Journal of Universal Computer Science, vol. 20, no. 7 (2014), 1006-1025
submitted: 31/1/14, accepted: 30/6/14, appeared: 1/7/14 © J.UCS

Let me Listen to Poetry, Let me See Emotions

Diana Arellano
(Filmakademie Baden-Wuerttemberg, Ludwigsburg, Germany
diana.arellano@filmakademie.de)

Cristina Manresa-Yee
(University of Balearic Islands, Palma de Mallorca, Spain
cristina.manresa@uib.es)

Volker Helzle
(Filmakademie Baden-Wuerttemberg, Ludwigsburg, Germany
volker.helzle@filmakademie.de)

Abstract: This paper presents the design, implementation and evaluation of the interactive installation The Muses of Poetry (MoP). In MoP the user interacts with a virtual character, who in turn recites poetry while manifesting the emotional content of the poem using visual cues (as facial expressions) and an affective voice. The novelty of MoP is that it combines real-time character animation, semantic analysis, natural voice interaction and poetry to create a unique and surprising experience for the user.
Key Words: Affective Computing, Human-Computer interaction, User experience
Category: J.5, H.5

1 Introduction

According to Andrew Stern, the creation of powerful emotional experiences, whether interactive or not, consists of inducing a reaction in an audience. In his experience, people look for situations where they can interact with something that seems alive, that has feelings, that they can connect with [Stern 2003].

That can be seen, for instance, in the use of virtual characters in applications where the user needs to interact with the machine, and these characters serve as the interactive interface. That is also the reason why researchers in affective computing are trying to enhance characteristics as believability, emotionality or empathy in those characters to provide better experiences.

Our search for new uses of emotional animated characters led us to join poetry, animated characters (or muses) and interactivity. Thus our motivation for The Muses of Poetry, or MoP for short, is to explore how virtual characters can be used in more affective applications. The idea of using poetry to this end comes from the fact that it is one of the most emotional and imaginative literature forms, which gave us space to play with different emotional elements. Moreover, MoP intends to bring people with little knowledge, or experience on poetry closer to this literary form.

The main objective of MoP was to create an interactive installation where animated characters recite poetry in an emotional way. The emotions are first extracted from the poems, and then manifested by the characters through visual cues (e.g. facial animations, movement, and so on) and emotional speech.

In MoP the user not only interacts with a virtual character by telling it a few words that will be included in the poem, but also experiences the intrinsic emotionality of the poem that is recited. From the technical perspective, the system built for MoP generates all the animations of the different emotional states, as well as the affective speech in real time. Therefore, every time a new poem is included in our repository, there is no need to wait for a new animation, or voice recording for each character, because they are automatically generated.

The rest of the paper is organized as follows. First we present a selection of previous works that have dealt with interactive poetry, with generated poetry and with virtual characters as interactive beings. Then, we explain how MoP was implemented and how each module of the system works. Afterwards, we describe the user experience evaluation carried on on a group of participants after their interaction with the installation. Finally, we give our conclusions and an overview of our future work.

2 Related Work

Poetry is considered one of the oldest and most creative forms of literary art. Over the years, many applications have used artificial intelligence, machine learning, human-computer interaction (HCI), among other techniques to create new representations, or produce new forms of poetry. In the following we will review a number of previous works related with these fields.

2.1 Artificial Poetry

Artificial intelligence has been widely used to "generate written poetry". As a result, it is the machine the one that creates new poems. That is the case of [Colton et al. 2012], who came up with a corpus-based poetry generation system that constructs poems according to a given rhyme, sentiment, word frequency and similarity. David Cope created the program "Alena" (Artificial Life Evolving Natural Affinities) to automatically write haiku, which were subsequently published in the ebook "Comes the Fiery Night" [Cope 2011]. [Gervás 2000] created WASP, a reasoning rule-based system that takes as input a set of words in Spanish and verse patterns and returns a set of verses. [Toivanen et al. 2012] made used of text mining methods, morphological analysis, and morphological synthesis to generate poetry in Finnish.

Regarding the poetry analysis, Tizhoosh and Dara [Tizhoosh and Dara 2006] and Tizhoosh et al. [Tizhoosh et al. 2008], analyzed text to distinguish between poem and prose, without understanding or interpreting the underlying poetic meaning. Other researchers worked on assessing the style of a poem, either to use it as a tool for the study of different types of poems and see how they affect the reader's perception of the poem [Kaplan and Blei 2007], or to figure out what makes a poem beautiful [Kao and Jurafsky 2012].

2.2 Interactive Media

Other ways in which poetry is often taken to new and fascinating levels are Media and Interactive technologies. In his book Digital Art and Meaning: Reading Kinectic Poetry, Text Machines, Mapping Art, and Interactive Installations, Roberto Simanowski delights us with a semiotic analysis of different installations focused on digital poetry [Simanowski 2011]. The main objective of these installations was to find new ways of representing the meaning of the poems, or providing users with a new way of experiencing poetry. An early example is MUSE, the work of [Tosa and Nakatsu 1998], where poems were created by exchanging poetic phrases between the user and the system; the latter being represented by a character which facial features are eyes, eyebrows and mouth. In Tosa's work, the facial expressions of the character changed according to the emotions conveyed in the phrases uttered by the user.

In the same direction of interactive systems, [Kwiatek and Woolner 2010] merged poetry into interactive storytelling based on still and video panoramas. The aim of their application was to develop interest not only in the life of the poet Charles Causley but also in his literary output. [Utterback and Achituv 1999] created an interactive installation where letters fall like rain or snow, landing on the heads and arms of the participant's projection, and forming entire words and sentences of a poem about bodies and language. The L'Oréal Poetry Harp [Small Design Firm Inc 2004] was also an interactive installation where 28 cords stretched from the ceiling can be plucked and pulled, releasing sinuous clouds of letters (projected on the opposite wall) from poems about a woman's worth.

2.3 Virtual Characters

Although they have been barely used for, or in poetry, it is not surprising to see that many researchers decide for Virtual Characters when it comes to create interaction. In this sense, the search for better and more believable human-computer interaction has been the main goal of the investigators, employing these characters in a variety of applications.

For instance,[Kasap et al. 2009] implemented Eva, a geography teacher capable of speaking and showing emotions through facial expressions and body gestures. [Reithinger et al. 2006] developed their own Virtual Human demonstrator system, taking as use case the World Cup in Germany 2006; there it can be seen two football experts and a moderator, and the users can ask the experts for advice and receive comments on their decisions.

[Kopp and Jung 2000] created the agent Max, an example of a virtual character that marked the cutting edge of research in artificial intelligence and artificial life at the turn of the millennium. Since 2004, Max is in permanent exhibition at the Heinz Nixdorf MuseumsForum [Pfeiffer et al. 2011], showing historical and concurrent technologies that can also be tested.

An example of a facial animation system based is Alfred [Bee et al. 2009]. The system allows creating facial expressions for Animated Agents in an easier, more intuitive and faster way than conventional animation tools. It has been used especially in studies about personality, eye-contact, and other non-verbal signals an agent should convey to create a favorable atmosphere.

3 Behind the System

The former works showed us that even though a lot has been done in the fields of interaction, poetry and virtual characters, no previous work resembles the Muses of Poetry. MoP can be described as a novel application of interactive virtual characters; as a way to explore how these characters can effectively transmit the intrinsic emotions of poems; and as an original combination of different areas like semantic analysis, real-time animation, voice generation, and voice recognition in noisy environments. Moreover, as an installation, it intended to go a step forward in the computational creativity direction by bringing into existence a "virtual poetry interpreter".

The last premise of MoP was to develop a novel autonomous system where all the processes, from the analysis of the poems to the generation of the animations were done automatically and in real time. As a computational system, MoP can be seen as a set of interconnected modules that provide the analysis of the poems and the manifestations of their emotional content. Fig. 3 shows an overview of the different modules of the system.

As for the poems, we only used free verse ones. This type of poems do not use a consistent meter pattern or rhyme, tending to look and sound like prose. Therefore, it was easier to perform a semantic analysis on them to extract the emotions, than on other kind of poems like Haikus (very short japanese poems, composed of 17 syllables and juxtaposed images). The poems for MoP were obtained thanks to the collaboration with the Australian online magazine Cordite Poetry Review, which served as intermediary between a group of Australian poets and us. Besides that, other poets kindly provided us their poems under the Creative Commons license.

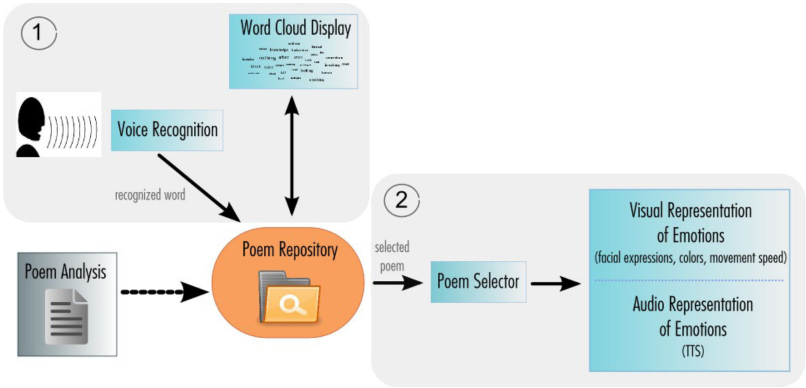

Figure 1: MoP System

3.1 Poem Analysis

In order to extract the emotional content of the poems, we developed a dedicated module that performs a semantic analysis of their text.

The input to this module is the text in XML format. This structure facilitates the extraction of the different parts of the poem: author, title and body. The body is then re-arranged, converting the lines of the verses in single lines that facilitate the semantic analysis, and subsequent tagging. Then, the analysis is done using an affective dictionary, obtaining the emotions in the poem. Finally, these emotions are translated into XML tags that will be used by other modules for the audiovisual manifestation of those emotions. Once the text is tagged, it is stored in the poems repository, becoming available for the rest of the system.

3.1.1 Verses into lines

A verse in a poem is usually formed by a number of lines, that together convey an idea. Therefore, each line is not enough to obtain the relevant emotion. That is why this module re-arranges the verses into single lines, adding punctuation marks that serve as separators (commas ‘,’, periods ‘.’, or double periods ‘..’); or if the line itself conveyed a complete meaning, it is left as it was.

3.1.2 Emotions Extraction

Once the poem is divided into lines, the semantic analysis is carried on to extract the emotional states. This analysis is done by using the Whissell’s Dictionary of Affect in Language (WDAL) [Duhamel and Whissell 1998]. The dictionary includes 10,368 English words with affective connotations, where each one is described with regard to the dimensions of Activation (or Arousal) and Evaluation (or Pleasantness). It also provides a detailed classification of each word in the poem according to the following states: pleasant, nice, fun, passive, sad, unpleasant, nasty, active, high imagery and low imagery.

Two reasons led us to decide for the WDAL instead of others like ANEW [Nielsen 2011] or WordNet-Affect [Strapparava and Valitutti 2004]. First, the WDAL has been created with words from literary and poetic texts, while the other two contemplate mostly affective words, which are not enough to assess the whole meaning of a poem. Second, the number of words contained in it is greater than in ANEW and Wordnet-Affect. As a result, the majority of words used in poems are mostly recognized by the WDAL and not by the other two.

Regarding the analysis, in a first step a *global* assessment of the poem is obtained, which indicates if the poem is in general pleasant, nice, fun, sad, unpleasant, or nasty. Additionally, a further classification is performed: if the poem was ranked as sad, unpleasant or nasty, then it is classified as negative. If the poem resulted pleasant, nice or fun, then it is classified as positive. Note that at the moment, the high and low imagery states are not used.

In a second step, each word in the poem is evaluated using the same states as for the *global* assessment. Using the emotional states of the words in a line, the module performs an analysis of that line, extracting the emotions associated to it. As a result of the semantic analysis, we obtained a *global* emotional state associated to the whole poem, and a *particular*, or line-per-line states associated to each of its lines.

3.1.3 Emotions into tags

The result of the semantic analysis gives a set of emotions that need to be converted into tags, and integrated into the text of the poem for the system to know how to interpret and manifest them.

There are two types of tags: *emotion tags* for visual manifestation of the affective content of the poem, and *prosody tags* to change the speech of the synthetic voice. The former ones are added at the beginning of each relevant line (i.e. lines with emotions similar, or in the same category as the global emotional state), e.g. [synthesis:emotion id=UNPLEASANT TRIGGER/], and then interpreted by the real-time animation module to trigger facial movements, or other visual representations. The latter ones contain prosody elements like "pitch" and "speed", and are added at the beginning of the line, and before the words with the line's prevailing emotional state, e.g. [synthesis:pitch level='79']. In order to simplify the prosody tagging, the module re-groups the states into: happy (i.e. pleasant, nice, or fun), unpleasant (i.e. unpleasant or nasty), and sad.

The following excerpt is the tagged version of the poem *For the Road* by Carol Jenkins:

`[synthesis:emotion id='JOY_TRIGGER'][synthesis:pitch level='99']` `[synthesis:speed level='100']` *First as a dare and then for the* `[synthesis:pitch level= '124']` *warm* `[/synthesis:pitch]` *languor of the tar, at midnight* `[synthesis: pitch level= '124']` *walking* `[/synthesis:pitch]` *to my house, we lay down our bodies on the middle of Moana Road and* `[synthesis:pitch level='124']` *kissed,* `[/synthesis:pitch]` *.. Those long dreamy kisses of abandonment, to each other, to the road, to the dark pines looking on, to the locked light of houses with blinds drawn tight on quarter acre blocks,* `[/synthesis:speed] [/synthesis:pitch ]`

3.1.4 Assigning values to Prosody Tags

While the *emotion tags* only need the name of the emotion obtained from the analysis, *prosody tags* need a value to modulate the speed and pitch of the voice. These tags are always located at the beginning of the line, and in the particular case of pitch, also before the word to emphasize. The values we use for speed and pitch are obtained from the global and line-per-line analysis of the poem, according to the following rules:

1. Compute the increment, or decrement, of the neutral pitch and speed, according to the poem emotional state (i.e. if it is happy, sad, or unpleasant). In the case of a human-like character, we use the following values:
   - If the state is *happy*, the neutral speed (i.e. neutral speed=90) is incremented by 5%, and the neutral pitch (i.e. neutral pitch=90) by 10%.
   - For an *unpleasant* state, the decrement in neutral pitch and speed is 8%.
   - For the *sad* state, the decrement in neutral pitch and speed is 5%.
2. Compute for each line its resultant emotional value by multiplying (a) the poem emotional state value obtained from the affective analysis using the WDAL (e.g. pleasant = 11.45) by (b) the line emotional state value, which is basically the number of words with the same emotional state of the line (e.g. if the line is "pleasant" and has 4 words rated as"pleasant", then the line state value will be 4)
3. To the speed increased in step 1, add (or subtract, depending on the emotional state) the emotional value obtained in step 2. This value will be used to tag the speed of the whole line.
4. The pitch value for the tag of the whole line is obtained from step 1. Moreover, to modify the pitch value of each word with an emotional state similar to the one of the line, the value obtained in step 2 is added to (or subtracted from, depending on the emotional state) the neutral pitch. If no words are rated with the same state as the line, then no pitch changes are applied.

The subtle variations in step 1 are due to the fact that abrupt changes in a realistic character are seen as very uncanny, diminishing the whole experience. However, this changes in the case of a more cartoon or abstract character, where the common guidelines followed in animation include the idea that cartoon characters should be exaggerated to better convey emotion and intent [Hyde et al. 2013].

Due to the static nature of the poems, the affective analysis is performed only once. The resultant tagged poem is stored in the poems repository, ready to be used by the other modules of the system [see Fig. 2]

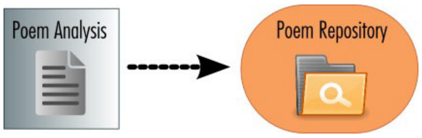

Figure 2: MoP - Poem Analysis Module

3.2 Voice Recognition - Interaction

One of the objectives of MoP was to achieve a natural interaction between the user and the system. "Natural" in this context meant to free the user from cables or other devices that should be held or attached to the body. The metaphor of interaction is a conversation between the user and the animated characters.

To achieve this we developed a module for voice recognition that made use of a clip microphone, and the state of the art third-party software VoCon® Hybrid [1] (SEMVOX) for the voice recognition itself. To give a better impression of natural interaction, we hid the microphone on the back of the second rib of the installation, right above the head of the participant [see Fig. 3].

As for the interaction itself, it starts when the user stands at the entrance of the installation and awakes the system with the greeting: "Hello Muses". Once this is acknowledged, the displayed character replies: "Hello, I am one of the Muses of Poetry. Select two words and I will recite a poem for you".

The words the user can select to interact with the character are displayed in the form of a "word cloud" over the character. The first time the user approaches the installation, the module displays the word cloud formed by a set of 30 words chosen from all the poems in the repository. After a word is said and recognized, the system looks for it in the current set of poems, keeping only the poems that

[1] http://www.nuance.com/for-business/speech-recognition-solutions/vocon-hybrid/index.htm

Figure 3: MoP - Microphone for voice recognition

contain that word. Then, the most frequent words of this reduced set of poems are again displayed as a cloud. Once the user selects a second word, and this is acknowledged by the system, then the control is completely passed to the Poem Selector module [see Fig. 3 ①].

The decision of choosing two words came after some previous informal experimentation. We noticed that with three words, the user felt it was too long to start listening to the poem. On the contrary, the use of just one word was not engaging enough to make the user feel part of the installation.

3.3 Animated Characters - Representation of emotions

Once the poem is selected, the system generates the corresponding animations for the character and the emotional speech [see Fig. 3 ②]. The animation of virtual characters is itself a very demanding task that requires both expertise and great amounts of time. Therefore, to accomplish a dynamic visual representation of emotions, the MoP system had to provide real-time animations for both 2D and 3D characters. In this way, there is no need of having pre-rendered animations for each poem, which represented a great advantage because any new poem could be added into The Muses without effort on the animation side. Furthermore, the visualization of the emotions needed to be multimodal, meaning that it could not be limited to only facial expressions given our varied repertoire of characters.

3.3.1 The Characters

– **Nikita:** she is a realistic 3D character, inherited from our previous applications with interactive virtual characters (e.g. [Helzle et al. 2011] and a conference guide). To make Nikita look more like an old-Greek "poetry teller", we added a veil, which also helped to avoid the time and resources consuming task of rendering her hair in real-time.

To animate Nikita we used the principles of the Facial Animation Coding System (FACS) [Ekman and Friesen 1978], specifically the Action Units (AU), which describe the facial movements produced due to facial expressions. Thus each AU is associated to a facial region, which is activated depending on the emotion to manifest. Each emotion has different representations, adding variability to the experience. Besides happiness, sadness, disgust and anger, Nikita can also be in an idle state and acknowledge the recognition of words with expressive facial expressions [see Fig. 4].

Figure 4: 3D Female Character

- **Particles Head:** this 3D abstract character was created as an alternative to Nikita. The received feeback pointed out that even though Nikita conveyed emotionality, she was too realistic, and users associated poetry with more abstract representations. As a result, an abstract character made of white-shaded particles, whose face resembled a dynamic mask, was designed. The dynamism was provided by the particles moving all over the face.

  The Particles Head was implemented as a real-time k-d tree based particle system, with a spatial hash table to further enhance its performance. The particles can be either emitted from a point, or randomized from a disk shape. Several types of affectors control the movement of the particles, and define the shape and flocking behavior of the particles cloud. The shape of the face is formed by an invisible mesh, which is used as a particle attracting affector, while point and plane affectors push the particles away from the face surface to create the impression of eyes and a mouth.

  To manifest the emotional content of the poem, we decided for a change in the movement speed of the particles, guided by a musical background. Using a set of six tracks for the emotions of happiness, sadness and unpleasantness (2 tracks per emotion), we varied the speed of the particles. For instance, if the poem was happy, the happy track would be played while the particles would move faster and close to the face, almost hiding the mesh. On the other hand, if the poem was sad, then the sad track would be played and the particles would move slower, with less strength between them [see Fig.5].

Figure 5: Particles Head design. From left to right: neutral, anger, happiness

– **Myself:** this was a cartoonish 2D character, whose facial features in the neutral state are a circle (for the face), two dots (for the eyes) and one line (for the mouth). To animate this character we use its visemes to create the impression of pitch. A viseme is defined as a generic facial image that can be used to describe a particular sound. The character had in total 16 visemes, corresponding to sounds like: a, e, v/f, p/b, and so on. Then, each time a viseme was triggered by the text-to-speech (TTS) tool, the corresponding pose was displayed, creating the animation of the speech [see Fig. 6].

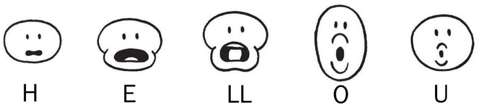

Figure 6: Myself: 2D Character

Regarding the manifestation of emotions, we were provided with sets of poses (up to 8) that formed a short clip animation (e.g. the character exploding, or turning into a line). Hence, when the TTS found an emotion in the poem, a clip was triggered manifesting the corresponding emotion [see Fig. 7].

Figure 7: Myself. Clip to show "anger"

- **Woody:** this character was created using materials like stones, branches and natural threads. Each viseme was achieved by changing the configuration of the material and then photographed to obtain the corresponding image. Therefore, the animation of this character was done in the same style as with Myself. To represent the emotions in the poem, the character changes its color. For happy poems, a color correction using bright yellow tonalities was used, while for sad poems the color correction was based in blue, and red for unpleasant poems. The transition between one emotional state to the other was done by blending the corresponding color channels [see Fig. 8].

Figure 8: Woody displaying emotions. From left to right: neutral, joy, sadness, anger

- **Sailor:** similarly as how Woody was created, the Sailor was made of clay, and then each viseme photographed, for which the changes were mostly done in the mouth. His rough appearance was left on purpose to make him an hybrid between an abstract character and humanoid one. Contrary to the other character, the Sailor did not manifest emotions visually, but with his voice. Given his comical appearance, we were capable of playing with the pitch and speed values to create real emotional voices (e.g. high pitched for happiness and very low pitched for sadness) [see Fig. 9].
- **Krel:** this 2D cartoonish creature was animated only by changing the position of the mouth, to resemble the corresponding visemes. In the same way as Woody, we decided for a change of color of its skin to manifest the triggered emotional state [see Fig. 10].

Figure 9: Sailor: 2D/3D Character

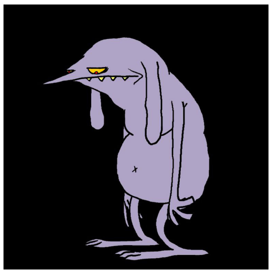

Figure 10: Krel: 2D/3D Character

3.3.2 Real-time animation

In order to produce all the dynamic and real-time animations, we used our own development framework named Frapper [2]. It features a node-based scene model with plugins for node types, a model-view-controller architecture, a Qt-based panel-oriented user interface and a viewport using the Ogre 3D render engine.

In the case of Nikita, we made use of the Facial Animation Toolset [3], and in particular of the Adaptable Facial Setup (AFS), which is part of a pipeline for the creation and animation of facial expressions in animated characters in Maya®. The result was then exported to Frapper in the form of files containing the references to the object and its animation properties.

Another important feature of Frapper is the automatic lip-sync. Thus, having previously defined the poses for each viseme, these are displayed when the corresponding viseme is triggered by the TTS. Additional rendering elements provided by the framework are tears simulation, eye redness, wrinkles generation, color correction and sound play.

[2] http://sourceforge.net/projects/frapper/
[3] http://fat.research.animationsinstitut.de

3.4 Voice Synthesis - Vocalization of emotions

One of the characteristics of poetry is its freedom of interpretation, which can lead to different ways of reading it aloud. Pauses, intonation, melody, and emotions are some of the elements that need to be taken into account when reciting poetry. Their correct use can enhance the poetic experience to a level that is capable of engaging a wider audience.

In MoP one of our objectives is to transmit the emotionality of the poems not only with visual representations, but also with changes in the speech. This was done through the use of the *prosody tags* in the poems, which are interepreted by the text-to-speech (TTS) software SVOX® (NUANCE), which modulates the generated synthetic voices.

The added punctuation marks and new lines (also from the poem analysis) are interpreted as pauses. The length of the pause depends on the mark. If it is a new line or double periods '..', then it is a long pause. On the contrary, if it is a comma ',' or a period '.', then it is a short pause (comma pause is in general shorter than the period pause).

One of the advantages of using a TTS is that we do not need to pre-record each of the poems in the system. In this sense, any new poem can be added without having to depend on the recordings of the poet, or other persons. The other advantage is that the voice can be made more emotional by changing the values of speed and pitch in the tags.

The main disadvantage of using a TTS is that intonation cannot be modified. Therefore, we cannot change the pitch through its syllables, but of the whole word, which is some cases does not sound natural enough.

4 User Experience

4.1 Installation

The Muses of Poetry was presented to the audience in the form of an interactive physical installation. The motivation for it was to create a scenario where the participant could feel somehow isolated from the surrounding world, and therefore, an integral part of the experience. As for the audience, we did not have a specific "target" in mind. On the contrary, we wanted to reach as many people as possible (artists, technical people, HCI experts, etc.) and see how they react and feel the characters and the poetry that they recite.

The result was a sort of cave created with flexible ribs that simulated the pages of a book seen in perspective [see Fig. 3]. Besides the rib structure, the installation consisted of: a fire-proof carpet that also served to muffle the noise, a projection screen, a beamer for back-projection, a sound-box, and a clip microphone. As for the measures, it was approximately 6 meters in length, 3 meters in width and 2.5 meters in height.

For a better interaction we recommended the users to stand at the "entrance" of the installation, preferably under the second rib, because of the microphone. The next step was to explain them how the interaction needs to be performed (i.e. awake the system by saying aloud *Hello Muses*, and then selecting two words from the word cloud). In order to asses how users perceived the characters included in the system, the person in charge of the installation selected every hour a different character to be displayed.

The complete interaction with MoP had a maximum duration of approximately 3 minutes. It mainly depended on the length of the poems, but we tried to keep them in the range of 1 to 2 minutes. As an open installation, MoP has been designed to allow more than one person to experience the poetry. It was also possible that two persons take turns to select the initial words [see Fig. 11].

Figure 11: Interactive Installation of The Muses of Poetry

To evaluate the installation as a system, and the characters and the manifestation of emotions, we took advantage of its exhibition during the FMX 2013. FMX is the annual conference on Animation, Effects, Games and Transmedia, celebrated in Stuttgart, Germany. For four days, The Muses of Poetry was presented in an exhibition floor where people interacted with the different characters and had a different poetic experience. After the interaction, each person was asked to fill a questionnaire to assess the system and the experience with The Muses. At the moment of this questionnaire, we had developed 3 out of the 6 characters that had the installation in its final state.

In the following subsections we explained the procedure and results of the evaluation of Nikita, Particles Head and Myself, and how this assessment led us to the creation and integration of Woody, Sailor and Krel.

4.2 Part I - Evaluation of Nikita, Particles Head and Myself

Participants experienced the installation alone with the support of a member of the development team, who explained how the system worked. Each participant interacted with only one character: Nikita, Particles, or Myself. At the beginning, the character asked the user to select two words from a word cloud displayed on the screen. Once the two words were recognized, a poem containing those two words was selected and recited. At the end of the experience, the participant completed a questionnaire regarding the installation, the character and his or her feelings towards the interactive system and the poem reading. The written questionnaire consisted in eight 5-point Likert scale questions ranging from "Strongly Disagree" on one end to "Strongly Agree" on the other:

Q1: The character is attractive as a poetry reader
Q2: The character conveys emotion
Q3: The character read the poem in a way I understood the topic of the poem
Q4: I think I would like to visit this installation frequently or I would recommend it to my friends
Q5: I felt a variety of emotions while listening to the poems
Q6: The system is pleasant
Q7: The system is inviting
Q8: The system is appealing

At the end 51 questionnaires were gathered: 35 from male participants and 16 from female participants, with ages ranging between 19 and 45 years (average age was 27).

Data was analyzed for each specific question using two approaches: separately for each character to find differences among the three poetry readers and together to conclude global insights from the interactive system.

Figure 12 shows the boxplots for each of the eight questions, considering the results of the three characters together. Based on the interquartile ranges (IQR) of questions Q1, Q2 and Q5, all related with the emotional aspect of the installation, we could not conclude if the user felt emotions when listening to the poems, or if he or she felt the characters emotional enough, because the bars are widely distributed. However, Q3, which evaluated the degree of empathy with poetry, throws *median*=4, which indicates that one of the objectives of the installation (make a wider audience understand poetry) might have been achieved. As for questions Q6, Q7 and Q8, despite the outliers, we can conclude that the installation was pleasant, inviting and appealing.

Figure 13 shows the boxplots for each of the eight questions, for each character. The boxplot on the right of figure 13 shows the IRQs when evaluating Nikita. For Q1, although the median value was above the mid point, it cannot be concluded that Nikita was attractive as a poet reader. The analysis of Q2

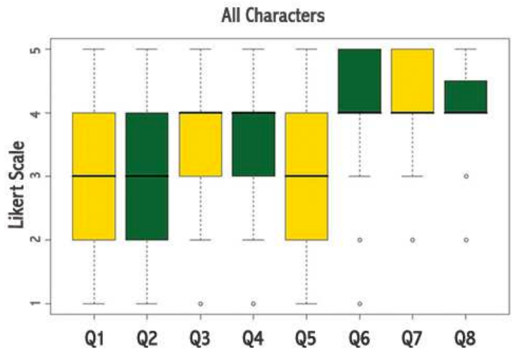

Figure 12: Inter-quartile analysis for each question and all characters

and Q3 do not throw decisive results either. However, people felt they understood the topic of the poem (Q4) and felt the installation more inviting (Q7) and appealing (Q8) than pleasant (Q9).

From the boxplot corresponding to Particles we can conclude that people indeed understood the content of the poems (*median*=4), although 50% of the sample is under this value. As for the appealing, pleasantness and inviting nature of the installation, results are not so satisfactory as with Nikita and Myself. Nonetheless, these elements were rated as positive.

Finally, from the boxplot with the evaluation of Myself, we cannot conclude that the character was attractive as a poet reader (Q1), but it did convey emotions (Q2) and 50% of the sample felt emotions while listening to the poems. Participants also felt that with this character the installation was pleasant, inviting and appealing. A last thing to note is that with Myself users would recommend the installation to friends, in a higher scale than with the other characters.

A last analysis focused on the mode and median differences based on genders showed that women found the interactive installation more inviting and pleasant than men, and in general they rated higher values than men in all question, as seen in Table 1. In a more detailed analysis, women also felt more likeliness for the Myself character, while men for Nikita. Regarding the mode, it was observed that women felt more emotions while listening to the poem.

4.3 Part II - Addition of new characters

The results of the user experience evaluation showed that the majority of the participants found the installation pleasant, inviting and appealing. However, these results did not allow us to conclude which character performed better as a poetry reader.

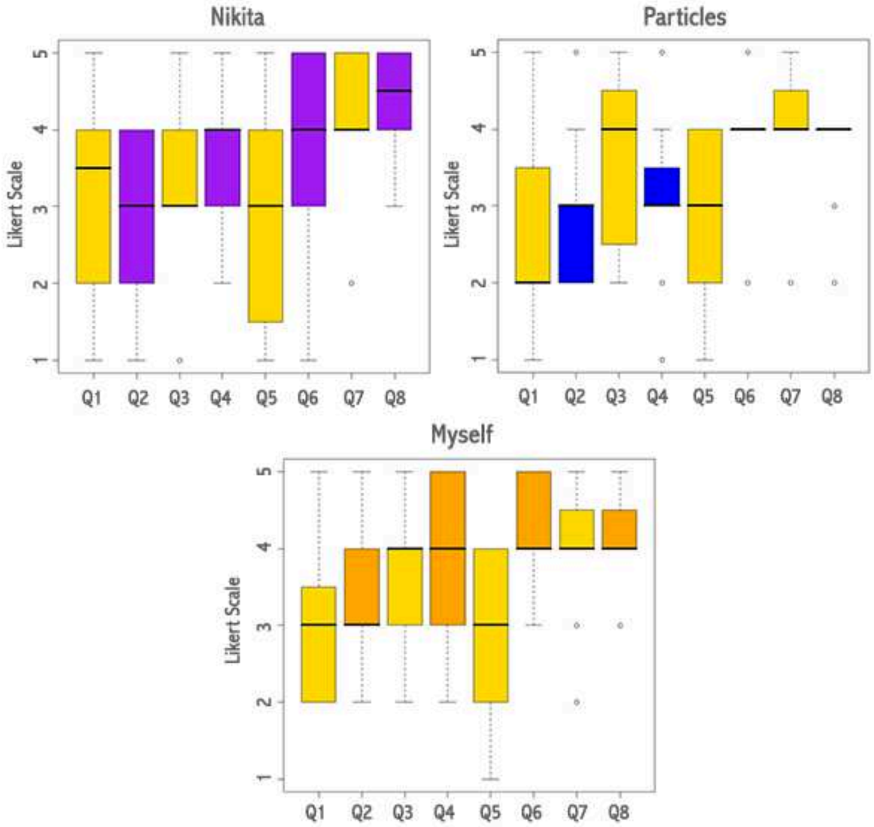

Figure 13: Inter-quartile analysis for each question and each character: Nikita (upper-left), Particles (upper-right), Myself (bottom)

As a consequence, we decided to introduce new and original characters that could enhance the interactive experience, and perform as creative poem interpreters. To that end, we carried on a workshop on lip-sync, where the students needed to come up with one character that should be animated by the end of the seminar. The chosen characters for Mop were Woody, Krel and the Sailor because they moved away from the traditional character design, and their animation was itself appealing enough.

5 Conclusions and Future Work

In this paper we have presented the concept and implementation of the interactive media installation named The Muses of Poetry, or MoP.

Table 1: Data of participants and evaluated characters

	Q1	**Q2**	**Q3**	**Q4**	**Q5**	**Q6**	**Q7**	**Q8**
median (women)	3	3	4	4	3	4.5	5	4
median (men)	3	3	3	4	3	4	4	4
mode (women)	3	3	4	5	4	5	5	4
mode (men)	2	3	3	4	2	4	4	4

Conceptually, our intention was to make users get a closer approach to poetry. This was achieved through the interaction with a virtual character who expressed the underlying affective content of each poem. The ease to add new poems in our framework made MoP an interesting and novel example of interactive poetry. It is worth noting that it was not our intention to replace the poet. On the contrary, we thought that having characters that manifest emotions in different ways could enhance the performance of a real poet.

In the future we intend to evaluate the 6 characters, in the same manner as we did with Nikita, Particles Head and Myself, and see which type of character conveys the emotionality of the poems in the best manner.

Acknowledgments

The project was funded by the Innovationsfonds Kunst of the Ministry for Science, Research and Art in Baden-Württemberg (54-7902.513/65). Special thanks to the creators of the charaters: Prof. Andreas Hykade (Myself), Amit Rojtblat (Particles head), Nils Zweiling (Unexpected GmbH), Tatiana Pushkareva (animations 3D female character), Irina Robina (Woody), Valentin Kemmner (Sailor) and Anselm Pyta (Krel); and to Jan Roth and Patrick Puszko for the music tracks.

References

[Bee et al. 2009] Bee, N., André, E., Tober, S.: "Breaking the Ice in Human-Agent Communication: Eye-Gaze Based Initiation of Contact with an Embodied Conversational Agent"; International Conference on Virtual Agents (IVA) (2009), 229-242.

[Colton et al. 2012] Colton, S., Goodwin, J., Veale, T.: "Full face poetry generation"; Proceedings of the Third International Conference on Computational Creativity (2012).

[Cope 2011] Cope, D.: "Comes the Fiery Night"; Amazon Books, NY.

[Duhamel and Whissell 1998] Duhamel, P., Whissell, C.: "The dictionary of affect in language", [computer software] (1998).

[Gervás 2000] Gervás, P.: "Wasp: Evaluation of different strategies for the automatic generation of spanish verse"; Symposium on Creative & Cultural Aspects and Applications of AI & Cognitive Science (2000).

[Ekman and Friesen 1978] Ekman, P., Friesen, W.: "Facial Action Coding System: A Technique for the Measurement of Facial Movement"; Consulting Psychologists Press.

[Helzle et al. 2011] Helzle, V., Spielmann, S., Zweiling, N.: "Emote, a new way of creating animated messages for web enabled devices"; CVMP (2011).

[Hyde et al. 2013] Hyde, J., Carter, E. J., Kiesler, S., Hodgins, J. K.: "Perceptual effects of damped and exaggerated facial motion in animated characters"; IEEE International Conference on Automatic Face and Gesture Recognition - FG 13 (2013)

[Kao and Jurafsky 2012] Kao, J., Jurafsky, D.: "A computational analysis of style, affect, and imagery in contemporary poetry", in NAACL Workshop on Computational Linguistics for Literature (2012).

[Kaplan and Blei 2007] Kaplan, D. M., Blei, D. M.: "A computational approach to style in american poetry"; Proceedings of the 2007 Seventh IEEE International Conference on Data Mining (2007), 553558.

[Kasap et al. 2009] Kasap, Z., Moussa, M. B., Chaudhuri, P., Magnenat-Thalmann, N.: "Making them remember - emotional virtual characters with memory"; IEEE Computer Graphics and Applications, 29, 2 (2009), 20-29.

[Kwiatek and Woolner 2010] Kwiatek, K., Woolner, M.: "Let me understand the poetry. embedding interactive storytelling within panoramic virtual environments"; Electronic Visualisation and the Arts (EVA) (2010) 199205.

[Kopp and Jung 2000] Kopp, S., Jung, B.: "An Anthropomorphic Assistant for Virtual Assembly: Max"; Working Notes Workshop Communicative Agents in Intelligent Virtual Environments, Autonomous Agents (2000).

[Nielsen 2011] Nielsen, F. A.: "A new anew: evaluation of a word list for sentiment analysis in microblogs"; Proceedings of the ESWC2011 Workshop on "Making Sense of Microposts": Big things come in small packages, 718 (2011), 9398.

[Pfeiffer et al. 2011] Pfeiffer, T., Liguda, C., Wachsmuth, I., Stein, S.: "Living with a Virtual Agent: Seven Years with an Embodied Conversational Agent at the Heinz Nixdorf MuseumsForum"; Proceedings of Re-Thinking Technology in Museums (2011).

[Reithinger et al. 2006] Reithinger, N., Gebhard, P., Löckelt, M., Ndiaye, A., Pfleger, N., Klesen, M.: "VirtualHuman- Dialogic and Affective Interaction with Virtual Characters"; ICMI'06 (2006), 24.

[Simanowski 2011] Simanowski, R.: "Digital Art and Meaning: Reading Kinetic Poetry, Text Machines, Mapping Art, and Interactive Installations"; University of Minnesota Press (2011).

[Small Design Firm Inc 2004] Small Design Firm Inc: "L'oréal poetry harp". Presented in Cambridge, USA (2004).

[Stern 2003] Stern, A.: "Creating Emotional Relationships with Virtual Characters"; Ed. R. Trappl, P. Petta, and S. Payr, MIT Press.

[Strapparava and Valitutti 2004] Strapparava, C., Valitutti, A.: "Wordnet-affect: an affective extension of wordnet", 4th International Conference on Language Resources and Evaluation (LREC 2004) (2004), 10831086.

[Tizhoosh and Dara 2006] Tizhoosh, H. R., Dara, R. A.: "On poem recognition"; Pattern Analysis and Applications 9, 4 (2006), 325338.

[Tizhoosh et al. 2008] Tizhoosh, H. R., Sahba, F., Dara, R.: "Poetic features for poem recognition: A comparative study"; Journal of Pattern Recognition Research 3, 1 (2008).

[Toivanen et al. 2012] Toivanen, J. M., Toivonen, H., Valitutti, A., Gross., O.: "Corpus-based generation of content and form in poetry"; Proceedings of the Third International Conference on Computational Creativity (2012).

[Tosa and Nakatsu] osa, N., Nakatsu, R.: Interactive Poem System. ACM Multimedia 1998 (1998), 115-118.

[Utterback and Achituv 1999] Utterback, C., Achituv, R.: "Text rain"; Permanent exhibition at the 21c Museum Hotel, Louisville, USA (1999).