

# ***Recipe Hunt: Engaging with Cultural Food Knowledge using Multiple Embodied Conversational Agents***

Sabiha Khan<sup>1</sup> ✉, Adriana Camacho<sup>2</sup>, and David Novick<sup>3</sup>

<sup>1</sup>Dept. of Communications, The University of Texas at El Paso, El Paso, TX USA

<sup>2</sup>Dept. of Computer Science, The University of Texas at El Paso, El Paso, TX USA

<sup>3</sup>Dept. of Eng. Educ. and Leadership, The University of Texas at El Paso, El Paso, TX USA  
{skhan2, novick}@utep.edu, caro4874@gmail.com

**Abstract.** The popularity in recent years of food media, particularly in the domain of documentary films, has brought the communicative potential of food to the fore. *Recipe Hunt* is an interactive documentary that simulates the cultural experience of connecting over food by sharing recipes. Embodied conversational agents (ECAs) are used to engage users with cultural food heritage from the U.S.-Mexico border. *Recipe Hunt* aims to use a distributed and participatory model of cross-cultural learning for users to engage with the culinary heritage from this region of the United States.

**Keywords:** Embodied Conversational Agents, cultural heritage, interactive documentary.

## **1 Introduction**

In this paper, we describe an approach to storytelling about the culture of food [1] through a docu-game, implemented through verbal interaction with virtual agents, that teaches users about regional traditions through food. Our approach reflects that of many documentary film about food, which take pains to actively engage the viewer through diverse approaches, including: political narratives of the food-industrial complex, such as *Food, Inc.* [2]; highly aesthetic treatments of culinary traditions, such as *Jiro Dreams of Sushi*; participatory endeavors to understand the agricultural system, such as *King Corn*; and personal narratives of food justice such as *Soul Food Junkies*. As these examples show, many documentary films draw in their viewership by taking otherwise dry subjects and enlivening them with a political, aesthetic, or personal angle that appeals to the emotions [3]. We describe a pilot application, *Recipe Hunt*, and the process of creating naturalistic dialog using our system for the application. We also discuss how we added handling agent-agent interaction and multi-agent-to-human interaction to our system.

*Recipe Hunt* is an interactive documentary about Borderland culture that employs embodied conversational agents (ECAs) to engage users with that culture through the familiar medium of food. The premise of *Recipe Hunt* is that a college student from the U.S.-Mexico border is far away from home and missing the food of her

hometown. She remembers some aspects of each recipe, but not all. She is a bit shy, so she enlists the user, her roommate, to help her roam around the city and ask for help locating the missing ingredients for chile con queso. Luckily, teleportation and time travel are possible in this game world, so the agent and the user potentially have the option of going to foreign countries and into the past as well. The roommate agent trusts that the user will help her, so will provide clues along the way to guide the user.

The underlying conflict in this scenario however is that the user is an outsider to the roommate's culture. That means the user must measure her responses to avoid offending both the roommate and other agents. This feature of the game enables the user to interact with the agents beyond the merely transactional business of hunting down missing ingredients. By making the agents' communication dependent on the politeness and sincerity of the user, *Recipe Hunt* engages the user as the best documentary films do: through both intellect and affect. The emotional engagement of the user with the agents in this way would eventually allow for a more natural interaction with the agents through subsequent stages of game play.

## 2 UTEP AGENT System

*Recipe Hunt* is based on the UTEP AGENT system for creating scripted interactions between embodied conversational agents (ECAs) and humans using XML-style scripting and the Unity game engine [4]. *Recipe Hunt's* design extended the existing system by featuring multiple agents interacting with the user. Its ECAs are fully automated rather than reliant on a Wizard of Oz (WoZ) scenario in which a human controls the agent behind the scenes. As *Recipe Hunt* portrays a specific culture, sustaining multivocality in the agents' representation of the culture in question becomes important. To assuage any potential awkwardness between user and agents, the game starts with a series of brief instructions to orient the user in the proper use of the system.

In this pilot version of *Recipe Hunt*, the agents lead the user through a series of activities and conversations while playing a game. We simulate a roommate relationship, where the user must collaborate with a set of agents to uncover the ingredients and preparation details of a traditional family recipe from the U.S-Mexico border. This simulation is built to be engaging, where both the user and the agent can interact with the same objects in virtual space. The storyline provides the necessary flexibility and decision making, without creating a completely open environment where tasks would otherwise be difficult to set up and evaluate.

The scenario comprises six scenes, each of which lasts approximately one to two minutes, depending on the player's choices and interaction speed. The first scene focuses on the user learning the environment and interaction style. There are no text, heads-up display, buttons, or notifications of any kind, and all interaction occurs through verbal speech [Fig. 1a]. Throughout every scene, each player is offered the same decision-making opportunities and asked the same personal questions. However, scenario choices, animations, and some agent responses depend on each player and can affect the behavior of the agent. *Recipe Hunt* involves agent-to-agent interaction as well as multi-agents-to-human interaction, automated and in real-time [Fig. 1b].



**Fig. 1.** (a) The first scene of *Recipe Hunt*, in the apartment shared between the agent and the user. (b) A scene in which the roommate agent debates with the shopkeeper agent over the ingredients used in a well-known dish from the U.S.-Mexico border.

Each agent is independently controlled through the XML script and can recognize and react to the user's speech accordingly. One of the major limitations is the ability to recognize to whom the user is speaking when there are multiple agents in the scene. Currently we address this problem by having the agent who last spoke to the user be the agent that is spoken to [see, e.g., Fig. 2].

```

<!--Shopkeeper to Roommate-->
<event id="s2a9" agent="Shopkeeper">
  <animation> telling_a_secret </animation>
</event>
<event id="s2e13" agent="Shopkeeper">
  <dialog>Well, yes actually. Sometimes I see people buying Velveeta cheese and
  RoTel tomatoes with chiles.</dialog>
</event>
<event id="s2e14">
  <dialog>It seems like those would go together pretty well.</dialog>
</event>
<!--Roommate to user-->
<event id="s2a10" agent="Roommate">
  <animation>talking</animation>
</event>
<event id="s2e14" agent="Roommate">
  <dialog>mmm...that's not what I expected, but it seems worth a shot.</dialog>
</event>
<!--Roommate to shopkeeper-->
<event id="s2e15" agent="Roommate"><dialog>Thanks</dialog></event>
<!--Roommate to user-->
<event id="s2e16" agent="Roommate"><dialog>Lets keep going.</dialog></event>

```

**Fig. 2.** Example of XML script for the market scene, during which two agents speak both to one another and to the user.

## 2.1 Creating a Naturalistic Dialog

When creating a dialog for the system, our process is like writing a story with a protagonist who has a goal, goes through a conflict, and eventually reaches a resolution. The system works with a declarative XML-style script which is parsed through C# code inside Unity and which in turn controls the 3D models inside a scene (including our agents and the environment) such as how to behave in terms of animation, speech recognition, sounds inside the environment, whether the agent's voice is synthesized

or recorded wav files, etc. In this version of the system, in addition to having information regarding one agent and a few triggers for the environment, we now have to specify to which agent the voice belongs and which agent is activated for a specific animation. Our team first writes a mock-up script without the technicalities of an XML script and performs the script to ensure that expected responses from the user are induced and anticipated. Our aim is to induce the user to respond within a known range of possibilities and to make it unambiguous for users whether the system expects a response from them. One way of doing this is by asking users specific questions but avoiding rhetorical questions, which could lead to interruptions of the agent when the user responds and the agent was expecting an answer.

### 3 Limitations and Future Work

*Recipe Hunt* is currently in its pilot stage and limited to a basic recipe played out over a few scenes. Further work will require developing scripts and scenarios based on extensive oral histories, interviews, and archival research. Future scenes may involve more involved actions, such as walking the user through the actual process of cooking a meal and more culturally specific environments that would place *Recipe Hunt* squarely in the docu-game genre. Beyond transmitting the basics of Borderland culinary heritage, *Recipe Hunt* promises to serve as a food-based, participatory model for cross-cultural engagement both within the field of interactive documentary and beyond to domains such as museum curation and education.

#### Acknowledgments

This material is based upon work supported by a National Science Foundation Graduate Research Fellowship under grant 2015195909. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. We thank Immersion for an updated xml interpreter based on the UTEP AGENT system.

#### References

1. Barthes, R. Toward a Psychosociology of Contemporary Food Consumption. In: *Food and Culture: A Reader* (pp. 23-30). Routledge (2013).
2. Carson, D. Baron, C., and Bernard, M. *Appetites and Anxieties: Food, Film, and the Politics of Representation*. Wayne State University Press (2014).
3. Miles, A. Interactive Documentaries and Affective Ecologies. In: *New Documentary Ecologies: Emerging Platforms, Practices, and Discourses*. Palgrave Macmillan (pp. 67-82) (2014).
4. Novick, D., Gris Sepulveda, I., Rivera, D. A., Camacho, A., Rayon, A., & Gutierrez, M. (2015, November). The UTEP AGENT System. In : *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction* (pp. 383-384). ACM.