showed a higher statistical level of boredom compared to the study group who listen to the sound cues. In an unpaired *t*-test between percentages none of the parameters (See Table 4.5). Our SVT did not show any difference between the two modalities, in fact both groups were consistent in only one mistake, the rest of the sentences were answered showed a significant difference at the .5 critical alpha level. We attribute this to the fact that the short story was not long enough to challenge memory or comprehension.

4.2.2 Dual-coding of verbal sonic information

In a search for other applications for our empathetic system, we decided to alter the sonic non-verbal information of the prototype made from the short story *Sleepy Hollow,* into verbal information. Instead of using non-verbal cues, such as sound effects and soundscapes, we designed an interaction that provides the reader with the audiobook of the text being read. The idea comes from verbal redundancy, where information is presented in a combination of modalities. These modalities, as our dual coding across senses are also based on Pavio's dual coding of auditory and visual channels, where readers process the same information twice. Penney [10] found that the words presented in these two modalities produces and enhanced memory recall. Thus, Moreno and Mayer [11] applied this verbal redundancy in multimedia learning, where the addition of on-screen text facilitated the narration of a scientific explanation. When words are presented in both channels, learners were able to select both modalities of information but with no cognitive overload, since working memory and auditory working memory are independent processes. Moreno and Mayer collected data from thirty-eight college students divided into two groups. One group heard only the narration of the scientific text on "how the lightening process works" and the other simultaneously hear and read the same narration (See Table 4.6). They measured the students' memory with a retention test where they had to explain the scientific text, and a transfer test, where they had to apply the acquired knowledge.

Students remembered significantly more, $F(1, 70) = 12.18$, $MSE = 143.09$, $p < .01$; they generated significantly more conceptual creative solutions on the transfer test, $F(1, 70) = 36.56$, $MSE = 53.81$, $p < .01$; and they correctly matched more items, $F(1, 70) = 3.82$, $MSE = 11.02$, $p = .05$, (with a tendency to be significant) when the verbal material was redundant than when it was not.

---

10. Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17,* 398–422.

11. Moreno, Roxana, and Richard E. Mayer. "Verbal redundancy in multimedia learning: When reading helps listening." *Journal of Educational Psychology* 94, no. 1 (2002): 156.

Table 4.6: (Moreno and Mayer 1989) Results

	Retention		**Transfer**		**Matching score**	
	Mean	Standard	Mean	Standard	Mean	Standard
Narration	5.21	2.78	0.86	0.94	4.63	2.09
Narration+Text	8.05	3.76	2.47	1.43	5.84	1.77
Redundant	9.70	4.03	2.81	1.39	6.49	1.74
No-redundant	6.92	3.57	1.11	1.05	5.70	2.07

Their results show a positive effect into providing the reader with the same verbal information in both channels. As in our first design the system interacts with the reader assisting her to better extract meaning, but this time it gives the reader phonic, intonation and rhythm. When the system, using the nose temperature metrics, finds that the reader is submerged into the reading experience it then stops the auditory aid, leaving the reader to applied the learned intonation and follow her own rhythm (See Figure 4.2).

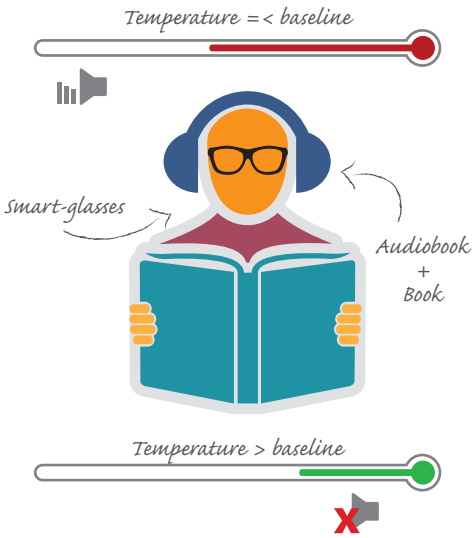


Figure 4.2: Sonic Interaction with Temperature.

Literature of alphabetical languages uses the printed letters as visual symbols to map units of sound, or phonemes of the spoken language.[12] On the other hand, the Chinese and Japanese language uses characters as units of intricate strokes which meaning is suggested by the overall visual shape. These characters map onto phonology at the monosyllable level, and have no units of sound like alphabetical writing. Moreover, it is never the case that Chinese character compounds can be phonetically mapped through the sub syllabic phonological representation of the character's different components. Chinese characters do no allow for segmental analysis, fundamental to the alphabetic system.[13] Therefore the system verbal sonic information can assit the reader with a second channel to process the phonological information. This allows the reader to use the audiobook as a training system for phonological assistance.

4.3 Sonic Editorial Design

In order to integrate sound in literature we present our methodology for an editorial process. The design of sonic interventions is thought to take part of an editorial process where editor and author include a sound designer to their team, in order to deliver a structured multimodal storytelling (See Figure 4.3). The sound designer works very closely with both editor and author to create sounds that will register in the reader's mind instantaneously. This sounds are created to call for the reader's attention surprising her with sonic information designed to assist the imagination. The addition of the sound designer into the editorial process means that the story needs to be studied and analyzed to include resonance. This is for the sound to be of assistance, not a disruption, and justify its presence within the text.

---

12. Coltheart M, Curtis B, Atkins P, Haller M (1993): Models of reading aloud: dual-route and parallel-distributed-processing approaches. *Psychol Rev* 100:589–608.

13. Tan LH, Hoosain R, Peng DL (1995): Role of presemantic phonological code in Chinese character identification. *J Exp Psychol Learn Mem Cogn* 21:43–54.

Once the sounds designer reads the manuscript he then analyses and discusses it with editor and author to design the sounds that will be included, thinking about the difficulties that the text might present to reader's with poor background knowledge. The sounds are then designed to present the required extra-textual information that allows the reader to better imagine the literary world.

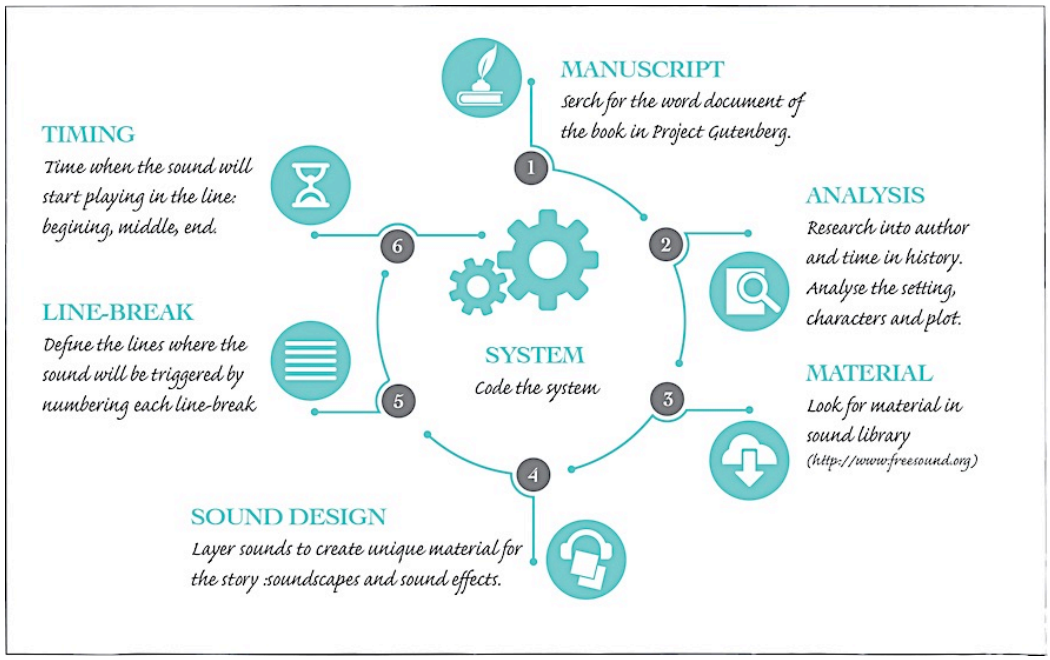


Figure 4.3: Sonic Editorial Process.

As pointed out before the use of sound is inspired by characteristics of orality, specifically a multimodal sensory processing of information. As described, sound carries meaning and information, and it can be used as extra-textual information. As described before there can be two kinds of sonic cues: when the story gives information on the cause of the sound (link to the source) and when there is no textual information about the origin of the sound (inferential).[14] Pauletto [15] also suggests the use of substitution and translation as a creative device for poetry, where ideas are connected by metaphors. For example, a car is cruising along the Mediterranean road, then the sound of a cork twisting out of a wine bottle is the metaphor for Mediterranean.

---

14. Chion, M. 1994. Audio-vision. New York: Columbia University Press.

15. Pauletto, Sandra. "Film and theatre-based approaches for sonic interaction design." *Digital Creativity* 25, no. 1 (2014): 15-26.

Listening supports the reader's imagination, using sound effects and soundscapes as multimedia aids. In general multimedia aids can be classified into three types depending on what these pretend to achieve .[16] For example, the text can have aids for "selecting information" helping the reader focus her attention on the target information to correctly process that information. The processing of a deviant stimuli such as the sound effect can linger in the immediate memory long enough to interact with the processing of the target task. If the sound is incompatible with the target task it might cause distraction, but if it is compatible distraction is reduced.[17]

Another kind of aid is for "building internal representations" that support internal connections of the information presented, organizing it into a coherent structure of the logical relations amongst idea units in the text. The sonic aids, sound effects and soundscapes, are designed to build internal representations, based on the information and meaning they carry. There are also aids for "text comprehension" which support external connections designed to help with the ideas in the text, by integrating them into existing mental models. On a macro level, we assume reading comprehension with sonic multimedia aids draws a parallel with the theory of Multimedia Learning.[18] This theory considers the reader as a knowledge constructor who actively selects relevant information and who actively selects relevant information, organizing it into coherent mental representations by integrating the newly constructed representations with one another. Thus, we expect the embedded sound effects to carry narrative information that will assist the reader for comprehension. Here we assume the process of perception is entwined with the process of naming, which in turn refers to the possibility of language to describe the recognition of sound sources and interactions.[19]

---

16. Mayer, R. E. (1984). Aids to text comprehension. *Educational Psychologist*, 19, 30-42.

17. Parmentier, Fabrice BR. "The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review." *Psychological research* 78, no. 3 (2014): 321-338.

18. Mayer, Richard E. "Multimedia learning: Are we asking the right questions?."*Educational psychologist* 32, no. 1 (1997): 1-19.

19. Metz, C. 1977. Essais Se´miotiques. Paris: Klincksieck.

4.3.1 Manuscript

Sound uses our experience of the world to infer what the specific sound source is. It is in this way that we find sound made theatre audiences to perceive what they could not see, stimulating their mental imagery with sonic cues for a better narration. As done in theatre the sound design can be integrated while editing the manuscript. Editor can replace screen directors, who work closely with sound designers to present the idea of the author in the best possible manner. In this way the sound is integrated to the already established editorial process of books creating a new department that works along with editors, authors and sound designers to create a multimodal and professional storytelling. In our prototypes we assumed the roles of editor and sound designer, working with the author to achieve what she wanted to communicate.

4.3.2 Analysis

To implement sound in literature we propose the use of short stories to design the first prototype. The short story was chosen because it is a form of narrative prose-fiction that encompasses the traditional elements of dramatic structure like exposition, complication, crisis, climax and resolution, but at the same time is not as complex as the novel, using only a few characters in only one setting. Unlike the novel, short stories communicate a limited sequence of events, experiences or situations in a closed order to create its own perception as a totality, in a limited continuity. While the novel tells a life, the short story tells a fragment of a life, called the moment-of-truth, a crisis that will change the central character forever. These characteristics have given the short story the capability to be used for formal experimentation; to introduce new subject matters into the literary arena; and to break the establishment. As an illustration, for Bret Harte,[20] American author, the short story singled the end of English dominance in American literature. He used this genre to express his country's history, and as a form of detachment from English literary patterns; being most famous for his short stories featuring the California Golden Rush.

---

20. Harte, Bret. *The Rise of the Short Story*. Anmol Pub., 1990.

Similarly in France, Maupassant considered one of the fathers of the modern short story, used this genre to break the taboos on sexuality and class, while Joyce used the short story to document modern Irish life. Finally, the genre of short story has been selected because of its oral asymmetry with the novel. While the written formats are similar to the novel, short stories have a closer tradition with orality. The short story has provided popular for regional cultures to reproduce the length of most oral speech events. In fact, orality can be regarded as an important factor for the flourishing of short stories in modern literature of Third World nations, where the established literary language is that of an oppressor.[21] Hence, as the goal is to re-introduce the subject of orality and break the establishment of silence in print, short stories become the right genre to present an our multi-sensory reading experience, to experiment and design sound for literature. Specifically, we have chosen four short stories by Carla Novi of her series *Encounters* and the short story of *The Legend of Sleepy Hollow*, by Washington Irving, contained in his book *The Sketch Book of Geoffrey Crayon, Gent.*

The series of short stories were picked because they depict unreal characters in real setting or real character in unreal settings. They call for the reader to use her imagination to understand not only the fictional world created by the author but also the message of each story. These short stories are set to challenge the reader to imagine what she cannot find in the natural world. On the other hand Irvin Washington's short story looks for the readers mental imagery of a place and a time in American history. Set in the post-revolutionary war America, in the rural boondocks of New York, Sleepy Hollow is a quiet Town by the Hudson River. The time and place are important to the story since the headless horseman is thought to be a soldier of the war. The story is narrated in third person, but in a she-said, he-said urban legend style. The narrator is Diedrich Knickerbocker, who heard the story from an old man, who probably read the story from Geoffrey Crayon' sketchbook, the pseudonym of Washington Irving.

---

21. Pratt, Mary Louise. "The short story: The long and the short of it." *Poetics* 10, no. 2 (1981): 175-194.

The third person point of view distances the reader from the story, since the narrator doesn't really know what is going on half the time. Knickerbocker is unable to give enough facts, forcing the reader to figure out the truth on her own. The writing style is not of much help either. Irving's sentences are littered with commas, semicolons, dashes and colons, reaching lengths of up to ten lines. This writing style, although not uncommon for 19th century writers, can give the story an old-time feel to today's readers. Irving is also guilty of being too descriptive in his short story. He uses sound and sight to describe his settings in such detail that the description feels as if to slow down the story, detracting from the little action there is in it. Overall, *The Legend of Sleepy Hollow* provided us with a short story that has an exact time and place from where to give extra-textual content; and a storytelling of long descriptive sentences that can benefit from the extra-textual stimuli of sound.

4.3.3 Sound Material

For the series *Encounters* sound effects are used to give the storytelling a creative point apart from the written word. They are artificially created or enhanced to emphasize the content of the narrative, calling for the reader's attention. The sounds can give information depending on their classification. For example, isolated sounds are real specific sounds like the flight's seatbelt announcement used on *The Watchmaker* or a the sound of someone using a hammer used in *The Last Space* which can be related to the thing that produced it. These are also called hard sounds. On the other hand, there are specialty effects, sounds that only exist in the fictional world the reader creates in her mind. These sounds enhance the fictional world by delivering a special sound that represents non-natural things. The sound can be, for instance of the create that lives in the attic in the short story *That Which Lives in the Attic* or the sirens of the human kennel in *John.*

On the other hand *The Legend of Sleepy Hollow* was first analyzed to define place and time to design appropriate sonic aids that would give information about the sounds found in that place at that time. Irving starts his short story defining the place, New York in 1790, and comparing the noisy town to the more quiet countryside. In the first paragraph he refers to the Hudson river, the market town or rural port to describe the a noisy dwelling. The author uses the brook, quails and woodpeckers to describe the more quiet setting of his story: Sleepy Hollow (See Figure 4.4).

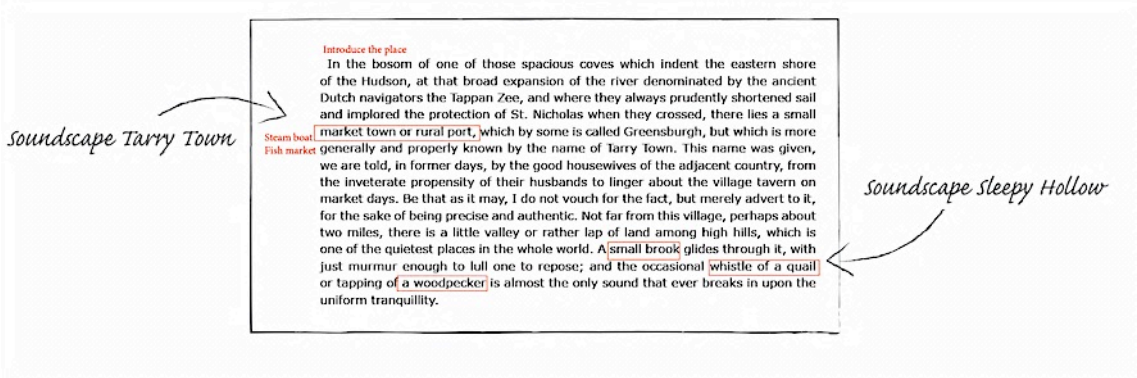


Figure 4.4: Embedding Sound.

The sound material also includes the audiobook from *Sleepy Hollow* which is used to aid the reader in a bottom up reading -whereas the use of non-verbal sonic cues are aids for a top down reading. The verbal dramatization of the words taken from Librivox is broken down to fit the textual contents of each page (See Figure 4.5). Each file is about a minute. This is thought to help the reader to find the right intonation within the paragraph long sentences of Irving Washington.

Figure 4.5: Sound Edition.

4.3.4 Sound Design

Soundscapes layering is used to create new sounds out of two or even three average sounds to create some background ambience. Ambience tracks provide the environment for the story as a subtle atmosphere that makes the listener to get a feel the setting. This sounds do not explicitly synchronize with the text but indicate the setting to the audience. Such atmosphere is created for the town of Sleepy Hollow. This place is clam and surrounded by nature as Knickerbockers narrates when describing the whereabouts of the town. But this town is also mysterious and scary, and therefore a sound layering between a strong wind and a synthetic sound to give a feeling of darkness is embedded to the description of a peaceful spot embossed in the great state of New York. But it is a strange place nonetheless, and the sonic information tells the reader exactly this.

To represent in sound these two different towns we designed soundscapes. Each soundscape includes sound marks, sounds unique to the area,[22] some provided by Irving, some others by researching into post-revolutionary America. Irving tells us that Tarry Town is a rural port or market by the Hudson river in 1790. Therefore the soundscape starts with a layering of voices in a market, a Dutch market, to emphasize the first settlers (Dutch Navigators) and to add an blab noise that the reader will not understand and be distracted by. In addition the soundscape has a sound mark of the Hudson river heard in the distance: a steamboat. There is no mention of a steamboat in the text but we know that the steamboat was introduced in America in 1787. Moreover, the Fulton's ferry service navigated the Hudson river from 1802. On the other hand the soundscape of the surroundings of Sleepy Hollow evokes nature, where birds live across a small brook. Irving mentions the quail and woodpecker and so the soundscape is a layer of running water, the hammering of a woodpecker and the whistle of a quail. For the reader, the sounds of the brook and woodpecker might be familiar. But maybe the whistle of the quail is not as familiar to most readers.

---

22. Schafer, R. Murray. The new soundscape: a handbook for the modern music teacher. BMI Canada, 1969.

By providing the three elements described by Irving into one soundscape we can give the reader with new information about how a quail should sound, supporting the connection between word and sound. The MixPad software is used to create soundscapes unique to the story (See Figure 4.6).

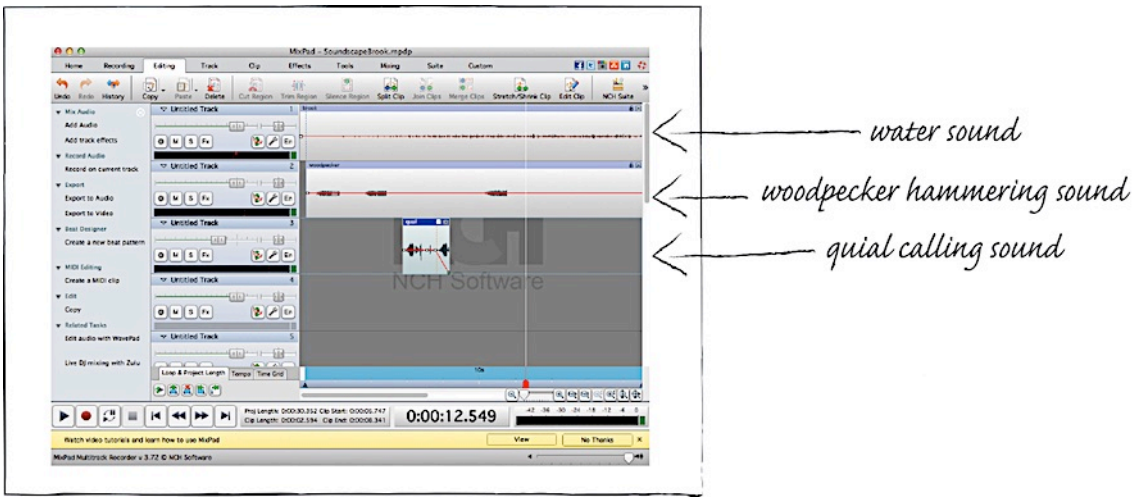


Figure 4.6: Sound Layering.

4.3.5 Line Break

The line break is a term to describe the number of lines within a page. These lines determine where in the text the sound should be embedded. Each page is divided by the number of lines and columns the text has. Then the sound is assigned to one of those lines depending on the sound design (See Figure 4.7).

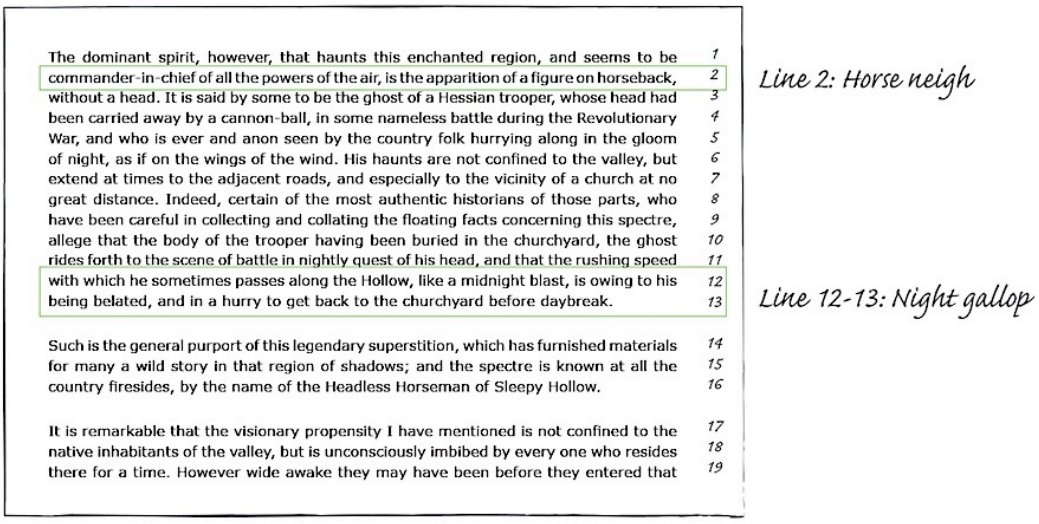


Figure 4.7: Line-Breaks.

This can also be done by creating a window, within a specific area within the text where the sound is embedded. In the design of the series of *Encounters,* we used two columns that help the eye concentrate in a smaller area (See Figure 4.8). Dividing the text by lines helps to embed the sound in a passage of the story more than in an exact word. It can be difficult to trigger the sounds in an exact moment; therefore we design our sounds with this in mind. Line breaks give us more room as to when the sonic cues will be activated.

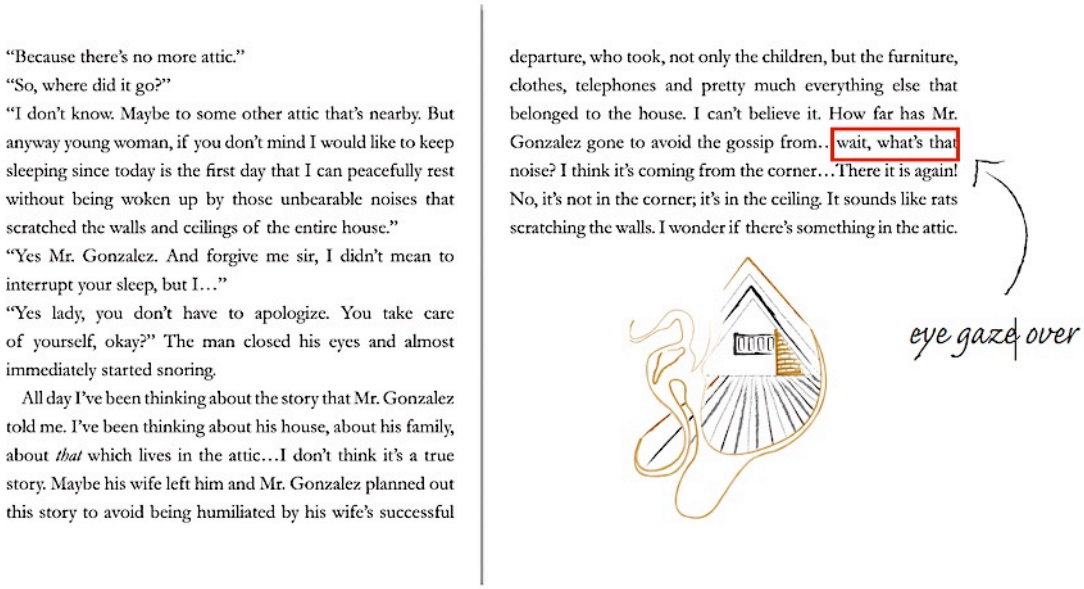


Figure 4.8: Gaze Over.

4.3.6 Timing

In average readers reach 250 words per minute, making time a good starting point to deliver the sound effect, specially the isolated sounds. If there is a need for more exactitude, once the line where to deliver the sound is determined, each sound could be triggered at the start of the line, middle or end. For that, we add silence at the beginning of the each sound to more or less adapt to the text. The text is designed to include approximately 75 character per line. This gives an average of 15 words per line, which is a small enough time frame to control. For instance in the example of line break of *Sleepy Hollow*, in line two the sound of a neighing horse comes right at the end of the line (See Figure 4.7). Taking in consideration the reading speed and an average of fifteen words per line, each line takes approximately 3.5 seconds to read.

We then divide the line into three sections: beginning, middle, end, and add the silent time required to get to that section. In the named example, the sound of a neighing horse is at the end of the sentence, thus we can add two thirds of the total time for each line of silence before hearing the neighing. The next sound is in line twelve, but the gallop of the horse should start at the beginning of the sentence, so no silence is added.

4.3.7 Coding

In order to create an application of a multi-sensory literature we used the Processing IDE and programming language to recreate the looks of a digital book (See Figure 4.9). We had done our previous study with sound using the editorial program InDesign to create the sonic cues. However, this could only be done calculating the time until the reader would be going through the desired passage. The problem is if the reader is distracted then the time does not stop. There fore we looked into Processing as it is a program accessible to artist and designers to recreate the process done in InDesign.

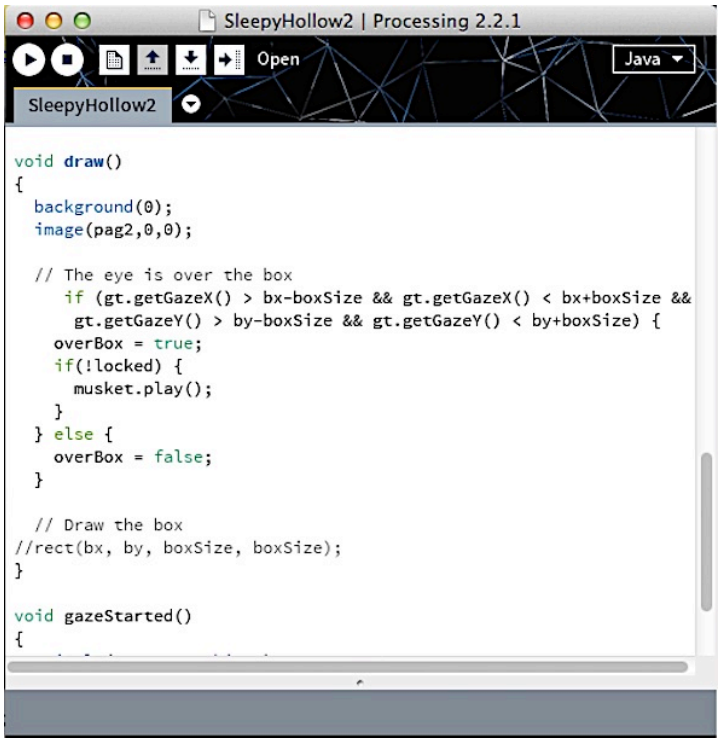


Figure 4.9: Code.

The program allows coding a digital book that is no different from what the reader might be expecting. Furthermore, Processing uses the same language as Arduino, and thus allows for a serial call to connect our prototype of smart glasses and the short story with the embedded sound.

4.4 Interactivity

Through this sonic editorial process is that we propose to bring back orality to literature. We want to exploit the spontaneity of sound that comes with a situation and vanishes with the situation itself. This multi-sensory literature plans to be spontaneous by coding sound into the narrative. Sound is fundamental to perceive and interact with the natural world and the use of it in literature is set to create an embodied experience where the reader can perceive the fictional world in a natural way, with more than one sense. This might force the mind to make connections that could otherwise not make due to the lack of knowledge of the name of a bird, for example. But if the reader is listening to a soundscape she could make the connections between the sound and what is being read.

In synthesis we can go back to the concept of an ecology of literary texts, where we are introducing a stimuli that was lost with print. In this new ecology sound, with it's many properties, assists the written word to communicate the author's literary creation. In this interaction the story is narrated through two different channels, the visual and auditory, that assist and complement each other. This sonic stimuli is a translation of Shakespeare's trick, a disruption that integrates with the meaning of the story.

Chapter 5

Augmented Narrative

Augmented Narrative is a responsive literary storyteller that interacts with its reader using sound to stimulate the imagination with sonic information. The system allows for an interaction between text and reader through the spontaneity of sound trigger by a biofeedback dependent on the nose temperature.

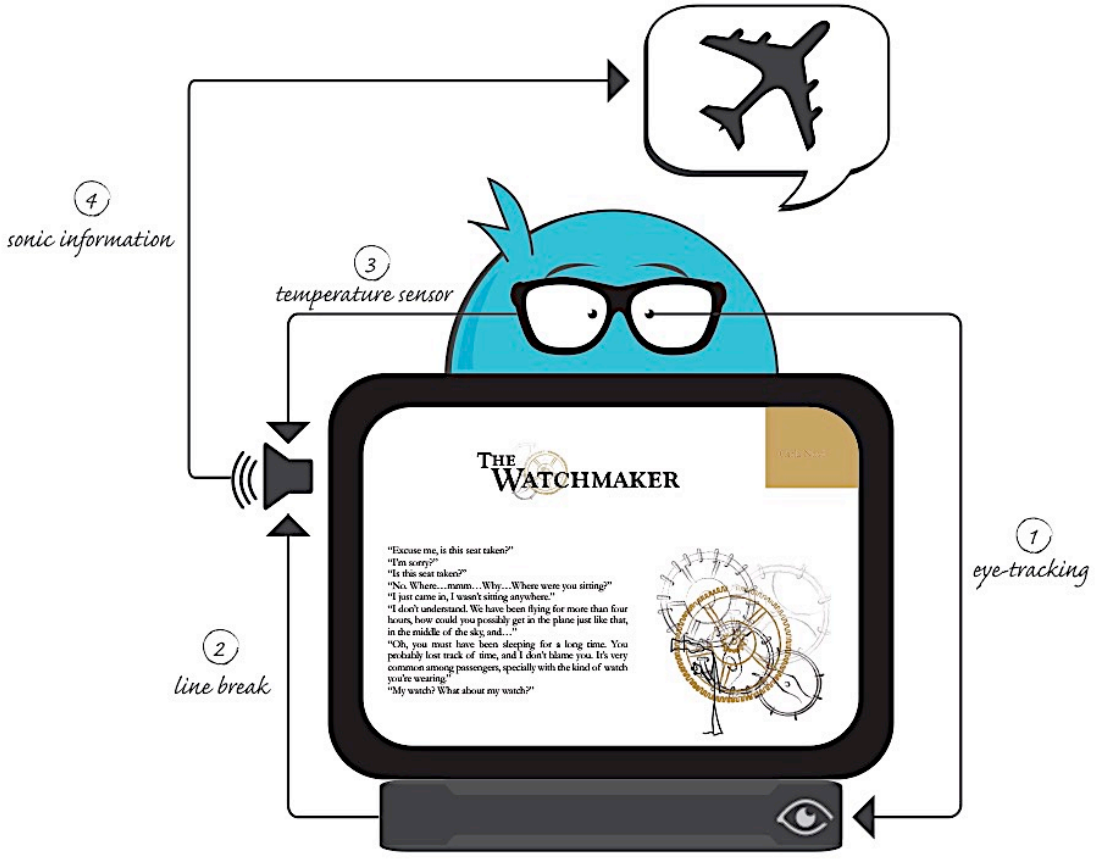


Figure 5.1: Augmented Narrative System.

With the application of the metrics of engagement using the smart-glasses, literature can intervene when the reader needs it. These interventions are proposed to provide the reader with a multimodal reading experience, where attention and perception are technologically mediated through the use of sound triggered by a lack of immersion in the story. These sounds, prompted by the involuntary reactions of mind and body, are used to call for the reader's attention, while at the same time provide with useful sonic information related to what is being read. This is set to operate as a dual coding across senses, where the primary information is delivered by the written word through the fictional literary text, and where extra-information is delivered through the auditory channel with sonic cues that act as a valid co-occurrence of what is being read. This means that the information of the sonic cues is not just only related to the text, but also gives the reader multimodal information to assist her perception and imagination. If the reader cannot create a mental image from what she can extract visually from the text, she might be able to do so with information delivered in a combination of text and sound; creating a redundancy from where to retrieve meaning.

This text-body interaction is set to excel transportation, transforming storytelling into a multimodal literacy suited to improve reading comprehension through the personalized use of unexpected stimulations of sound that promote a meaningful reading experience. The system follows the reader's engagement by measuring her mental workload, reflected on a decrease on the nose temperature. The system provides the reader with extra-textual information until it finds that the temperature of the nose has decreased. This signals the system that the reader is finally immersed into the story and is in no need for extra stimuli (See Figure 5.1). On the contrary, if the reader has not had her nose temperature decreased, the system understands she has not yet been immersed in the story and thus still requires assistance. In this way, the medium extends from hand to eye, to include the ear, in what we propose as a cognitive ecology of literature. The endeavor is to extend the reader-response criticism to include a technologically mediated unexpected stimulation that will not only call for the reader's attention, but also fill in information gaps that the text could be missing.

5.1 Eye-tracking

The system uses off the shelf eye-tracking technology to link the perception of the text with the location of the eyes. The smart glasses can know when there is engagement or a lack of it, but are not able to determine where the reader's gaze is within the text. For the sound to be delivered in accordance with the text, the system needs to link the engagement data with the eye-gaze of the reader. To solve this problem, we use an eye-tracker add on. The eye-tracker can follow the reader's gaze through the text, determining which sound could be triggered, if the system finds that there is no transportation into the story. The eye-tracker serves to follow the reading progress while the smart glasses follow the perception of that progression. This allows the system to leave each individual to read at their own pace in their own time. The design relies on the works of Biedert, Buscher, Dengel and Schwarz's Text 2.0, who had an eye-tracker to act as the cognitive agent in an interaction where gaze responsive annotations embedded into a webpage are detected on the span region.[1] In their prototype they use an interaction with the text trough commands of fixation, gaze over, and on read. In fact, there are already different applications using Text 2.0, such as the EyePad.[2] This system uses the eye-tracker to link what is being read with an animation of the text, showing the progress of the reader within the animation. The EyePad also allows an interaction with voice, where the reader can fixate on the unknown work and ask a question about that word. The system then delivers an audio recording of the meaning. Be that as it may, Augmented Narrative uses Text 2.0 only as an input of location, leaving the interaction to the smart-glasses. In our system the user's cognitive process is automatically obtained through the skin temperature of the nose, without causing any distractions that interrupt the reading exercise.

---

1. Biedert, Ralf, Georg Buscher, Sven Schwarz, Jörn Hees, and Andreas Dengel. "Text 2.0." In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pp. 4003-4008. ACM, 2010.

2. El Hosseiny, Mostafa, Ralf Biedert, Andreas Dengel, and Georg Buscher. "The eyePad-Tom Riddle in the 21st Century." In *Proceedings of the 2nd Workshop on Eye Gaze in Intelligent Human Machine Interaction held in conjunction with IUI 2011*. 2011.

The system of Augmented Narrative follows the eye gaze of the reader to understand where she is in the text using the tobii eyeX eye-tracker, placed at the bottom of the screen. By following the eye gaze, the system knows when to deliver the sound (See Figure 5.2). We continue using the open source Processing IDE and programming language to connect eye-tracker, text and sound together (See Figure 4.9). In it we utilize the Gaze Track Plugin developed by Augusto Esteves [3] to connect the eye-tracker to the Processing sketch. This plugin allows using the tobii eyeX as if it were part of the Text 2.0 framework developed by Ralf Biedert [4] who also developed a core plugin that accepts raw gaze information from a remote eye-tracking server. The code starts by adding the text and sound using the mouse as the test tracking system. Once the system works with the mouse tracking it can then be modified to use the Gaze Track Plugin and the eye-tracking data. The audio was set using the Minim audio library [5] that uses JavaSound API, Tritonus and Javazooom's MP3SPI. This allows for an easy integration of the audio in the sketch without the need of callbacks or arrays.

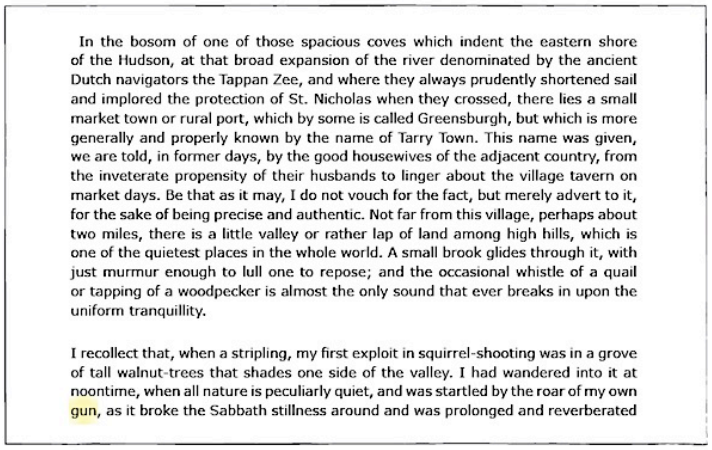
Through the analytical reflection that spontaneous is good, this system allows literature to carefully plan to be spontaneous using the embedded sound to address particular passages that vanish with the passage itself; just as seen in verbal communications. The text is carefully edited and planned to stimulate the reader with different sounds that complement the story through a dual-coding, activated at a specific time by following the reader's progress throughout the text. Once the reader is reading through the lines where the embedded sound is, the system triggers such sound. The multi-sensory experience brings the empathetic identification with the known from the natural world of oral cultures in an interaction shaped in form and content by an anticipated response of the text.

---

3. https://github.com/AugustoEst/gazetrack/tree/master/GazeTrack

4.Biedert, Ralf, Georg Buscher, Sven Schwarz, Manuel Möller, Andreas Dengel, and Thomas Lottermann. "The text 2.0 framework." In Workshop on Eye Gaze in Intelligent Human Machine Interaction, pp. 114-117. 20105

5.http://code.compartmental.net/tools/minim/



In the bosom of one of those spacious coves which indent the eastern shore of the Hudson, at that broad expansion of the river denominated by the ancient Dutch navigators the Tappan Zee, and where they always prudently shortened sail and implored the protection of St. Nicholas when they crossed, there lies a small market town or rural port, which by some is called Greensburgh, but which is more generally and properly known by the name of Tarry Town. This name was given, we are told, in former days, by the good housewives of the adjacent country, from the inveterate propensity of their husbands to linger about the village tavern on market days. Be that as it may, I do not vouch for the fact, but merely advert to it, for the sake of being precise and authentic. Not far from this village, perhaps about two miles, there is a little valley or rather lap of land among high hills, which is one of the quietest places in the whole world. A small brook glides through it, with just murmur enough to lull one to repose; and the occasional whistle of a quail or tapping of a woodpecker is almost the only sound that ever breaks in upon the uniform tranquillity.

I recollect that, when a stripling, my first exploit in squirrel-shooting was in a grove of tall walnut-trees that shades one side of the valley. I had wandered into it at noontime, when all nature is peculiarly quiet, and was startled by the roar of my own gun, as it broke the Sabbath stillness around and was prolonged and reverberated

Figure 5.2: Trigger Sound.

Pauletto [6] found that even though the synchronic effect when sound and action are synchronized has a strong impact in meaning making, unrealistic matches still allow for interpretation. In fact the plausible and implausible associations between sounds and actions can create expectation and surprise that can even be engaging for the audience. This gives the embedded sound effects, specially the isolated sounds a bigger window to be triggered. They do not necessarily need to come in an exact cue, such as a word they are supporting. This makes possible to use lines breaks, using an eye-tracker, to define the line where the sounds will be triggered. By defining a window of lines, the eye-tracker allows the system to follow the reading speed of each individual, which may vary not only between readers but also between situations of the same reader.

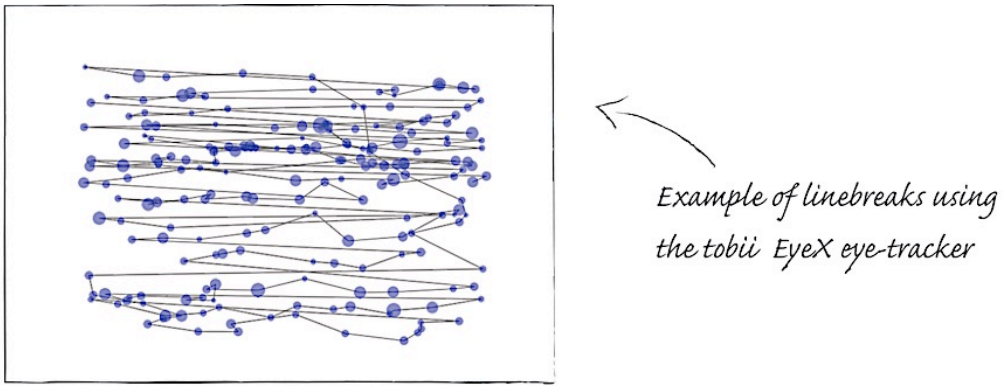


Figure 5.3: Using Eye Activity to Count Line Breaks.

---

6. Pauletto, Sandra. "Film and theatre-based approaches for sonic interaction design." *Digital Creativity* 25, no. 1 (2014): 15-26.

5.2 Line Break

Kunze et al.[7] in their study of mobile eye-tracking used line breaks to create a wordometer to estimate the number of words read on each line. Line breaks are an easy eye-movement to detect because they go against the regular reading order, from the end of the line to the start of the new line (See Figure 5.3). Kunze detected the line breaks using ($g1,\ldots,gl$) to represent subsequent gaze points, being recognized as the line break if the area that circumscribes the l gazes is larger than a certain threshold and the direction of the eye movements from the l - k th to the l th gaze is opposite. However, Kunze et al. found that short lines were a source of error in the line break detection, since detection was based within the area of a bounding box of gazes. This means that the number of line breaks might need calibration depending on each page.

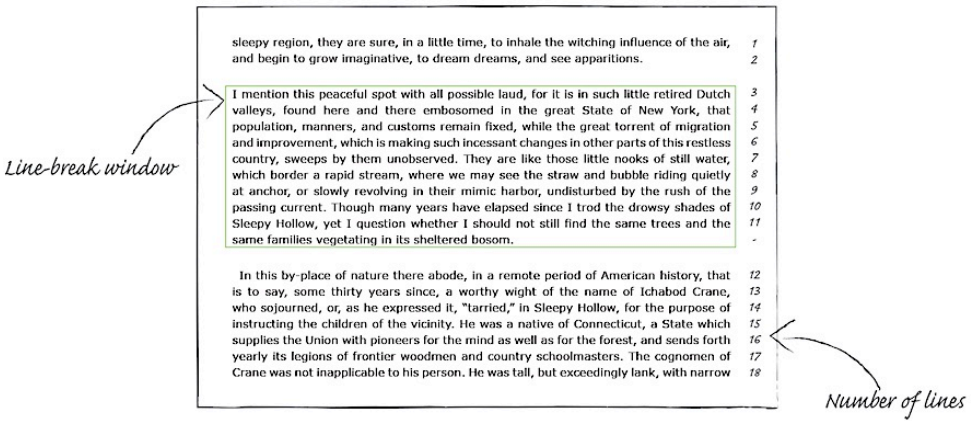


Figure 5.4: Sound Window.

The same method can be used to define the length of atmospheric sound effects and soundscapes. Depending on how many lines the sound needs to cover the, is multiplied by 3.5 seconds to define it's length. For instance, an atmospheric sound that describes the quiet but strange surroundings of Sleepy Hollow is designed to cover from line three to line eleven (See Figure 5.4).

---

7. Kunze, Kai, Hitoshi Kawaichi, Kazuyo Yoshimura, and Koichi Kise. "The wordometer--estimating the number of words read using document image retrieval and mobile eye tracking." In *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, pp. 25-29. IEEE, 2013.

Counting the last half line, not for line break purposes, the sound should last for nine and one half lines (9.5 x 3.5) giving a total of around 33 seconds. Since the atmospheric sound is part of the whole paragraph, this would not need silence added to it.

5.3 Temperature Sensor

The temperature measures the effectiveness of the narrative exploiting psycho-physiological changes as a bridge-for-understanding between reader and text.

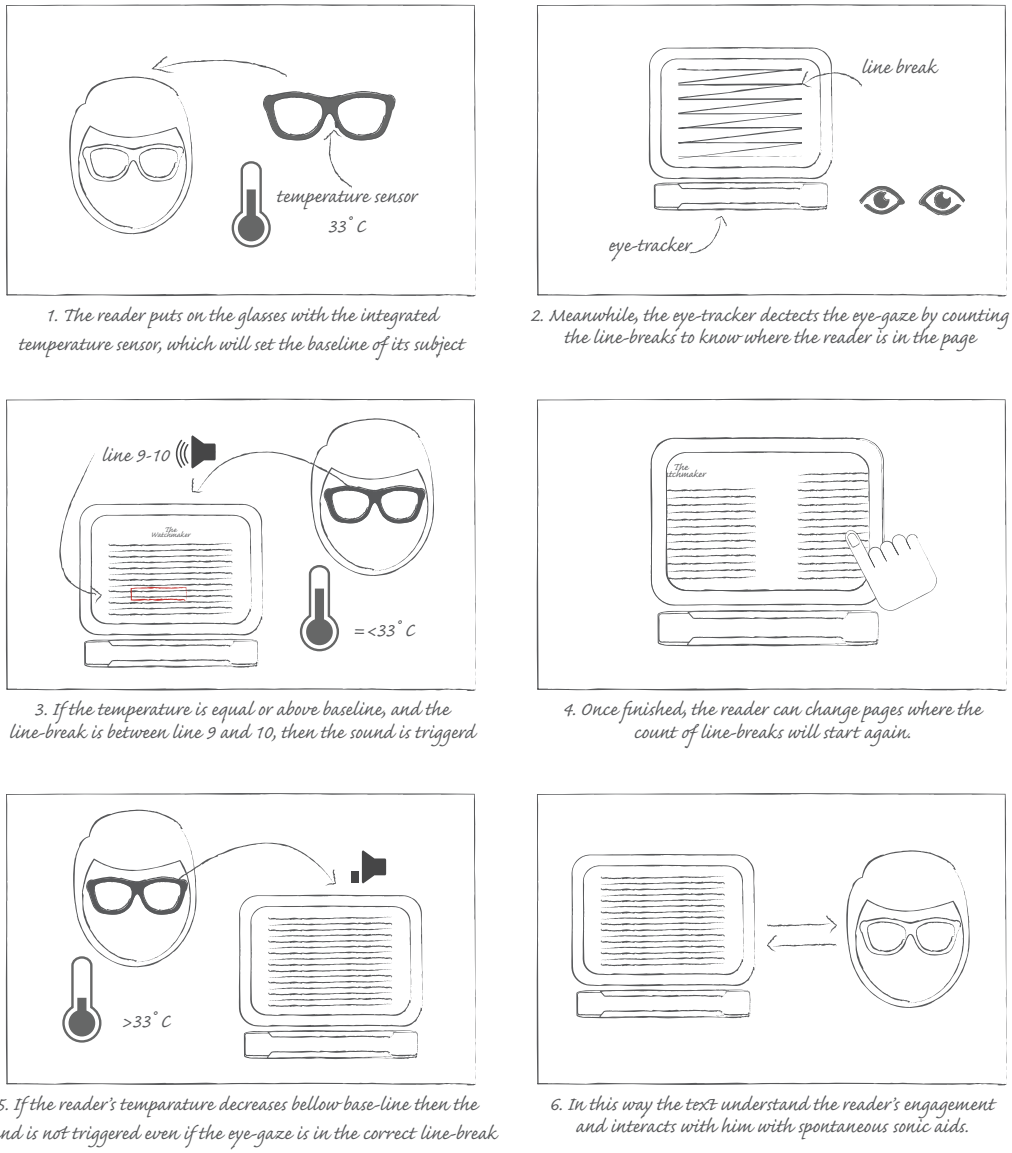


Figure 5.5: Sound-Temperature Storyboard.

The digital literary text, influenced by orality, incorporates an empathetic interaction where the text is able to stimulate the reader when needed, calling for her attention and aiding her to create context, filling in gaps that prevent mental simulation. The system needs to find two conditions. The first one is to identify if the person is reading within the window where the sound has been embedded - with the eye-tracker. Using our smart-glasses for engagement the temperature of the nose defines the second condition. If the temperature does not decrease, then the sound should be activated (See Figure 5.5). This is how attention is technologically mediated, creating a medium that is receptive to the reading experience, just as actors are when performing a play.

5.4 Sonic Information

Once the system knows where the reader is in the text with the input of the eye-tracker, and perceives that there is no engagement yet with the input of temperature, it then can decide to release the embedded sound. The idea of the sonic agency in literature is to bring the concept of theatre-sound-design to create context and enhance mental simulation. This idea has also been considered to give extra-textual information for the blind. Dr. Mariana Lopez,[8] a major in medieval theatre, is carrying out a pilot study to replace audio description in films for the blind, with acoustical information. Focused on sound effects, surround sound, and sound layering, she has developed an audio film that can be followed by a visually impaired audience, without the need of a narration track. The system of an Augmented Narrative also presumes that acoustical information can be as conveying as the textual narration. The multi-sensory experience brings the empathetic identification with the known from the natural world of oral cultures, using the embedded sound to address particular situations that vanish with the situation itself. By adding sonic agency to the literary text, the extra-embedded sound could perform a relationship with the environment giving an enriched meaning to the narrative.

---

8. Julieta Lopez, Mariana, and Sandra Pauletto. "The design of an audio film for the visually impaired." (2009).

Literature, which is consumed in a spatio-temporal context, can convey meaning through the narrative qualities of sound.[9] This means that the visual channel can be enhanced or even corrected by adding appropriate auditory information. The mix of these two senses, enables the literary work with extra-textual content that brings simultaneous information about what is occurring in the narrative text. The sonic cues act as valid co-occurrences, where the sensory information coming from one event is perceived by different senses. These co-occurrences are set to produce a cross-modal interaction where interpretation of read actions is influenced by information available in the auditory sense; taking advantage of an intermodal redundancy. During these co-occurrences, the embedded multimedia operates in a cognitive ecology of multidimensional content, where the reader, assisted by the sonic information, interacts with sound and text to remember, feel, think and imagine. This literary ecology aims to resemble theatre performances, where the actor is not the solely in charge of the narrative. For instance, Shakespeare's plays utilized sounds effects and music to command attention, whereas contemporary theatre relies more on lighting. However, this theatrical ecology has been seen, until now, opposed to literature which is only associated with alphabetic writing,[10] being defined as works whose only point of reference is as written textuality.[11]

5.5 Encounters

Encounters is the series of four short stories that tell the strange situations of different characters. The four short stories are embedded with no mare than four different sound effects each that complement the narrative. As with *Sleepy Hollow*, the four short stories are presented in the Dell Venue Pro tablet, with the tobii eye-tracker and the smart glasses (See Figure 5.7). The difference with the *Sleepy Hollow* is that these *The Watchmaker, John, The Last Space* and *That Which Lives in the Attic* are very short, around four to six pages each. The series *Encounters* uses the gaze over feature and thus allows to present the text in two columns. This set of short stories is not designed to stop the sonic interventions when engaged, since they are too short and the reading lasts no more than five minutes. Instead they are designed to provide a dual coding across senses.

The smart glasses are used to find if the reader is being engaged using this dual coding. The smart-glasses are not used to intervene but to study the effectiveness of the multimodal literacy. Even though the system is the same as the one used in *Sleepy Hollow,* the use of the smart glasses can be different,. This design is to provide an example of the different application that the smart glasses can have.

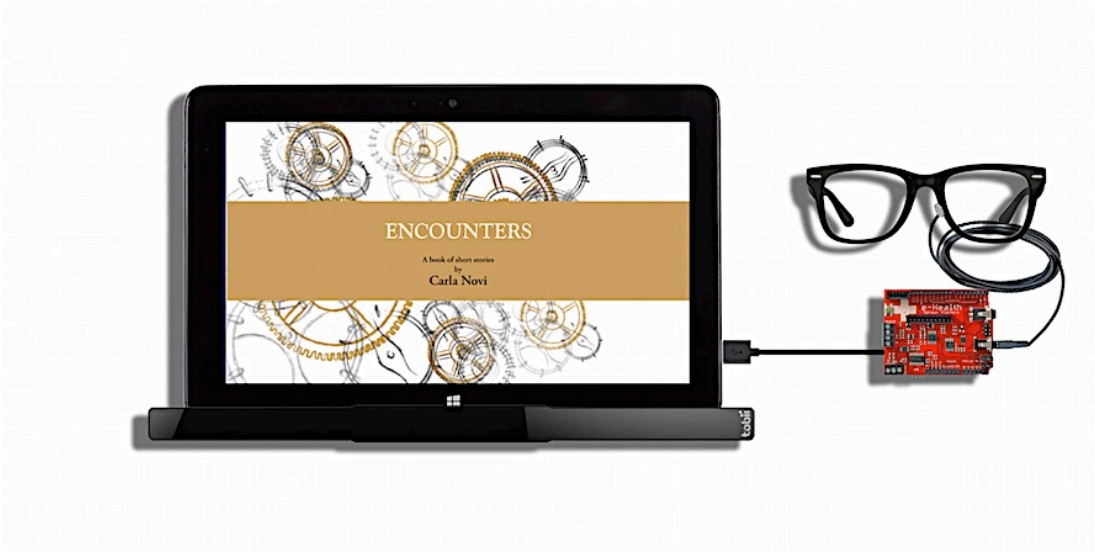


Figure 5.6: Prototype for Sound Effects.

5.6 Sleepy Hollow

Once the design of the sonic cues is completed, it can then be integrated to the system into line-breaks, to create the Augmented Narrative of Sleepy Hollow. The augmented short story is then presented to the reader in a Dell Venue 11 Pro with the tobii Eye X eye-tracking add-on. The tablet is also connected to Arduino board that uses the eHealth platform shield to connect the temperature sensor to the computer (See Figure 5.6). The sensor, mounted on the eyeglass frame, is then given to the reader to measure her engagement.

---

9. Davenport, Glorianna, Thomas Aguirre Smith, and Natalio Pincever. "Cinematic primitives for multimedia." *IEEE Computer Graphics and Applications* 11, no. 4 (1991): 67-74.

10. Finnegan, R,. “The How of Literature” Oral Tradition 20/2: 164-187. (2005)

11. Widdowson, Peter. Literature. London: Routledge (1999)*Applications* 11, no. 4 (1991): 67-74

The eye-gaze plug is used to read the eye-tracker signal, while a serial call with Arduino is used to extract the skin temperature data within the Processing IDE.

Dual-coding of verbal information

The system of Augmented Narrative is then used to design a literary storyteller that assists second language learners, specifically learners that use Chinese characters in their native language. We argue that the difference in phonological processes may cause difficulty when reading alphabetical languages, and that the redundancy of the audiobook might help these readers to extract enough information form the first pages to be able to engage through the rest of the story.

This design presents the story on the same tablet Dell Venue Pro 11 which presents the reader with the literary text and the audiobook at the same time. The smart eyeglasses with the temperature sensor are used in the same way as the initial design, only this time they will stop or continue the sonic verbal information of the audiobook (See Figure 5.7). This first second of Augmented Narrative is done with the same Processing language, but needs no use of the eye-tracking device. Each page has the auditory extract of the text content that is activated once the reader turns the page. As the auditory cue determines the reading speed, there is no need to follow the eye-gaze. However, the system still uses the temperature sensor for engagement.

Figure 5.7: Prototype for Audiobook.

5.8 Application

The system of Augmented Narrative with the smart-glasses for engagement, sound design and eye-tracking technology was used to augment five different short stories. Four of them from the same author with whom we worked closely to in order to follow her ideas, and another one that was worked from an analysis of the author and narrative. This means that the system can be applied to new literature, where it can become part of the process; or it can be applied to already existing narratives that might need help to attract readers.

Working with an author augmenting her four short stories into multimodal narratives, we found that coding extra-textual cues can be part of the creative process, adding other dimensions to the written word. It can be as simple as to scare the reader with the sound of a siren; to a landscape described not by words but by the sound of it. We found sound could be used to create more complex interaction, giving sonic clues that readers can only associate at the very end of the story. The different channels can allow the writer to expand her creativity. We find that by transforming the medium into a cognitive agent that communicates the author's work beyond the written language we can create an embodied experience. This gives the cognitive ecology the opportunity to include more stakeholders. In that sense we believe the user is the author, who can code language and non-verbal information of sound into her narrative. Augmented Narrative can be used as literary tools to create:

- Redundancy: Sounds effects can act as a redundancy when for example reading: the phone rang. The sound can complement the text by giving information about that phone: mobile, rotary dial or touch-tone dialing phone.
- Valid co-occurrences: Bring simultaneous sensory information of an event where for example, the narrative talks of an explosion.
- Inference: Many times authors give cues to their readers to infer certain situations and emotions, such as a blush or a sweaty hand to tell us a certain character is in love. These cues can be given with the sound of a palpitating heart.

- Differentiate: Sonic cues could be use to differentiate between parallel worlds or parallel times using a soundscape, or even between characters using haptic cues in a conversation.
- Ambience: The description of the story's setting is critical to understand the fictional world. It helps to establish the where and when of the story. An sonic perception of this setting soundscape could spark the readers previous experience giving the setting a public and perhaps also a private meaning.

On the other hand, the work of the author using the tools of an Augmented Narrative can create new experience for the end user: the reader. With the multimodal narrative the end user will have a different experience of the text, which allows for an embodied experience of the fictional world, which is the main focus in the present work. But in this cognitive ecology there is also the Publishing House, who can be part of an Augmented Narrative by doing the job we have done in this project, assisting authors with sound designers to meet the end user's expectations of an exciting literary experience. These publishing houses could also use the system to bring back books that have been forgotten in the backlist, re-designing them into a multimodal experience. Overall by creating a framework of cognitive ecologies we believed we have given the concept of Augmented Narrative the opportunity to integrate into the already existing editorial processes.

Evaluation

We find that sound effects can assist the reader imagery, which in turn can increase comprehension, as the sonic cues aided our readers to mentally simulate the context of the story. The subjects that were benefited by Augmented Narrative showed better comprehension than without Augmented Narrative. One of our initial hypotheses was that engagement would be directly related to comprehension, where higher levels of engagement would show higher levels of reading comprehension. However, we did not find engagement is necessary for comprehension. Our readers did not need to like the short story or showed high levels of engagement when there was good comprehension. If we think that in primary school or even high school, students can go through their classes and pass their subjects even when these subjects are not appealing. On the other hand we found that the redundancy of verbal information with the audiobook forced a top-down reading, where our subjects were encouraged to retrieve meaning using small slices of information they were able to understand, and not fixate in word by word comprehension. In conclusion we find that sonic cues of non-verbal information are better for readers with good reading skills, who can integrate the information of both channels. But, if there is not enough reading skills to retrieve meaning from the written word, then the sonic cue becomes pointless. For these low skilled readers, the redundancy of sonic verbal information is better as they are assisted in rhythm and intonation.

7.1 Smart-Glasses Evaluation

The smart-glasses were designed to transform literature into an empathetic agent that uses the reciprocal communication to intervene when the reader needs it to intervene. This new interaction took our subjects by surprise. They had to adjust reading behaviors to incorporate reciprocity in their reading experience. However, we found that after the second or third interaction all readers were able to understand what was happening and re-adapted to the new experience. The smart glasses for engagement worked on nineteen out of a total

of twenty-three participants. This represents 80% accuracy where the temperature reflected the reader's subjective immersion. Thus, we find that the temperature of the nose seems a reliable physiological metric to measure engagement through the involuntary physiological reactions of mental workload. With the improved design we also find that the skin contact is better and with less hassle. Non of the participants on our second experiment felt uncomfortable with the smart glasses, even when they had to wear them in a public space. In fact, eight out of ten found the design aesthetic.

Temperature

A statistical significance was found in the experimental setup with our stationary participants allowing us to think nose temperature can work as a indication of cognitive load and thus for engagement. However this metric of engagement is not the only one with which the concept of Augmented Narrative can work, as HRV and eye behavior could work too. Furthermore we believe temperature can also give an indication of an aversion. This is visible in an initial spike in increased temperature in the Dual-Coding study (See Figure 7.1). We believe this can be a sudden reaction such as blushing. We also found this spike in subject one who considered the sound disturbing.

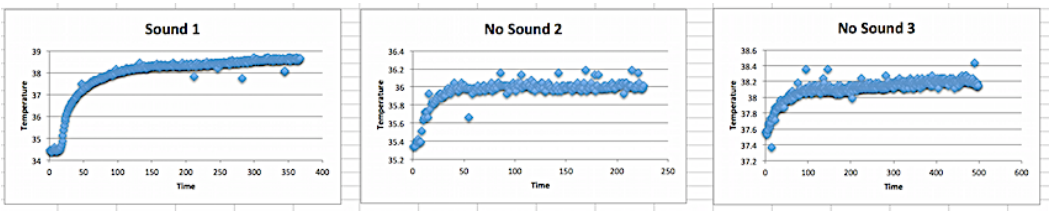


Figure 7.1: Temperature spikes.

7.2 Sonic Evaluation

The sonic assistance took most users by surprise, they didn’t know how to use this aids. However, even when the sonic cues disrupted the usual reading experience their biological minds were able modify and integrate the sonic information. The non-verbal sounds were capable to force the reader’s mind to imagine the setting of the story or even to correct their perception of the narrative, as some of them trusted more the recognizable sounds than the words, giving more weight to the sonic aid to retrieve meaning. On the negative side this sonic aids can also cause more frustration defeating the system of

Augmented Narrative, especially in low skill readers. However, we believe that we can correct this by adding to the system an algorithm that suppresses the sound if the temperature is incrementing in more than 1 degree in a short period of time.

Sonic Disruption

Our postulation was that we would be able to translate Shakespeare's grammatical disruptions into disruptions of sound, which would stimulate the reader's imagination and call for her attention. The temperature data of short story the Watchmaker, read by subject three and subject eight (Dual-Coding study), show this disruptions of sound can make the temperature go down (See Figure 7.2). In the case of subject three they were enough to keep her temperature down at the end of the reading exercise but not in the case of subject eight. In both cases subjects made long regressions after hearing the sound of the crying woman, trying to understand what is that they have missed. We believe this regression is the valley in both graphs. This is interesting data since this is what we expect to find in an application of a whole novel; where readers can loose attention and immersion at certain point of the narrative, but hopefully we can make them take long regressions and look for what they have misses increasing their curiosity for the narrative text.

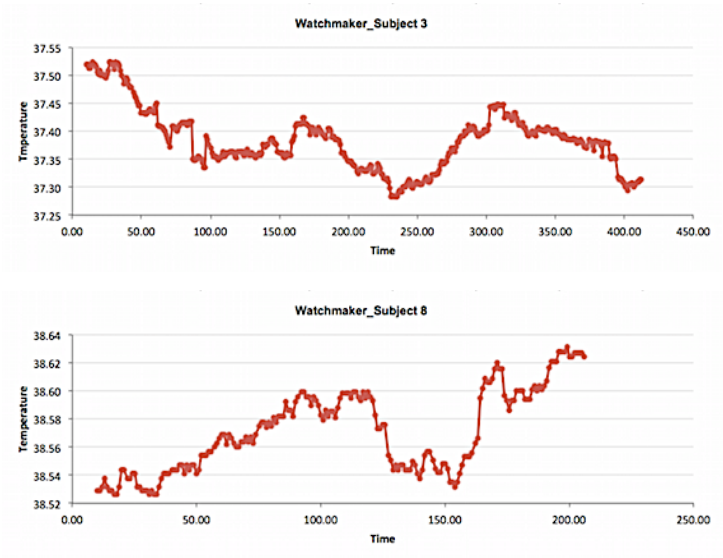


Figure 7.2: Similarities on sonic disruption.

7.3 User Scenario

The target user for the system of Augmented Narrative is defined by the

digital publishing demographics. Pew Research Center in a phone survey taken form December 2011 to January 2014 found that almost fifty percent of readers under 30 read an e-book in the pas year (See Figure 7.3). In the same line Pew Research Center's Internet Project Omnibus Survey from Jan 2011- Jan 2014, found that among e-book readers who owned e-reader, tablet, computer and cell phone, the first two devices were the most popular to read (See Figure 7.4).

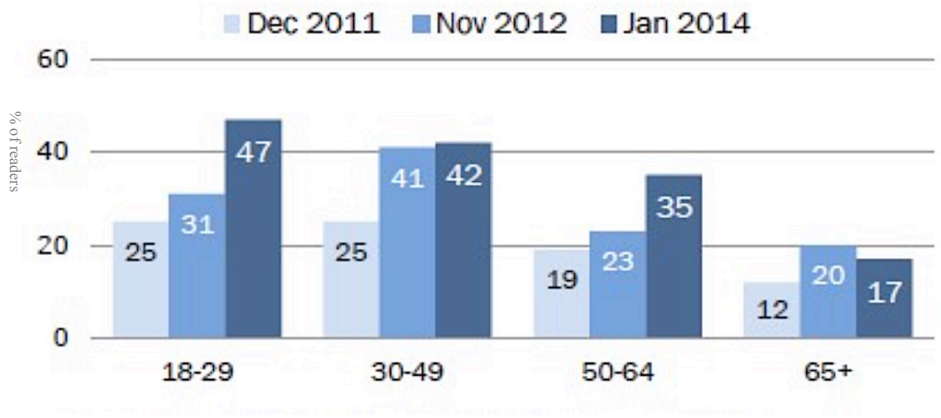


Figure 7.3: (Pew Research Center surveys, Dec2011-Jan 2014)
Percentage of readers that read an e-book in the past year.

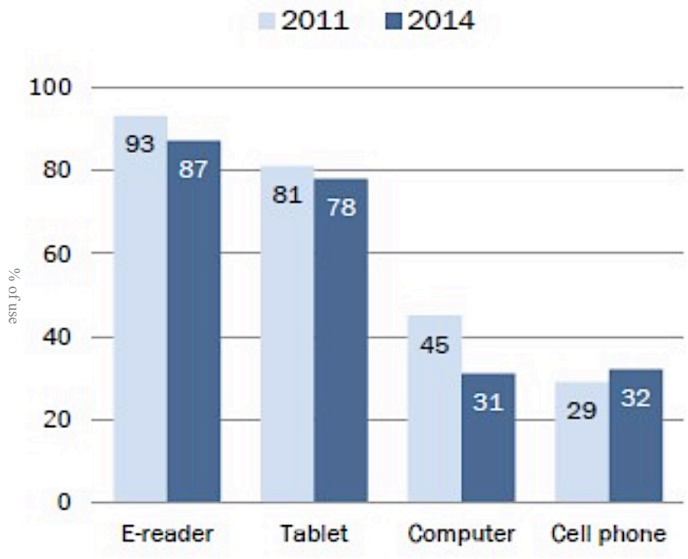


Figure 7.4: (Pew Research Center surveys, Jan 2011-Jan 2014)
Popular devices for e-books.

7.3.1 Persona

Sally is a 25-year-old architect, who lives on her own in a small

apartment in the city center (See Figure 7.4). Even though she is not up to date with the entire latest version, she owns a smart phone, a tablet and a laptop. Sally likes to read and often downloads e-books on her tablet, but rarely sits down to actually read them. She gets distracted with other applications of social media. As an architect she is educated and likes to read, but has problems finding a novel that engages her fast enough to want to keep reading, and her busy agenda does not allow her to spend time looking for books, she just relies on recommendations from friends and family.

Sally has found this new wearable device that allows e-books to know when she is not engaging with the story, intervening with sonic cues to produce an embodied experience of the fictional world. The smart-glasses (See Figure 7.5) remotely connect with her tablet to produce the multi-sensory experience of an Augmented Narrative.

Figure 7.5: Augmented Narrative's User

7.3.2 Product Presentation

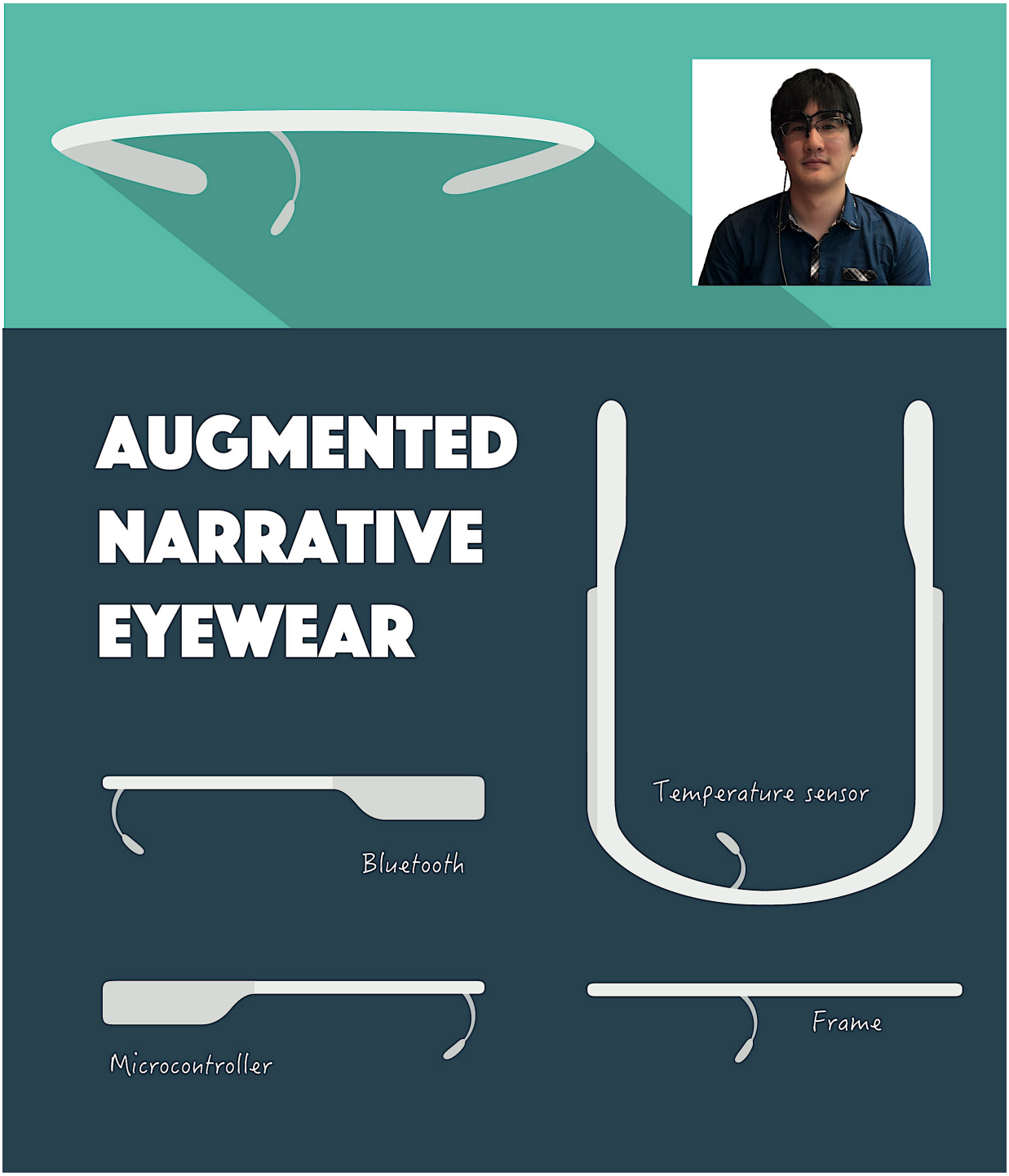


Figure 7.6: Augmented Narrative's smart glasses

7.3.3 Market Integration

The technology positions itself in the already exiting market for

publishers that have gone digital. It targets reader's that feel comfortable with digital content as well as readers that have not yet transitioned. The new reading experience is designed to lure users into reading more, enjoy and immerse into the written word. In that sense, Augmented Narrative does not change the interface of digital books, or how they are consumed. It is set to remodel reading behavior, by changing the relationship with the books, creating an interactivity that enriches literary storytelling.

Figure 7.7: Consumer's behavior

7.3.4 Benefits

Augmented Narrative can also be beneficial for publishers and digital distributors to follow and understand the reading behavior by following the

reader's experience. This means companies do not need to rely in subjective feedback, but continuously monitor the market to improve and provide a better product. The fact that a book with Augmented Narrative requires sound design means it can create employment within the publishing industry. A publishing house could create a department that incorporates Augmented Narrative to the editorial process.

Augmented Narratives benefits go beyond the reader's own experience:

- Personal communication with the book
- Control attention
- Increase imagery
- Improve comprehension

7.4 Limitations

Sound can make the temperature go up if there is an aversion to it, so the disruption of sound will not work to increase immersion, as the aversion to it will only create an unwanted distraction. But the system is designed to provide sound when temperature goes up, and thus acting against the purpose of aiding the reader. Another limitation is the different hardware needed for this application as the eye-tracker can stop working if the reader forward or backward. This can cause the application to fail and provide the sound. Regarding our smart-glasses at this point we still have a noisy data that needs to be cleaned in an average over time. This will give the system stability to decide on the temperature trend of each reader.

7.5 New Hypothesis

From the temperature data we found three cases where the plot is not consistent with what we have found before. In this research we were looking for a clear line going up or down, even stable for some periods. However, these three cases in the Multimodal Stimuli study suggest that we might also be able to determine if there is a top-down or bottom-up reading (See Figure 7.7).

In our first study of Dual-Coding we only found temperature data that seems to be a top- down approach, where subjects make sense of the narrative as a whole. In this form of reading we can find a mental workload where the

temperature goes up or down. However, we believe these three cases are a bottom up reading. This might not be the case in a bottom up reading, where the subject might concentrate in reading perfectly word by word. This effort in attention of word-by-word changes the dynamic of the temperature plot.

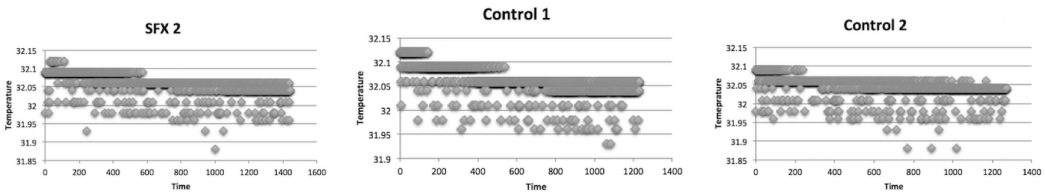


Figure 7.8: Example of Bottom-Up Reading.

Unfortunately we were not able to verify this hypothesis within this study as we were not expecting to have found such a difference in the visualization of the nose temperature. Nonetheless, we believe we might be able to corroborate this hypothesis using the eye-tracker, where we trust we will find long fixations all across the text.

Chapter 8

Conclusion

To be able to present the concept of an empathetic literary narrative, the design of the system Augmented Narrative has followed a historical progression in a co-evolutionary process between literature and technology using familiar instruments and spaces to exploit the brain's natural strengths. Following Gutenberg's approach, where the success of print rests in preserving the quality of scripts, as well as solving production strains, Augmented Narrative looks for a sensible interaction that preserves and at the same time enhances literary narratives. This interaction with sound, changes the paradigm and opens the reader's cognitive system to new possibilities. It improves reading comprehension by building higher correlations of distributed perceptual modalities that stimulate senses other than the visual, without modifying the written word.

With print and the emergence of the novel, academic texts and fictional narratives started differentiating themselves from each other by developing different relations with their oral past. Emerging, in our view, into two distinct areas to design digital interactions. It was only by researching into Literary Criticism that the differences between fiction and non-fiction literature became evident. For instance, in the digital age, academic texts connect to a primary orality by going back to rhapsody-like episodes in a hyper-textual navigation of the world wide web, where a vast quantity of information is linked. This simulates oral cultures where acquired knowledge was constantly repeated stitching together prefabricated parts to effectively administrate wisdom in memory.

These rhapsodies were not linear, and thus, are part of a hyper-textual structure that adapted to particular audiences encouraging them to respond vigorously to skilled narrators with the ability to adjust to different situations. In the same way, digital content adjusts to students' needs building knowledge by linking information, imitating the hyper-textual structure of the rhapsody. These literary texts come from academic rhetoric, that disconnected from oral cultures separating readers from the known, disengaging them from the natural world, where stories of human action were used to store, organize and communicate knowledge, based on real life experiences. These narratives were performed to audiences achieving a close empathetic communal identification, but academic rhetoric drifted oral rhetoric skills into writing. We can see the influence of these hyper-textual structures in Interactive Storytelling.

In short, academic rhetoric became the literary style of nineteenth century, with the exception of female author. Female authors, not being trained in academic studies, expressed themselves in a less oratorical voice, giving raise to the novel; which is more like a conversation and less of a performance making it more appropriate to the commercial essence of print. The novel, unlike academic texts, raised from print and sets itself apart from episodic structures by developing into linear texts that represent the word of the author in final form; since print is comfortable only with finality. Tight plotting is extended into lengthy narratives with a sense of closure in a fixed point of view, providing a shared understanding with the reader. It is here that Augmented Narrative reconnects writer and reader, keeping the text-enclose shared understanding and imitating, at the same time, oral narrators who interacted freely with a live audience. Based on the fact that societies were conceived with the assistance of oral speech, becoming literate only late in history, we find that human communication and thought are strongly relate to sound, but also involve other senses like touch, and smell. Fortunately, literature is infinitely adaptable, even though it obliterated its oral antecedents, allows for the reconstruction of the consciousness to cope with the multimodality of an interactive text.

The system reinvents storytelling in digital format using remediation, where new technologies take the place of the old ones in a cultural competition among media; reorganizing the characteristics of writing and reading in order to reform the text's cultural space. However, in digital societies change is the rule and stability the exception, proving digital technologies to be one of the most traumatic remediation for text. They change the look and feel of writing and reading continuously, making it important for any remediation to follow a coevolutionary process between narrative and media to provide the correct solution. For this reason we believe there must be a balance in integrating literacy and orality respecting the writer's work and preserving the linearity of the narrative on the bases that no technology can serve as a writing space in the absence of human writers or readers; even when it may predispose a cultural construction of natural writing. Thus we make use of new technologies to reconfigure the practices of writing and the ideal this practices want to express. Augmented Narrative redesigns the relationship between author, book and reader, putting them in a cognitive ecology where the book becomes a cognitive agent that understands and interacts with its reader. Culturally, written language has defined and empowered new groups of reader. For instance, the eighteen-century's growing English middle class with the novel. In the same way Augmented Narrative's endeavor is to define a multimodal digital content that empowers digital readers. This new reading experience present itself to existing groups of reader by delivering the known linear narrative, and to new groups that find multimodal interactivity useful to reconnect with literature.

8.1 Further Applications

Taking the idea of dual coding across senses further, in a collaboration work for another application, we decided to redesign and adapt our framework of multi sensory literacy, but this time putting more emphasis on intermodal redundancy to give information about context. For this application we used one of our original short stories by Carla Novi, *John*. The short story was augmented with multi-sensory cues that communicate an atmosphere of a rainy day (See Figure 8.1).

John is having the worst day of his life, and this is reflected in the weather. Even though the text does not explicitly tells the reader it is raining, this can be inferred from the words "John' s shirt was soaked". In this way the interaction is used as a literary tool to express emotions through a multimodal experience of the narrative. This emotional aspect of the story is reinforced by haptic visual and auditory cues of raindrops. The cross-modal cue of rain was design as an 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.

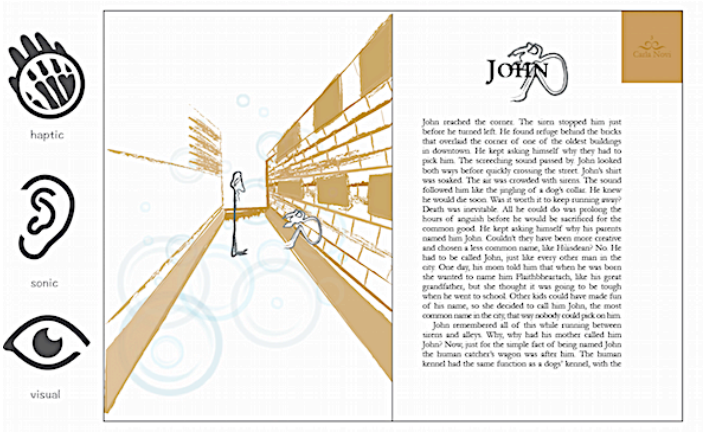


Figure 8.1: Haptic, sonic and visual literature

In this prototype the text and illustrations of the short story are shown in a double page format. The visual cues are the only sensory cues that are divided into verbal and non-verbal. While the text is visual verbal information, the illustrations are non-verbal. Adding to these non-verbal cues is the animation of raindrops that seem to be falling into the pages of the book. These animation change from left to right page, depending on which page the reader's eye gaze is - not to disturb the reading task. In order to reinforce the animation of rain drop sound of rain is also embedded to the story. The sound of the rain with the visual effects of the raindrops on the pages starts reinforcing the idea that within John's fictional world it is in fact raining. Finally in addition to the a haptic interaction was added using the iPad prototype of an e-Reader for Multimodal Literacy, designed by Heng Gu (See Figure 8.2). [1]

---

1. Heng, Gu, Sanchez, S., Kunze, K., Masahiko, I. "An Augmented e-Reader for Multimodal Literacy." In proceedings for Ubicomp (2015)

Figure 8.2: Haptic add on for multimodal literacy

This prototype made it possible to incorporate our framework of dual- coding across senses, adding to it the haptic stimulation. This haptic stimulation was designed to work together with sound and visual effects and reinforce the concept of a rainy day. The hand feels the sensation of raindrops hitting her hand while holding the iPad's add-on that contains the transducer. The large surface transducer is used to simulate sensation and textures through a series of impulses. A wide variety of haptic effects can be achieved through the same method, where textures can be simulated through vibration patterns. Beyond simulating existing textures, haptic stimulation could also create compelling tactile experiences that don't naturally occur.

The iPad is used to present the user with the short story, while at the same time it serves to power the multi-sensory add-on for sound and tactile cues. The Processing language is used to display the pages and generate the vibration and audio signals. The signals for sound and haptic cues are split between the two stereo channels: the left to achieve the tactile effect, and the right channel for the auditory signal. This is because the haptic stimulus, as the sound stimuli, is created with .wav files. The Processing applet pulls from a library of prerecorded .wav files that were specially designed for the split haptic and audio channels. The multi-sensorial reading-device, uses a large surface vibration transducer to simulate haptic sensations and textures through vibration patterns, while the small speaker provides auditory stimuli. Both connect to the iPad mini through a stereo jack, which splits the input into vibration and audio channels. The iPad camera is used to detect the reader's relative page position: left or right to change and interact according to progress.

In a first evaluation we showed and discussed the augmented story and the initial prototype that includes haptic sensations, with three Human Computer Interaction experts. All three experts have been within this field of researchers for over ten years; one of them with background in haptic sensation. Their feedback was in general positive. They believe the application can support and enhance immersion in literature. To study on how the implementation of this application works for the general public, we proposed to apply the same immersion and hedonic questionnaires, plus the retell interview. We argue that by providing the reader with a mix multiple sensory-non-verbal extra-textual cues, we can continue to enhance the reading experience, increasing mental imagery by adding haptic feedback.

Overall, the discussed application with haptic interactions, serves to illustrate the possibilities of our concept of a dual coding across senses. The interaction created between textual sonic and tactile information serves to communicate a deeper message and allow other senses to be used as literary tools to create a narrative, bringing ear and tact into a storytelling that until now has only allowed the eyes.

The finality of these multimodal interventions has been to translate what is found in oral communications, where language is not the only channel capable to hold meaning. In oral communications we hear, see and feel what is happening, and put all the information together to communicate and understand others. With Augmented Narrative this multi-sensory communication can be pass on through literature.

8.2 Future Work

Since the idea of the incorporation of sounds is for it to be an unexpected stimulation, just as Shakespeare did with his verbing, we would like to find evidence of this through the study of event related potentials (ERP). The hypothesis for future work is that through Augmented Narrative we are translating Shakespeare's grammatical disruptions into disruptions of sound. Thus bringing his literary tricks to stimulate his readers to other authors, to apply spontaneous stimulations without having to change their writing style. This multi-sensory literary tool can bring new ideas to writers to find new way of storytelling that respect their individual narrative work and capabilities. We believe we can do this with accessible electroencephalography technology to try and do EEG recordings. Using the Emotiv [2] headset and software we believe we can find and document the event related potential (ERP) P300 while reading the augmented literary text. The study will look for the P300 instead of the P600 found in Shakespeare's study since P600 relates to unexpected linguistic stimulating and our objective is to find unexpected non-linguistic stimulations. The P300 appears across modalities where perceptual distracters such as sound, occur during the visual stimuli of words, and has been conventionally elicited using the oddball paradigm, where visual or sonic simulations evaluate the subject's reactions to unpredictable yet recognizable events.

---

2. "Emotiv" http://www.emotiv.com/

In addition, previous studies have revealed that memory and attentional mechanisms are connected in processing perceived information, where attentional processes updates memory processes. [3,4,5,6] Thus, engagement during reading also contributes to the P300 amplitudefrom the pre-stimulus baseline due to the processing demands of perception. To find the P300 this study will follow the steps of Campbell [7] who used the Emotiv (See Figure 8.3) to design NeuroPhone, a brain-mobile-phone interface in a task stimulus exercise where the phone flashes a sequence of photos of contacts from the address book while the EEG looks for a P300 wave, sign that the shown matches the person whom the user wishes to dial.

Figure 8.3: Emotive Epoc.

---

3. Chapman, Robert M., and Henry R. Bragdon. "Evoked responses to numerical and non-numerical visual stimuli while problem solving." (1964): 1155-1157.

4. Sutton, Samuel, Margery Braren, Joseph Zubin, and E. R. John. "Evoked-potential correlates of stimulus uncertainty." Science 150, no. 3700 (1965): 1187-1188.

5.Sutton, Samuel, Patricia Tueting, and Joseph Zubin. "Information delivery and the sensory evoked potential." Science (1967).

6. Polich, John. "Updating P300: an integrative theory of P3a and P3b." Clinical neurophysiology 118, no. 10 (2007): 2128-2148.

7. Campbell, Andrew, Tanzeem Choudhury, Shaohan Hu, Hong Lu, Matthew K. Mukerjee, Mashfiqui Rabbi, and Rajeev DS Raizada. "NeuroPhone: brain-mobile phone interface using a wireless EEG headset." In Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds, pp. 3-8. ACM, 2010.

To acquire the P300 the electrodes are placed in the 10-20 system and the data average to increase the signal to noise ratio (SNR) and filtered with a 0-9 Hz band pass over multiple trials. In this way Campbell found how to hack the Emotiv headset to find the ERP needed. If finding the P300 proves to be unattainable the study can use the Emotive's Affectiv™ Suite to monitor the user's emotional states in real-time: engagement, frustration and excitement.

Similarly to these studies for future work we would like to look into ERP to prove or disprove the disruptions of sound in the system of Augmented Narrative can create the same brain stimuli as Shakespeare's grammatical disruptions (See Figure 8.4).

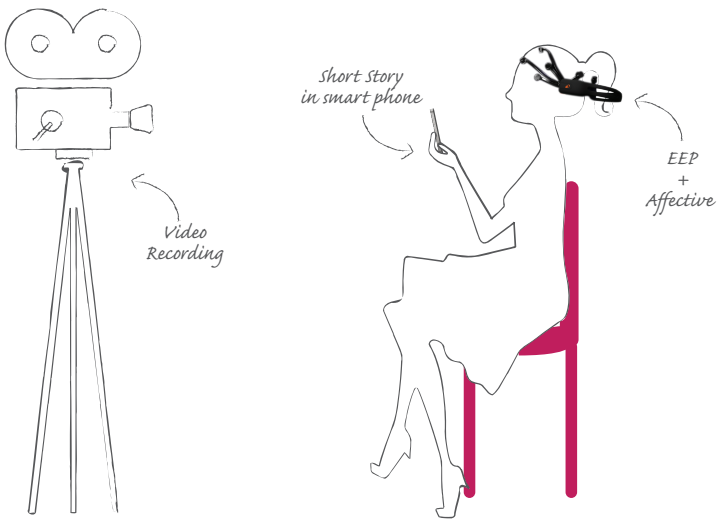


Figure 8.4: Proposed future study EEG.

Bibliography

Adelmann, Pamela K., and Robert B. Zajonc. "Facial efference and the experience of emotion." *Annual review of Psychology* 40, no. 1 (1989): 249-280.

Angelotti, Michael, Ralph R. Behnke, and Larry W. Carlile. "Heart rate: a measure of reading involvement." *Research in the Teaching of English* 9, no. 2 (1975): 192-199.

Ashby, J., and Rayner, K. (2012). Reading in alphabetic writing systems: evidence from cognitive neuroscience. *Neuroscience in Education: The good, the bad, and the ugly*, 61.

Bal, P. Matthijs, and Martijn Veltkamp. "How Does Fiction Reading Influence Empathy? An Experimental Investigation on the Role of Emotional Transportation." *PLOS ONE* 8, no. 1: e55341. (2013)

Beierl, Barbara Hardy. "The sympathetic imagination and the human–animal bond: Fostering empathy through reading imaginative literature." *Anthrozoos: A Multidisciplinary Journal of The Interactions of People and Animals* 21, no. 3 (2008): 213-220.

Berka, Chris, Daniel J. Levendowski, Michelle N. Lumicao, Alan Yau, Gene Davis, Vladimir T. Zivkovic, Richard E. Olmstead, Patrice D. Tremoulet, and Patrick L. Craven. "EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks." *Aviation, space, and environmental medicine* 78, no. Supplement 1 (2007): B231-B244.

Bertelson, Paul, and Béatrice de Gelder. "The psychology of multimodal perception." *Crossmodal space and crossmodal attention* (2004): 141-177.

Biedert, Ralf, Georg Buscher, Sven Schwarz, Jörn Hees, and Andreas Dengel. "Text 2.0." In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pp. 4003-4008. ACM, 2010.

Buscher, G., Dengel, A., and van Elst, L. Eye movements as implicit relevance

feedback. In Ext. Abs. CHI 2008 (2008), 2991–2996.

Biedert, Ralf, Georg Buscher, Sven Schwarz, Manuel Möller, Andreas Dengel, And Thomas Lottermann. "The text 2.0 framework." In Workshop on Eye Gaze in Intelligent Human Machine Interaction, pp. 114-117. 20105.

Campbell, Andrew, Tanzeem Choudhury, Shaohan Hu, Hong Lu, Matthew K. Mukerjee, Mashfiqui Rabbi, and Rajeev DS Raizada. "NeuroPhone: brain-mobile phone interface using a wireless EEG headset." In Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds, pp. 3-8. ACM, 2010.

Chang, Yu-Chieh, and Shwu-Lih Huang. "The influence of attention levels on psychophysiological responses." International Journal of Psychophysiology 86, no. 1 (2012): 39-47.

Chapman, Robert M., and Henry R. Bragdon. "Evoked responses to numerical and non-numerical visual stimuli while problem solving." (1964): 1155-1157.

Chion, M. 1994. Audio-vision. New York: Columbia University Press.

Cleary, Anne M., Anthony J. Ryals, and Jason S. Nomi. "Intuitively detecting what is hidden within a visual mask: Familiar–novel discrimination and threat detection for unidentified stimuli." *Memory and cognition* 41, no. 7:

989-999 (2013)

Coltheart M, Curtis B, Atkins P, Haller M (1993): Models of reading aloud: dual-route and parallel-distributed-processing ap- proaches. *Psychol Rev* 100:589–608.

Corradini, A., and Antonietti, A. "Mirror neurons and their function in cognitively understood empathy." *Consciousness and cognition* (2013).

Coulson, Seana , King, Jonathan W. and Kutas, Marta : Expect the Unexpected:

Event-related Brain Response to Morphosyntactic Violations, *Language and Cognitive Processes,* 13:1, 21-58 (1998)

Csikszentmihalyi, Mihaly, and Mihaly Csikzentmihaly. *Flow: The psychology of optimal experience*. Vol. 41. New York: HarperPerennial, 1991.

Davenport, Glorianna, Thomas Aguirre Smith, and Natalio Pincever. "Cinematic primitives for multimedia." *IEEE Computer Graphics and*

Applications 11, no. 4 (1991): 67-74.

Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. Nat Rev Neurosci 2010; 11: 201--211.

Dobra, Susan. “The gift of Theuth: Plato on writing.” Retrieved from: www.csuchico.edu/phil/sdobra _mat/platopaper.html. September 2nd, 2014.

Eason, Sarah, Lindsay Goldberg, Katherine Young, Megan Geist, and Laurie Cutting. "Reader-Text Interactions: How Differential Text and Question Types Influence Cognitive Skills Needed for Reading Comprehension." *Journal of Educational Psychology* 104, no. 3 (2012): 515-528.

Fairclough, Stephen H., Louise Venables, and Andrew Tattersall. "The influence of task demand and learning on the psychophysiological response."International Journal of Psychophysiology 56, no. 2 (2005): 171-184.

Finnegan, R,. “The How of Literature” Oral Tradition 20/2: 164-187. (2005)

Fludernik, Monika. *Towards a'natural'narratology*. London: Routledge, 1996.

Folkerth, Wes. *The sound of Shakespeare*. Routledge, 2014.

Fournier, Lisa R., Glenn F. Wilson, and Carolyne R. Swain. "Electrophysiological, behavioral, and subjective indexes of workload when performing multiple t tasks: manipulations of task difficulty and training."International Journal of Psychophysiology 31, no. 2 (1999): 129-145.

Gambrell, Linda B., Barbara Martin Palmer, Rose Marie Codling, and Susan Anders Mazzoni. "Assessing motivation to read." *The Reading Teacher* (1996): 518-533.

Gilroy, Stephen, Julie Porteous, Fred Charles, and Marc Cavazza. "Exploring passive user interaction for adaptive narratives." In Proceedings of the 2012 ACM international conference on Intelligent User Interfaces, pp. 119-128. ACM, (2012)

Giordano, Bruno L., and Federico Avanzini. “Perception and Synthesis of Sound-Generationg Materials”. In *Multisensory Softness*, pp. 49-84. Springer London, 2014.

Gladwell, Malcolm. *Blink: The power of thinking without thinking*. Back Bay Books, 2007

Gladwell, Malcolm, and Madelon E. Ruiter. "Blink: the power of thinking without thinking." *Gedragstherapie* 41, no. 2 : 199. 2008

Goldman, A. (2005). Imitation, mind reading, and simulation. In S. Hurley, and N. Chater, Perspectives on Imitation II (pp. 80-81). Cambridge, MA: MIT Press

Gouvea, Ana Cristina. How to examine the P600 using language theory: what are the syntactic processes reflected in the component. *Revista do Programa Pos-Graduaçåo en Linguística da Universidad Federal do Río de Janeiro* Vol 7 (2): 88-97 (2011)

Goodman, K. (1967). Reading: A psycholinguistic guessing game. *Journal of \ the Reading Specialist*, 6, 126-135.

Graesser, Arthur C., and Danielle S. McNamara. "Computational analyses of multilevel discourse comprehension." *Topics in Cognitive Science* 3, no. 2 (2011): 371-398.

Green, Melanie C., Sheryl Kass, Jana Carrey, Benjamin Herzig, Ryan Feeney, and John Sabini. "Transportation across media: Repeated exposure to print and film." Media Psychology 11:4 (2008): 512-539.

Guo, Fu, Yaqin Cao, Yi Ding, Weilin Liu, and Xuefeng Zhang. "A Multimodal Measurement Method of Users' Emotional Experiences Shopping Online."*Human Factors and Ergonomics in Manufacturing and Service Industries* (2014).

Harte, Bret. *The Rise of the Short Story*. Anmol Pub., 1990.

Hassenzahl, Marc. "The effect of perceived hedonic quality on product appealingness." *International Journal of Human-Computer Interaction* 13, no. 4 (2001): 481-499.

Havelock, Eric A. *Preface to plato*.Belknap Press, Harvard University Press, 1963

Havelock, Eric A. "The oral composition of Greek drama." *Quaderni Urbinati di Cultura Classica* (1980): 61-113.

Healy, Sean Desmond. *Boredom, self, and culture*. Rutherford, NJ: Fairleigh Dickinson University Press, 1984.

Helms, N.R. "Conceiving Ambiguity: Dynamic Mindreading in Shakespeare's Twelfth Night." *Philosophy and Literature* 36, no. 1 : 122-135. (2012)

Heng, Gu, Sanchez, S., Kunze, K., Masahiko, I. "An Augmented e-Reader for Multimodal Literacy." In proceedings for Ubicomp (2015)

Herman, David. "Narratology as a cognitive science." *Image and Narrative* 1 (2000): 1

Houix H, Lemaitre G, Misdariis N, Susini P, Urdapilleta I (2012) A lexical analysis of environmental sound categories. J Exp Psychol Appl 18:52

Hug, Daniel. "Investigating narrative and performative sound design strategies for interactive commodities." *In Auditory Display*, pp. 12-40. Springer Berlin Heidelberg, (2010).

Hutchins, Edwin. "Cognitive ecology." *Topics in Cognitive Science* 2, no. 4 (2010): 705-715.

I. Stephanos, S. Ebisch, T. Aureli, D. Bafunno, H. Alexi Ioannides, D. Cardone, B. Manini, G. L. Romani, V. Gallese, and A. Merla. "The Autonomic Signature of Guilt in Children: A Thermal Infrared Imaging Study." *PloS one* 8, no. 11 (2013): e79440.

Ioannou, Stephanos, Vittorio Gallese, and Arcangelo Merla. "Thermal infrared imaging in psychophysiology: potentialities and limits." *Psychophysiology* 51, no. 10 (2014): 951-963.

Irwin, D. E., and Thomas L. E. "Eyeblinks and cognition." *Tutorials in visual cognition* (2010): 121-141.9

Jalilehvand, M.(2012) "The Effects of Text Length and Picture on Reading Comprehension." *Asian Social Science* 8 (3): 329.

Jennett, Charlene, Anna L. Cox, Paul Cairns, Samira Dhoparee, Andrew Epps, Tim Tijs, and Alison Walton. "Measuring and defining the experience of immersion in games." *International journal of human-computer studies* 66, no. 9 (2008): 641-661.

Julieta Lopez, Mariana, and Sandra Pauletto. "The design of an audio film for the visually impaired." (2009).

Kadir, Djelal. "Comparative Literature in a World Become Tlon." *Comparative Critical Studies* 3, no. 1 (2006): 125-138.

Kaye, Deena, and James LeBrecht. *Sound and Music for the Theatre: The Art and Technique of Design*. Taylor and Francis, 2013.

Keller, Johannes, Herbert Bless, Frederik Blomann, and Dieter Kleinböhl.

"Physiological aspects of flow experiences: Skills-demand-compatibility effects on heart rate variability and salivary cortisol. *Journal of Experimental Social Psychology* 47, no. 4 : 849-852. (2011)

Kim, Young-Suk, Richard K. Wagner, and Elizabeth Foster. "Relations among oral reading fluency, silent reading fluency, and reading comprehension: A latent variable study of first-grade readers." *Scientific Studies of Reading* 15, no. 4 (2011): 338-362.

King, Andrew J., and Gemma A. Calvert. "Multisensory integration: perceptual grouping by eye and ear." *Current Biology* 11, no. 8 (2001): R322-R325.

Kintsch, Walter. "The role of knowledge in discourse comprehension: a construction-integration model." *Psychological review* 95, no. 2 (1988): 163.

Klingner, Jeff, Barbara Tversky, and Pat Hanrahan. "Effects of visual and verbal presentation on cognitive load in vigilance, memory, and arithmetic tasks."*Psychophysiology* 48, no. 3 (2011): 323-332.

Knapp, R. Benjamin, Jonghwa Kim, and Elisabeth André. “Physiological signals and their use in augmenting emotion recognition for human–machine interaction.” Emotion-Oriented Systems (2011): 133-159

Kort, Barry, and Rob Reilly. "Theories for deep change in affect-sensitive cognitive machines: A constructivist model." *Journal of Educational Technology and Society* 5, no. 4 (2002): 56-63.

Kuhl, J. (1986). Introduction. In J. Kuhl and J.W. Atkinson (eds.) *Motivation, thought, and action* (pp. 1-6) New York: Praeger.

Kunze, Kai, Hitoshi Kawaichi, Kazuyo Yoshimura, and Koichi Kise. "The wordometer--estimating the number of words read using document image retrieval and mobile eye tracking." In *Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, pp. 25-29. IEEE, 2013.

Kunze, Kai, Susana Sanchez, Tilman Dingler, Olivier Augereau, Koichi Kise, Masahiko Inami, and Terada Tsutomu. "The augmented narrative: toward estimating reader engagement." In *Proceedings of the 6th Augmented Human International Conference*, pp. 163-164. ACM, 2015.

Kuo, Yu-Chen, and Chih-Chi Tseng. "Recognizing the emotion of learners by physiological sensors to improve English learning performance." In

Biomedical Engineering and Informatics (BMEI), 2011 4th International Conference on, vol. 4, pp. 2152-2156. IEEE, (2011)

Lemaitre G, Heller LM (2013) Evidence for a basic level in a taxonomy of everyday action sounds. Exp Brain Res 226(2):253–264

Leslie, Lauren, and Jo Anne Caldwell. "19 Formal and Informal Measures of Reading Comprehension." *Handbook of research on reading comprehension* (2009): 403.

Lindsay, Peter H., and Donald A. Norman. *Human information processing: An introduction to psychology*. Academic Press, 2013.

Lyn F., and K. Rayner. "Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences." *Cognitive psychology* 14, no. 2 (1982): 178-210.

Mahlke, Sascha, and Michael Minge. "Consideration of multiple components of emotions in human-technology interaction." In *Affect and emotion in human-computer interaction*, pp. 51-62. Springer Berlin Heidelberg, 2008.

Mansfield, Brenda E., Bruce E. Oddson, Josee Turcotte, and Roger T. Couture. "A possible physiological correlate for mental flow." *The Journal of Positive Psychology* 7, no. 4: 327-333. (2012)

Matthews, Gerald, Sian E. Campbell, Shona Falconer, Lucy A. Joyner, Jane Huggins, Kirby Gilliland, Rebecca Grier, and Joel S. Warm. "Fundamental dimensions of subjective state in performance settings: task engagement, distress, and worry." *Emotion* 2, no. 4 (2002): 315.

Mayer, R. E. (1984). Aids to text comprehension. *Educational Psychologist*, 19, 30-42.

Mayer, Richard E. "Multimedia learning: Are we asking the right questions?."*Educational psychologist* 32, no. 1 (1997): 1-19.

McAdams, Stephen Ed, and Emmanuel Ed Bigand. "Thinking in sound: The cognitive psychology of human audition." In *Based on the fourth workshop in the Tutorial Workshop series organized by the Hearing Group of the French Acoustical Society.* Clarendon Press/Oxford University Press, 1993.

McCombs B.L. (1991). Unraveling motivation: New perspectives from research and practice. *Journal of Experimental Education, 60,* 3-88.

McCraty, Rollin, Mike Atkinson, Dana Tomasino, and Raymond Trevor Bradley. "The coherent heart: heart-brain interactions, psychophysiological coherence, and the emergence of system-wide order." *Integral review* 5, no. 2 (2009): 10-115.

McCraty, Rollin, Mike Atkinson, William A. Tiller, Glen Rein, and Alan D. Watkins. "The effects of emotions on short-term power spectrum analysis of heart rate variability." *The American journal of cardiology* 76, no. 14 (1995): 1089-1093

McLuhan, Marshall. *The Effect of the Printed Book on Language in the 16th Century*. Gingko Press, 2005.

McLuhan, Marshall. "The Gutenberg Galaxy." *Toronto, UTP* (1963).

McLuhan, Marshall, and W. Terrence Gordon. *The classical trivium: The place of Thomas Nashe in the learning of his time*. Gingko Pr Inc, 2006.

McNorgan, Chris. "A meta-analytic review of multisensory imagery identifies the neural correlates of modality-specific and modality-general imagery." *Frontiers in human neuroscience* 6 (2012).

Metz, C. 1977. Essais Se´miotiques. Paris: Klincksieck.

Michaels, Claire F., and Claudia Carello. *Direct perception*. Englewood Cliffs, NJ: Prentice-Hall, 1981.

Mibillard, Amanda. Shakespeare in Print. *Shakespeareonline*. 20 August 2004. Retrieved from: www.shakespeare-online.com/ biographyshakespeareinprint.html.

Middleton, Joyce Irene. "Orality, Literacy, and Memory in Toni Morrison's Song of Solomon." *College English* (1993): 64-75.

Moreno, Roxana, and Richard E. Mayer. "Verbal redundancy in multimedia learning: When reading helps listening." *Journal of Educational Psychology* 94, no. 1 (2002): 156.

Or, Calvin KL, and Vincent G. Duffy. "Development of a facial skin temperature-based methodology for non-intrusive mental workload measurement."*Occupational Ergonomics* 7, no. 2 (2007): 83-94.

Paivio, Allan. Mental representations: A dual coding approach. Oxford University Press, 1990.

Pauletto, Sandra. "Film and theatre-based approaches for sonic interaction design." *Digital Creativity* 25, no. 1 (2014): 15-26.

Parmentier, Fabrice BR. "The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review." *Psychological research* 78, no. 3 (2014): 321-338.

Parry, Milman, and Adam M. Parry, eds. *The making of Homeric verse: The collected papers of Milman Parry*. Oxford University Press, 1987.

Penke, L., S. Muñoz Maniega, M. E. Bastin, MC Valdes Hernandez, C. Murray, N. A. Royle, J. M. Starr, J. M. Wardlaw, and I. J. Deary. "Brain white matter tract integrity as a neural foundation for general intelligence." *Molecular psychiatry*17, no. 10 (2012): 1026-1030.

Penney, C. G. (1989). Modality effects and the structure of short-term verbal memory. *Memory and Cognition, 17,* 398–422.

Perlovsky, Leonid I., and Roman Ilin. "Mirror neurons, language, and embodied cognition." *Neural networks.* (2013)

Phelan, James, and Peter J. Rabinowitz. *Understanding narrative*. The Ohio State University Press, 1994

Polich, John, ed. *Detection of change: event-related potential and fMRI findings*. Springer, 2003.

Polich, John. "Updating P300: an integrative theory of P3a and P3b." Clinical neurophysiology 118, no. 10 (2007): 2128-2148.

Pollard, Alfred W. *Shakespeare's Fight with the Pirates and the Problems of the Transmission of His Text*. Cambridge University Press, 2010.

Pratt, Mary Louise. "The short story: The long and the short of it." *Poetics* 10, no. 2 (1981): 175-194.

Purves, Alan C., and Victoria Rippere. "Elements of Writing About a Literary

Rabinowitz, Peter J. *Before Reading: Narrative Conventions and the Politics of Interpretation. The Theory and Interpretation of Narrative Series*. Ohio State University Press, 1070 Carmack Road, Columbus, OH 43210-1002, 1997 Work: A Study of Response to Literature." (1968)

Rani, Pramila, Changchun Liu, Nilanjan Sarkar, and Eric Vanman. "An empirical study of machine learning techniques for affect recognition in human–robot interaction." *Pattern Analysis and Applications* 9, no. 1 (2006): 58-69.

Rayner, Keith, Kathryn H. Chace, Timothy J. Slattery, and Jane Ashby. "Eye movements as reflections of comprehension processes in reading." *Scientific Studies of Reading* 10, no. 3 (2006): 241-255.

Royer, James M., C. Nicholas Hastings, and Colin Hook. "A sentence verification technique for measuring reading comprehension." *Journal of Literacy Research* 11, no. 4 (1979): 355-363.

Russell, James " A circumflex model of affect". Journal of personality and Social Psychology 39: 1161-1178 (1980)

Sadoski, Mark. "An exploratory study of the relationships between reported imagery and the comprehension and recall of a story." *Reading Research Quarterly* (1983): 110-123.

Schafer, R. Murray. The new soundscape: a handbook for the modern music teacher. BMI Canada, 1969.

Schütz, Alfred. "Making music together: A study in social relationship." *Social research* (1951): 76-97.

Shokoff, James. "What is an audiobook?." *The journal of popular culture* 34, no. 4 (2001): 171-181.

Sousa, David A., ed. *How the brain learns to read*. Corwin Press, 2005.

Sousa, T. L. V., Carriere, J. S., and Smilek, D. The way we encounter reading material influences how frequently we mind wander. Frontiers in psychology 4 (2013)

Spacks, Patricia Meyer. *Boredom: The literary history of a state of mind*. University of Chicago Press, 1995.

Stern, John A., Larry C. Walrath, and Robert Goldstein. "The endogenous eyeblink." *Psychophysiology* 21, no. 1 (1984): 22-33

Stueber, Karsten R. "Varieties of Empathy, Neuroscience and the Narrativist Challenge to the Contemporary Theory of Mind Debate." *Emotion Review* 4, no. 1: 55-63. (2012)

Sutton, John, and Evelyn Tribble. "Cognitive Ecology as a Framework for Shakespearean Studies."Shakespeare Studies 39:94-103 (2011)

Sutton, Samuel, Margery Braren, Joseph Zubin, and E. R. John. "Evoked-potential correlates of stimulus uncertainty." Science 150, no. 3700 (1965): 1187-1188.

Sutton, Samuel, Patricia Tueting, and Joseph Zubin. "Information delivery and the sensory evoked potential." Science (1967).

Tan LH, Hoosain R, Peng DL (1995): Role of presemantic phonological code in Chinese character identification. *J Exp Psychol Learn Mem Cogn* 21:43–54.

Thierry, Guillaume, Clara D. Martin, Victorina Gonzalez-Diaz, Roozbeh Rezaie, Neil Roberts, and Philip M. Davis. "Event-related potential characterisation of the Shakespearean functional shift in narrative sentence structure."*Neuroimage* 40, no. 2: 923-931. (2008)

Town, Stephen, Katherine Wood, and Jennifer Bizley. "Multisensory integration in ferret auditory cortex: Effects of inactivating visual cortex." *Multisensory Research* 26 (2013): 221-221.

Turkyılmaz, Mustafa, Remzi Can, Kasim Yildirim, and Seyit Ateş. "Relations among Oral Reading Fluency, Silent Reading Fluency, Retell Fluency, and Reading Comprehension." *Procedia-Social and Behavioral Sciences* 116 (2014): 4030-4034.

Verstraete, Pieter. "The frequency of imagination: auditory distress and aurality in contemporary music theatre." (2009).

Villon, Olivier, and Christine Lisetti. "A user-modeling approach to build user's psycho-physiological maps of emotions using bio-sensors." In *Robot and Human Interactive Communication, 2006. ROMAN 2006. The 15th IEEE International Symposium on*, pp. 269-276. IEEE, 2006.

Wallentin, Mikkel, Andreas Højlund Nielsen, Peter Vuust, Anders Dohn, Andreas Roepstorff, and Torben Ellegaard Lund. "Amygdala and heart rate variability responses from listening to emotionally intense parts of a story."*Neuroimage* 58, no. 3 : 963-973. (2011)

Walter, Ong. "Orality and Literacy." *The technologizing of the word, TJPress, London* (1982).

Wandell. "Development of white matter and reading skills." *Proceedings of the National Academy of Sciences* 109, no. 44 (2012): E3045-E3053.

Weaver, Pascha. "RADIO DRAMA: A “VISUAL SOUND” ANALYSIS OF JOHN, GEORGE AND DREW BABY." PhD diss., University of Central Florida Orlando, Florida, 2012.8

Widdowson, Peter. Literature. London: Routledge (1999)*Applications* 11, no. 4 (1991): 67-74

Wigfield, A. (1994) Expectancy-value theory of achievement motivation: A developmental perspective. *Educational Psychology Review, 6*(1), 48-78.

Yeatman, Jason D., Robert F. Dougherty, Michal Ben-Shachar, and Brian A.

Zajonc, Robert B. "Emotion and facial efference: A theory reclaimed." *Science*228, no. 4695 (1985):15-21.

Zhao, Huiwen. "Emotion in Interactive Storytelling." PhD diss., Communication University of China, 2013.