

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/328215976>

# Spoken Conversational AI in Video Games–Emotional Dialogue Management Increases User Engagement

Conference Paper · October 2018

DOI: 10.1145/3267851.3267896

---

CITATIONS

38

---

READS

8,043

3 authors, including:



[Ioannis Papaioannou](#)

Heriot-Watt University

10 PUBLICATIONS 132 CITATIONS

[SEE PROFILE](#)



[Oliver Joseph Lemon](#)

Heriot-Watt University

372 PUBLICATIONS 6,144 CITATIONS

[SEE PROFILE](#)

# Spoken Conversational AI in Video Games – Emotional Dialogue Management Increases User Engagement

Jamie Fraser  
Heriot-Watt University  
Edinburgh, UK  
jdf4@hw.ac.uk

Ioannis Papaioannou  
Heriot-Watt University  
Edinburgh, UK  
i.papaioannou@hw.ac.uk

Oliver Lemon  
Heriot-Watt University  
Edinburgh, UK  
o.lemon@hw.ac.uk

## ABSTRACT

In a traditional role-playing game (RPG) conversing with a Non-Playable Character (NPC) typically appears somewhat unrealistic and can break immersion and user engagement. In commercial games, the player usually selects one of several possible predefined conversation options which are displayed as text or labels on the screen, to progress the conversation. In contrast, we first present a spoken conversational interface, built using a state-of-the-art open-domain social conversational AI developed for the Amazon Alexa Challenge, which was modified for use in a video game. This system is designed to keep users engaged in the conversation – which we measure by time taken speaking with the character. In particular, we use emotion detection and emotional dialogue management to enhance the conversational experience. We then evaluate the contribution of emotion detection and conversational responses in a spoken dialogue system for a role-playing video game. In order to do this, two prototypes of the same game were created: one system using sentiment analysis and emotional modelling and the other system that does not detect or react to emotions. Both systems use a spoken conversational AI system where the user can freely talk to a Non-Playable-Character using unconstrained speech input.

## CCS CONCEPTS

• **Computing methodologies** → **Natural language generation**; *Information extraction*; • **Human-centered computing** → *Natural language interfaces*;

## KEYWORDS

Speech, dialogue systems, video games, sentiment analysis, emotion

## ACM Reference Format:

Jamie Fraser, Ioannis Papaioannou, and Oliver Lemon. 2018. Spoken Conversational AI in Video Games – Emotional Dialogue Management Increases User Engagement. In *IVA '18: International Conference on Intelligent Virtual Agents (IVA '18), November 5–8, 2018, Sydney, NSW, Australia*. ACM, New York, NY, USA, 6 pages. <https://doi.org/10.1145/3267851.3267896>

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

IVA '18, November 5–8, 2018, Sydney, NSW, Australia

© 2018 Copyright held by the owner/author(s). Publication rights licensed to ACM.

ACM ISBN 978-1-4503-6013-5/18/11...\$15.00

<https://doi.org/10.1145/3267851.3267896>

## 1 INTRODUCTION

Many developers of today's role-playing games strive to find a balance in creating a game that is both immersive and enjoyable to play [2, 3]. To better understand immersion in this setting, we can distinguish two different kinds of immersion: *diegetic immersion*, where the player is immersed into the game itself, and *situated immersion*, where it describes the immersion in the experience of also acting within the game space [6, 10, 14]. This has often been done by developing ultra-high resolution photorealistic graphics with seamlessly smooth animations, and well developed emotionally stimulating storylines. Many games have moved to create a less linear experience where the player's choices or action can shape and guide the games story down different paths. Developments of branching dialogue trees can be used to make player's experiences seem unique to them, but the core of video game dialogue systems has not changed a large amount in the last 25 years. The players are usually presented with several choices of what dialogue the game character should say, and upon selection the the conversation continues along the pre-scripted path that the game developer designed (see e.g. figure 1). While such a system is very simple and easy for the players to use and progress with, as very little thought must be put in on the player's part, it also appears quite unrealistic and can break a game's immersion in otherwise engaging game worlds.

One way to increase player engagement could be to implement a spoken dialogue system or "conversational AI" in place of the traditional dialogue menu used to interact with NPCs. Laird [8] explains the limitations of modern video games' interfaces, and the role of Natural Language in this context. Using specially designed speech chatbots that are tailored to appear natural in the game world could make every player experience unique. A player could talk freely and naturally to the NPC, which would help enhance the immersion that game developer seeks to achieve. This could also make games more challenging as players would have to think more about what they are going to say, rather than selecting a predefined option.

Manipulating a video game character's emotion is another aspect that has been developed over the years. Freeman [4] explains the advantages of such a feature in video games and how it can improve the narrative. Many different games have used emotion to change and shape the conversation between an NPC and the player. However due to the nature of traditional conversational systems in games the choice is often clear and very plain to see. Players are presented with several options and of those options there may be some types of emotion present, but these options are clearly identified and labelled.



Figure 1: Simple Choice Dialogue menu in “The Legend of Zelda: Breath of the Wild” [Nintendo, 2017]

This project combines Natural Language Processing (NLP) and emotion recognition to evaluate the value of emotional driven spoken dialogue systems in a role-playing video game. To evaluate this, an emotion-driven spoken dialogue system game was developed. Users/players were required to interact with an NPC and learn certain pieces of information. Those pieces of information are only revealed by the NPC after the user has emotionally engaged in a positive (polite and kind) or negative (rude and aggressive) manner and pushed the NPC to become friendly or scared by the user. This system was tested and compared next to a simpler version of the dialogue system that lacks the emotion module. The evaluation compared the users’ assessment of the different systems and measures how emotionally-driven dialogue affected the user experience in terms of immersion, enjoyment, and engagement – the latter measured in terms of the time that the user spends talking with the character.

## 2 TYPES OF DIALOGUE SYSTEMS IN VIDEO-GAMES

Throughout the years, there have been many different ways to perform dialogue in video games, and to present dialogue options to the user. Here, we briefly note the most common ones:

### 2.1 Simple Choice/Hub

This form of interaction and dialogue consists of providing the player with options of things to do for example “trade”, “threat”, or “sell” (See Figure 1). These choices are typically actions rather than choosing precisely what the character says, however typically brief dialogue follows selecting a choice. This form of dialogue is used in many different types of game, however is often found being used in conjunction with other dialogue methods in RPGs.

### 2.2 Non-Branching Dialogue

One of the simpler forms of dialogue interaction in games is Non-Branching Dialogue. With this style, the player chooses to talk to an NPC who then delivers their scripted dialogue. There may be back and forth dialogue between the in-game player character and the NPC, but often the player will have no choice on what is said. This form of dialogue exists in a vast number of games from role playing games to first person shooters. The reason for this is that it’s the easiest form of dialogue to implement, it also serves to progress



Figure 2: Dialogues systems revealing too little information. Fallout 4’s standard dialogue menu (left) vs. custom “modded” menu (right) [Bethesda Softworks, 2016]

story elements more quickly while keeping the player informed and aware. In role playing games, it is often used in conjunction with Branching Dialogue systems to add some more substance to game worlds. [9]

### 2.3 Branching Dialogue Tree

The branching dialogue tree is a more advanced dialogue system that is found in a large number of role playing games and adventure games. When conversation is initiated with an NPC, the player is provided a list of dialogue options on screen. Once the player selects an option, the conversation will play out and the player is provided with a new set of options in response. Selecting the different options presented to the player will direct the conversation in a different way depending on how the narration is written. Each different dialogue option “branches” to different options to provide a more in-depth conversational experience that is driven by the player’s choices.

In some more linear games however, the choices made by a player in a conversation don’t have any impact on the overall story. The inclusion of giving players conversation options might just be there to provide the player with additional information on the game world but not all players will be satisfied with their choices being inconsequential. Another controversial aspect of dialogue in games is that often what the player’s in-game character says is somewhat different from what the selected dialogue option was. While this is not an issue in all games, some games are notoriously bad for misrepresenting dialogue. *Bethesda Fallout 4*’s dialogue system (see Figure 2) paraphrased dialogue options so much so that community members developed a mod that show precisely what was going to be said instead of a loosely summarised version [5].

Some games provided additional information of each dialogue options tone. *Mass Effect: Andromeda*’s dialogue system included icons that specified what different tones of conversation options were. A “friendly” dialogue option had a different icon from a “professional” response, which had a different icon from a “flirty” response. This let players understand what direction their selected dialogue options will go in. Older *Mass Effect* games included a simpler version of this system that identified certain options as “Paragon” or “Renegade”, BioWare’s interpretation of good and bad. This sort of system aids understanding of the direction of the conversation and makes progression easier in games however, according to a survey conducted by the Interactive Digital Software Association

2002, “71.4% of gamers play games because they are challenging” [13].

To change these traditional text-based systems into a spoken dialogue system could make games more challenging and engaging as the players would have to think about what they want to say rather than just selecting a pre-defined option.

### 3 CONVERSATIONAL SYSTEM ARCHITECTURE

The developed spoken conversational system is split into three main components: 1) the Unity engine that handles the video game aspects of the prototype, and the input of the user from the Automated Speech Recognition (ASR), 2) IBM Watson’s Cloud services, used for Tone Analysis and Speech Synthesis [7], and 3) the Alana conversational AI [12] dialogue management, response generation, and emotional modelling. At its core the system takes a speech input from the user and returns an output string to be used as the response from the NPC via speech synthesis.

The Unity game is run on a windows machine where it takes the user input and uses HTTP requests to communicate with IBM’s cloud services. It then sends the responses from those services to a lightweight version of the open-domain social conversational AI system *Alana* presented in [11, 12].

#### 3.1 Sentiment Analysis

In order to receive the emotional input of the user a method of analysing the users’ utterances and tone must be in place. Sentiment Analysis is the processes of determining the emotional connotations of a series of words. This can be done using a variety of platforms such as online view platforms and social media websites [1].

#### 3.2 Unity Game Prototype

The completed implementation resulted in a small 3D game environment very similar to that of a fantasy role playing game such as *Skyrim*. The player of the game can control the in-game character by moving and looking around the room, the room being a medieval tavern setting. The player’s goals are to converse with the barkeeper of this tavern to extract information on the whereabouts of a magical sword, needed to complete a quest. To do this the player has to approach the barkeeper to initiate spoken conversation. To test the emotional aspects of this game, certain conversational locks were put in place that withhold the information the user requires. To overcome this the user is required to “play” with the NPC’s emotional state by being very polite and nice, or else being rude and aggressive. Both emotional extremes open the locks on the dialogue allowing the user to learn about the relevant information and complete the quest.

Additionally, the room features two extra NPCs that sit at a table in the corner. These two NPCs chat to each other, also using the *Alana* system, though the implementation and functionality details of those NPCs are outside the scope of this paper.

The result of the level design provided a themed word that fits well with the dialogue to establish more of a real video game world, rather than a single test area.

#### 3.3 Speech Recognition

The ASR used in this project was Microsoft Windows Speech, also used by the Windows 10 built-in assistant, Cortana, which is Unity compatible.

Speech recognition in this prototype is managed in “turns” of dialogue, meaning that only the user or the NPC may be talking at the one time (no “barging-in” is allowed as yet). Once the user approaches the bartender NPC, the NPC initiates the interaction. To alert the user to whether they can speak or not, a microphone icon is placed in the left corner of the screen, indicating when the system is listening to the player by use of green and red (ASR on and off respectively). In addition to the changing colour of the icon when the system is listening, the text “Recording ...” is displayed and animated to further enforce the notification.

#### 3.4 IBM Watson: sentiment analysis and TTS

Once the user’s utterance is transcribed into a string by the ASR, it is sent to IBM Watson’s cloud service for tone analysis on the text. The user’s utterance is then annotated using the normalised distribution over all possible recognised emotions of the service, which are: *joy, fear, sadness, anger, analytical, confident, tentative*.

The emotions used in this game’s dialogue system were *joy, sadness* and *anger*. This module also handles the response audio synthesis via IBM Watson’s text-to-speech (TTS). After the Dialogue Manager (described in section 3.5) has processed a response it will be sent to IBM Watson’s Text to speech service, then an audio file is return and played though the Unity Game.

#### 3.5 Alana Conversational AI

As stated earlier, this system is using a modified version of the *Alana* system, described in [11, 12]. *Alana* uses an ensemble of bots each producing one or more different responses to the user’s utterance and the prior context, which are then gathered and ranked to return the most contextually and engagingly appropriate response. An entire dialogue system pipeline is being applied in this system (preprocessing, postprocessing, co-reference resolution, etc).

A main element of the system is a “Persona”-bot which is a “*rule-based system implemented in AIML (http://www.alicebot.org/aiml.html) whose main purpose is to maintain personality-based responses consistent across turns*” [12]. The bulk of the Persona bot uses AIML that uses pattern matching to find an appropriate response to whatever the user’s input was. AIML uses the “<Pattern>” tag to match an input and return a response nested in “<template>” tags.

A real in-game character in *The Elder Scrolls V: Skyrim* was used to be the subject of the personality of the bartender in the prototype. This character is named “Hulda” and is “*a Nord innkeeper and the owner of The Bannered Mare in Whiterun*” [15]. The AIML scripts that contain her personality use some information and dialogue lines that the character says in the actual video game, and the remainder of her persona was designed with Wiki articles as points of reference.

Next, we present in detail how sentiment analysis builds the emotional model of the system.

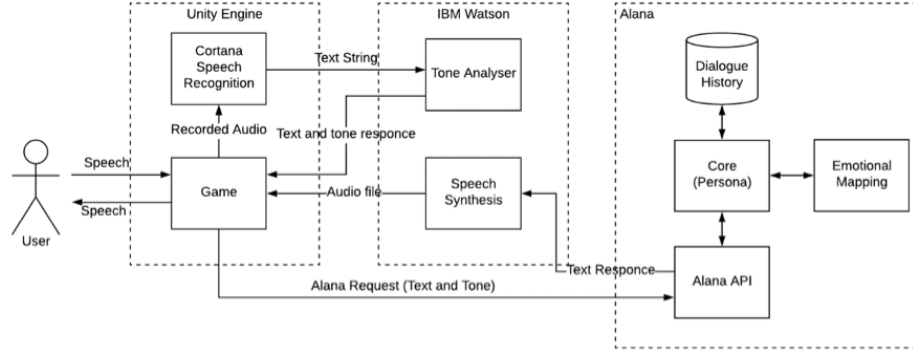


Figure 3: Emotional Conversational System Architecture

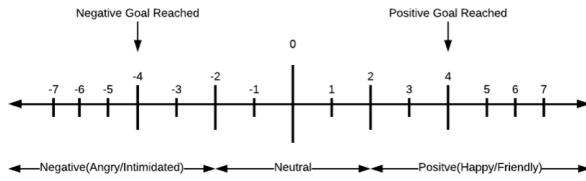


Figure 4: NPCs Emotional Model state representation

#### 4 SENTIMENT ANALYSIS AND EMOTIONAL MODEL

As described in section 3.1, there are several ways of extracting the sentiment of the user utterance. The emotional recognition element in the system is taken from two parts, firstly the sentiment analysis extracted by IBM Watson’s *Tone Analyser*, and secondly by manual assessment (see figure 4) from various inputs in *Persona*’s AIML patterns. Both methods will positively or negatively affect the emotional model of the NPC, currently just a simple scale used to represent how happy (at a score of 1.0) or sad (at a score of -1.0) the NPC is.

The emotional score of the NPC is then used to restrict certain information from being given to the player. The goal information is locked from the player until a certain positive or negative score is satisfied. This simple emotional model can be represented as a scale depicted in figure 4. Story-wise, the NPC barkeeper will only give the information to the player if they are either very nice to them and are friendly, or if the player is rude and intimidates the NPC into giving the information. This is done by a series of conditional statements that check whether the score reaches a value of either 0.4 or -0.4. This value was chosen as it requires the user to actively engage in emotionally manipulative conversation, while also not being too restrictive and drawing the conversation out too long giving the impression that emotion has little effect. If the user asks about the goal information before either of those emotion scores have been satisfied then the NPC will respond in a positive or negative way, depending on their current emotional state, denying the user of that information and asking why they should reveal it.

Depending on the user utterance, the emotional score of the NPC may decrease, increase or have no effect. Once the NPC’s state

is pushed to or past either the positive or negative goal, the NPC will comply with the user’s demand or request for the specific goal information and reveal it.

The two methods of extracting the user’s sentiment are:

##### 4.1 Tone Analyser and Emotion Model

Although IBM Watson’s *Tone Analyser* can recognise a variety of different tones, only three tones were processed by the dialogue manager in this system, these tones were “Anger”, “Sadness” and “Joy”. In addition to these three tones, if none were strong enough there is the default tone of “neutral” which would not have any effect on the outcome of the system.

After processing which tone was found, a simple Emotion Model for the character is used. The system chooses whether the tone will have a positive or negative impact on the NPC’s emotion. Any response that recognised anger will have a negative impact on the NPC’s emotion, reducing the NPC’s emotion score by 0.1. If Joy is recognised, then the system has a positive impact on the NPC’s emotion increasing it by 0.1. If the system recognises sadness, the system applies a positive score to the NPC, increasing it again by 0.1. This is done because the most common occurrences of sadness are being detected when an apology is being made which would make the NPC more friendly.

At different levels of the NPC’s emotion score, the NPC’s attitude changes. Between 0.2 and -0.2 the NPC is considered to have a “natural” attitude. Over 0.2 the NPC takes on a friendlier attitude. If the NPC’s score is less than -0.5 the NPC will be annoyed, and less than -0.7 the NPC will be frightened of the user. To make the conversation more dynamic and realistic a multiplier was added to emotions that acted against the current emotion of the NPC. For example, if the user had been rude to the NPC and forced the NPC’s score to -0.6 making the NPC angry with the user then made an apology to the NPC, the current NPC’s score would be multiplied by two making the NPC’s new score -0.3. The means that changes in attitude on the user’s side effect the NPC’s emotion more, to further prompt emotional manipulation.

The multiplier also works in reverse, meaning that if the NPC’s score was 0.6, then any angry insult or statement would reduce it to 0.3, making the NPC closer to being unfriendly with the player. The feature adds to the realism of the conversation, as no one in a

USER: I hate you! (*anger detected, decreases score, sets NPC emotion score to -0.4*)  
 SYSTEM: Well I don't like you much either!  
 USER: I'm sorry (*sadness detected, increases score and applies modifier, sets NPC emotion score to -0.2*)  
 SYSTEM: It's OK. What's your name?

**Figure 5: Example of the NPC's emotion model**

real conversation is comfortable with being insulted just because they had previously been flattered and been kind to.

## 4.2 Emotion Recognition using Persona

The other method used for recognizing emotion is done as part of the *Persona* module's AIML pattern matching within the Alana system. *Persona*'s rule based nature helps recognising emotion in user utterances that are less focused on tonal connotations than they are in the actual meaning of the utterance in a conversation with this specific character. For example, if the user says "This is a very nice tavern you have here" the system will recognise this as the user being nice to the NPC leading to an increase in the NPC's emotional score. The AIML code pattern matches "\^ NICE \^ TAVERN\^" and then applies an extension call to the emotion script to increase the emotional value by 0.1.

Emotional responses found in this manner are scripted but still can be very usefully in capturing specific occurrences that the user is more likely to say. It is also useful in finding specific things that a specific NPC character might be more emotionally malleable towards. For example, the owner of an inn or a merchant would become offended if a patron claimed that the goods or services that were being provided were too expensive. This method of emotional recognition is done throughout *Persona* and is used to catch both positive and negative emotions. Not all statements alter the emotional state by the same amount, where different approaches alter the NPC's emotional score differently (e.g. attempt to bribe).

After either a positive or negative emotion score threshold has been reached, the NPC will provide the goal information when asked. However, the responses given from a negative attitude are worded differently than the positive response. The answers given from an emotionally positive NPC are larger and provide more information, whereas the answers given from an emotionally negative NPC are shorter and provide less information.

## 5 EVALUATION

Our main hypothesis is that the use of *emotional* conversational AI will increase user engagement with the character, measured by the length of time spent speaking with the character. However, the length of a conversation alone does not necessarily mean a better interaction, because ASR errors as well as several other factors can prolong a conversation needlessly. Therefore we also measured the user's enjoyment in addition to the length of the conversation.

In order to assess the value of emotion detection in speech-driven role-playing video games, we therefore evaluated the proposed system by conducting a between-subject experiment with two conditions: one with the emotion detection features **enabled** and one where they were **disabled**. We tested the system with 16 participants of various ages and of various familiarity with video-games

and voice interfaces. Each participant was given a set of instructions, and was asked to fill in both a pre-experiment as well as a post-experiment questionnaire.

Each participant was to assume the character of an adventurer seeking information about a magic sword, where last he heard the town's barkeeper knows something about it. They were instructed to converse freely with the NPC using a microphone, or a text based input from the keyboard as a fallback method. Once the participant gathered all available information from the NPC (3 pieces of information in total) the "quest" would be marked as completed and end the experiment.

After the task was completed, they were asked to answer a set of questions given in a 4-point Likert scale regarding their interaction: *Q1: I enjoyed playing this game, Q2: I felt like the game was immersive, Q3: I found the game easy to play, Q4: I felt like the Non-Playable-Character reacted to my emotions* along with some general questions about the game like: *What did you like about this game? What did you dislike about this game? How would you improve this game? Do you think speech recognition detection could be valuable in role-playing video-games, and why?*

**Table 1: User Enjoyment: Q1**

	Emotions enabled	Emotions disabled
Strongly disagree	0	0
Disagree	0%	12.5%
Agree	75%	50%
Strongly agree	25%	37.5%

**Table 2: Emotion effect on Immersion: Q2**

	Emotions enabled	Emotions disabled
Strongly disagree	0	0
Disagree	25%	75%
Agree	62.5%	25%
Strongly agree	12.5%	0

It is clear from these results the impact that emotions have in both the user's immersion as well as the user's enjoyment of the game scenario. Players rate the emotional version of the game as more enjoyable and more immersive.

### 5.1 Player engagement

Moreover, we measured the time taken for the participants to successfully complete the task/quest, where we found a statistically significant difference between the conditions with a *p-value* of  $p < 0.05$ . Here we see that players of the emotional version of the game spent longer in the conversation (average time of 9 minutes and 43 seconds versus 5 minutes and 5 seconds for the non-emotional version), while at the same time Table 2, shows that users generally enjoyed the game while in the "emotions-enabled" condition more. This indicates that the emotional version of the game is more engaging, without the additional time spent in the conversation (due to the emotion gating feature of the condition) making the game less enjoyable.

## 6 EXAMPLE CONVERSATIONS

The following are real examples of spoken conversational interaction with different versions of the system. A lot of variation in user inputs is supported by the Alana system [12].

### 6.1 Sample Conversation with Emotion Deactivated

HULDA (BARKEEPER): Hello there, how you doing?  
 USER: I'm fine thanks.  
 HULDA: Good, good. So, tell me, you don't look from around here. Are you from Crydiil or Highrock?  
 USER: I am from Highrock.  
 HULDA: Ah! Knew it, I've got a keen ear for accents, looking for something specific? You'd be surprised how much I know.  
 USER: What do you know about dragonbane?  
 HULDA: Oh Dragonbane aye? That's an ancient sword that was used by The Blades long ago. But as for where to find it, I'm afraid I don't know. [...]  
 USER: Thanks, and who are The Blades?  
 HULDA: The Blades were supposed to be the protectors of the Emperors of Tamriel, but I heard they were also dragon hunters! They have that hideout at Sky Haven Temple I heard.  
 USER: And where is sky haven temple?  
 HULDA: I don't know much about it unfortunately, but I think it was supposed to be in somewhere in and around mountains [...]  
 USER: Thanks, goodbye.

### 6.2 Sample Conversation with Emotion (Positive)

HULDA (BARKEEPER): Hi there, how are you?  
 USER: Hi, I'm fine thanks (*Positive emotional increase*)  
 HULDA: Glad to hear that. What's your name?  
 USER: My name is Jamie  
 HULDA: It's nice meeting you Jamie. Interesting name. So you don't look from around here. Are you from Crydiil or Highrock?  
 USER: Yes I am from Highrock  
 HULDA: Ah! Knew it, I've got a keen ear for accents, So Jamie, Do you want some food or drink?  
 USER: Yes please, can I have some food? (*Positive emotional increase*)  
 USER: Sure, give me a few minutes to heat up the stew.  
 USER: Thanks, do you know anything about the sword dragonbane? (*Positive emotional increase*)  
 HULDA: Oh I might know something about that, perhaps I can be persuaded...  
 USER: Okay then, so tell me about yourself (*Positive emotional increase*)  
 HULDA: My name is Hulda, I'm the owner of this place.  
 USER: This is a nice tavern you have. (*Positive emotional increase*)  
 HULDA: Ah thanks, this place is my pride and joy. You know what, you're alright! What is it you wanted?

USER: I am searching for the sword dragonbane.  
 HULDA: Oh Dragonbane aye? That's an ancient sword that was used by The Blades long ago. But as for where to find it, I'm afraid I don't know. [...]  
 USER: Thanks, Goodbye

## 7 CONCLUSIONS

We presented a new conversational AI system allowing unrestricted user speech input when talking to game characters, based on a state-of-the-art open-domain social conversational system originally developed for the 2017 Amazon Alexa Challenge [12].

In a preliminary evaluation, we have shown that a spoken conversational interface augmented with emotional processing is more engaging (measured in terms of time spent talking with the NPC) for players than one without the emotional dialogue capabilities, while players found that the emotional speech dialogue interface was enjoyable to use.

Future work involves automatically generating different emotional conversational game characters at scale, with dialogue content trained on data such as game wikis. We will also integrate an upgraded version of Alana which has been developed for the 2018 Amazon Alexa Challenge, with better performance and user engagement.

## ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Commission's H2020 programme under grant agreement No. 688147, MuMMER project.

## REFERENCES

- [1] E. Cambria. 2017. *A Practical Guide to Sentiment Analysis (5th ed.)*. Edinburgh: Springer.
- [2] Benjamin Cowley, Darryl Charles, Michaela Black, and Ray J. Hickey. 2008. Toward an understanding of flow in video games. 6 (07 2008), 1–27.
- [3] Mihaly Csikszentmihalyi. 2008. *Flow: the psychology of optimal experience*.
- [4] Arthur M. Freeman. 2003. *Creating Emotion in Games: The Art and Craft of Emotioneering*. New Riders Publishing, Thousand Oaks, CA, USA.
- [5] K. Hamilton. 2017. <https://kotaku.com/fallout-4-s-full-dialogue-mod-makes-the-game-way-better-1750181929>. Retrieved October 29 (2017).
- [6] Ioanna Iacovides, Anna Cox, Richard Kennedy, Paul Cairns, and Charlene Jennett. 2015. Removing the HUD: The impact of non-diegetic game elements and expertise on player involvement. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM, 13–22.
- [7] IBM-tone. 2018. (2017). *Tone Analyzer The science behind the service*. Retrieved April 18 (2018).
- [8] John Laird and Michael VanLent. 2001. Human-level AI's killer application: Interactive computer games. *AI magazine* 22, 2 (2001), 15.
- [9] J. Matulef. 2017. *Analysis: Smooth Talk - The Evolution Of Dialog Trees*. Retrieved October 26.
- [10] Steven Pace et al. [n. d.]. Immersion Flow And The Experiences Of Game Players. ([n. d.]).
- [11] Ioannis Papaioannou, Amanda Cercas Curry, Jose L. Part, Igor Shalymov, Xin-nuo Xu, Yanchao Yu, Ondřej Dušek, Verena Rieser, and Oliver Lemon. 2017. Alana: Social Dialogue using an Ensemble Model and a Ranker trained on User Feedback. In *Proc. AWS re:INVENT*.
- [12] Ioannis Papaioannou, Amanda Cercas Curry, Jose L. Part, Igor Shalymov, Xin-nuo Xu, Yanchao Yu, Ondřej Dušek, Verena Rieser, and Oliver Lemon. 2017. An Ensemble Model with Ranking for Social Dialogue. *CoRR abs/1712.07558* (2017). arXiv:1712.07558 <http://arxiv.org/abs/1712.07558>
- [13] Abdenour El Rhalibi and D. E. 2005. *EXTENDING SOFT MODELS TO GAME DESIGN*. Liverpool: Liverpool John Moores University.
- [14] Laurie N Taylor. 2002. Video games: Perspective, point-of-view, and immersion. (2002).
- [15] Elder Scrolls Wiki. 2018. (2011). *Hulda, Skyrim Wiki*. Retrieved April 12 (2018). <http://elderscrolls.wikia.com/wiki/Hulda>