

Furhat Culinary Instructor

Group members: Ioana Cotutiu, August Heddini, Sophia Hettich, Siyana Ivanova

Supervisor: Dimosthenis Kontogiorgos

Video link (Vimeo): <https://vimeo.com/499644460>

Video link (YouTube): https://www.youtube.com/watch?v=0NqJQjho_zU

Abstract

This report outlines the details of a project and accompanying pilot study that focused on the effect of using a Furhat Culinary Instructor helper robot when learning a new recipe. The interface consists of a virtual Furhat who delivers a recipe step by step and is interactive, with the ability to repeat steps and answer clarifying questions. Output modalities utilized in the creation of the Culinary Instructor included speech and gestures (facial expressions). The small-scale study included 8 participants divided into two groups, and showed that there is a demonstrable difference in time efficiency and user attitudes when using the Furhat Culinary Instructor in the learning process.

Introduction

During the COVID-19 pandemic, the popularity of cooking at home and learning new recipes (especially baking recipes) has been rising due to the fact that many people are stuck in their homes. Social media is full of pictures of freshly baked bread and of people trying out different cuisines which they had

not considered exploring before. Given this context, we think that newly minted home chefs could benefit from a tool which helps them hone their skills and makes the process of learning to cook easier, even for complete beginners. We believe a robotic culinary instructor could fit into this niche, while also being quite novel to many people, entertaining and engaging to work with: this is why we chose to create the Furhat cooking assistant and evaluate its effect on learning culinary skills and enjoying the cooking process.

The Furhat cooking instructor comes with pre-selected recipes (one example is available in Appendix 2) and does not have the option for the user to add their own recipe due to time limitations. The lists of ingredients were also provided to the users in advance, as a complete, fully autonomous robot was not needed for the sake of answering the research questions, and having the full list of ingredients ready in advance prevented problems in the testing process.

We aimed to provide an easy and seamless learning experience regardless of the user's level of culinary knowledge. The Furhat instructor comes with the

options to ask the robot questions, repeat steps, or recap the ingredient list. Furthermore, we wanted the user to interact with the robot and receive encouragement and cues like facial expressions and nods - details like these make the experience more pleasant and the instructions clearer [1, 6]. In a session which flows without a hitch, Furhat should be able to walk the user through the entire recipe from start to finish - save for the ingredients list, which is viewable in advance, but still available from Furhat at the user's request - and assist with a variety of difficulties that may occur throughout the process.

In this project we aimed to find out whether the introduction of the Furhat culinary instructor improved the experience of learning to cook at home compared to simply following a written recipe. Specifically, we wanted to know how it affected the users' thoughts and feelings about the process, and whether cooking with Furhat was more efficient in terms of perceived effort and time taken to complete the recipe than cooking without it, or both would yield similar results. To put this to the test, we formulated the following set of hypotheses.

H0. There is no quantifiable difference in the learning experience with the help of the robot and without it.

H1. Users will take less time when they have robotic help.

H2. Overall, users will prefer to learn with the robot rather than on their own.

H3. Overall, users will feel more motivated to continue learning when they have received help from the robot.

H4. Users will find working with the robot easier than working on their own.

The null hypothesis is naturally that there is no difference between the two conditions. The rest of the hypothesis statements were crafted to reflect the following questions that aim to qualify a "better" or "worse" experience with learning to cook: Does Furhat help with time efficiency when learning? Do people prefer learning new skills with the help of the robot? Does the experience motivate further learning to a greater extent than the "regular" way of trying a new recipe? Is it easier to learn with robotic help? Based on the feedback we gathered from potential users in the preliminary study stage (see the Preliminary Inquiry section), we came to believe that these questions are likely to be answered in the positive, resulting in the above hypothesis statements. This report details our findings with regard to these statements and the general research question of whether the help of Furhat improves the users' learning experience.

Background

One of our most important concerns when creating the cooking instructor had to do with the interaction between it and the user and the different ways to make that interaction smoother and more natural. With that in mind, we had to consider the literature on facial expressions, gaze, and their effects on human-robot interaction. Zhang et al.

discuss the effects of “eye contact” between users and social robots. The paper found that mutual gaze, i.e. the Furhat robot and the user “looking” at each other, improved the social connection (based largely on where the user’s visual attention was directed) between them [6]. That is a relevant finding to this project because of the incorporation of both gaze and facial expressions to facilitate communication between the instructor and the user: this is reflected in the experimental setup, where Furhat was always placed in a way that allowed and encouraged the participants to look at it throughout the session.

Al Moubayed et al. (2012) also consider facial expressions and how they factor into Furhat’s communication with the user, although that paper has more to do with the physical Furhat robot, which we are not using. Nevertheless, the paper supports the notion that utilizing gaze in the interaction between Furhat and the user affects the experience, particularly when it comes to the perceived gaze affecting turn-taking behaviour on the part of the users. In the 2D condition, where the virtual Furhat is used, as is the case in our project, the subjects took cues from the robot’s gaze to locate the intended target with increased accuracy [1].

Continuing on the topic of user interaction, we must also consider the conversation between the robot and the home cook using it. The 2017 paper by Hastie et al. discusses strategies for building the interface of teaching robots, especially the cadence of their

speech and the required level of courtesy they pay to the user. The paper’s findings point to politeness being key in instructor robots’ effectiveness [4]; this informed our decisions on how to formulate the cooking instructions for optimal results and the best user experience. Also relevant to the discussion of dynamic dialogue with the robot is the 2013 work by Al Moubayed et al. which introduces Furhat. It provided us with insights on its capabilities for multimodal dialog, including facial gestures, lip movements, and intuitive spoken conversation, and underlined the importance of these components to a seamless human-robot interaction [2].

Additionally, we had to consider the inevitability of failure states. It is impossible to predict the full range of situations and conversations that might happen in a session with a user, and therefore we have to make decisions related to handling fail scenarios and the interaction between Furhat and the user in case of an error - to inform those decisions, we referred to the 2020 work by Kontogiorgos et al. relating to people’s perception of errors and failure in robots. The paper found that failures in a human-like robot had less of a negative impact on people’s stated intention to interact with it frequently than failures in a speaker embodiment [5]. The 3D Furhat was used in the study, raising the interesting question of whether failures would be perceived the same way in our 2D cooking instructor, which might be worth exploring in further work.

Finally, it is worth mentioning the multitude of benefits that come with learning to cook - robot or no robot. A 2017 meta-analysis found that on top of the obvious nutritional boons, cooking appears to have a beneficial effect on psychological outcomes [3]. It is our hope that if taken further, our project or a similar one may contribute to people enjoying those physical and mental health benefits.

Preliminary Inquiry

In the planning stage of the project, we performed a quick, cursory preliminary user study in order to get some sense of what people think about and would expect of a robotic cooking instructor, and what their attitudes to learning cooking skills generally are. We interviewed three people of varying ages and levels of skill, asking them three questions we thought were most important to consider before entering the coding stage of the project. The questions are as follows, and the users' full answers can be found in Appendix 1.

Q1. Do you prefer to learn new recipes on your own or have them taught to you by someone else? Why?

Q2. What is the most important feature a cooking instructor robot should have?

Q3. What do you imagine the most annoying thing about a cooking instructor robot would be?

After considering the answers the three participants in the preliminary inquiry gave to these questions, we had an idea of what our potential users would like to see and what they would consider

inappropriate in a product like ours, as well as some insight in how they might view the process of learning to cook itself. The two novice users expressed a desire to be taught cooking rather than try to teach themselves while looking at a recipe, while the advanced user reported not using recipes very often at all. All participants expressed that clarity and flexibility in the robot's instructions would be important to them, and one mentioned that the instructor should not be creepy when it comes to voice and facial expressions. Everyone stressed the importance of being understood by the robot when interacting with it.

Technical Procedure

Overview

The Furhat Culinary Instructor is based on the virtual Furhat interface and written in Kotlin. Due to the time constraints, we adhered to the divide-and-conquer principle and broke down possible responses or user input into multiple categories, each then divided into subcategories, until we achieved a structure similar to a grammar, but with a lower degree of flexibility. More precisely, the resulting structure allowed us to separate keywords essential for different stages of the interaction into files, allowing easy refactoring and maintainability, as well as associating each entry in the file with multiple surface forms to allow greater flexibility in accepted speech input. Even if modern speech recognizers have a good performance and recognition rate for most supported languages, they are still prone to error,

especially when dealing with uncommon words or phrases as well as homonyms. Our choice of structure allowed fine-tuning and improving recognition by specifying which keywords must be present to indicate intent. Finally, to increase the chances of a matching response and to accommodate a larger spectrum of possibilities, we included partial response triggers and handled their received response accordingly.

The robot encapsulates a multitude of states throughout the flow of the application. To coordinate the flow, we used a combination of state transitions and event handling. A special case is the wait state, in which the robot enters a hibernation mode after not receiving any further requests from the user for two minutes or when the user explicitly instructed it that they do not need further help. While in this state, the robot enters a parallel continuous listening state which allows the user to cook and speak with others without being interrupted by the robot.

Gestures are associated with speech throughout the interaction. These include but are not limited to encouraging smiles, brow frowns, head tilts, and sometimes even winking. For improved usability, some states - idle, wait, interaction - are marked in the user interface with various colours. Some of them (green and red) are accompanied by gestures like nods and smiles for expressing agreement or head-shaking and a disapproving expression for disagreement.

Application flow

The general application flow (see Appendix 3) starts in the idle state. When a user appears, Furhat moves to its start state. Next, a recipe can be chosen based on the available alternatives or Furhat's suggestion. The robot will optionally provide the list of ingredients. When requested, Furhat can take short breaks between ingredients (allowing the user to check availability), repeat a certain ingredient, or find alternatives for an ingredient. The user can listen to the instructions step by step and ask for clarifications and repetitions, receiving both verbal and non-verbal (in the form of facial expressions and gestures) feedback. Here, the option to swap out ingredients is also present: a limited number of valid replacements are supported for each recipe. If the user does not interact with Furhat for two minutes or explicitly states that they do not need help, the robot goes into a waiting state. Similar to other "assistants," Furhat can be awoken by keywords less likely to be present in a normal conversation, like "Furhat" or "chef." After waking, Furhat assists the user based on the next request. When finished with the recipe, the robot congratulates the user and goes back to its idle state.

Error handling

In order to handle potential problems with the interaction, Furhat has several error handlers which provide responses in case of misunderstanding or an invalid input, thus avoiding crashes or infinite loops of unmatched responses. If a nonexistent recipe is chosen, the robot will admit to not knowing it and

offer a choice between the ones it already has. If an invalid ingredient swap is requested, the robot will show an irritated facial expression and specify that the change is not possible. If the user lacks an ingredient which cannot be replaced, they are provided with the option of trying another recipe. If it does not understand a word or phrase, Furhat will politely ask the user to repeat what they said.

Physical Furhat

Even though circumstances disallowed including this in the evaluation, the project was designed with the possibility to use the physical Furhat in mind, the interaction and wait colour values were chosen to be appropriate for use during extended periods of time without overheating, as per the official documentation's recommendation. In addition, some constraints and actions are included which do not directly affect the interaction with the virtual robot, but are useful for the physical one: when the user exits Furhat's interaction visual field, the robot remains in the waiting state for ten more minutes, after which it switches to idle; to prevent overheating, the LED strip in the idle and wait states has a timeout.

Method

Experimental Setup

For the evaluation, we recruited 8 volunteers to participate. There were two experimental groups of 4 each, one working without the cooking instructor and one with it. In the first group, participants received not only the ingredients list, but also the entire

written recipe, split up into steps, to follow at their leisure. They had to complete it by themselves without help from an experimenter. In the second group, the participants interacted with the Furhat cooking instructor for the length of the session. They saw the digital Furhat face on a screen and listened to the instructions given by the robot, all the while having the opportunity to interrupt and ask questions. It was important that the digital Furhat is positioned in such a way that one can simultaneously cook and not only have a conversation with it, but also look at it and see its facial expressions. Each participant in the Furhat group had to complete a simple recipe with the help of the culinary instructor. The recipes were chosen so that they are not too long, but have a good chance of being new to the participants, so that they would have to engage with the robot more and potentially ask it clarifying questions. Due to the pandemic, the experiment had to take place in each participant's home environment. We decided to provide access to Furhat through Zoom rather than through installing our interface onto the participants' own computers, because in that way it was less likely that problems with the software setup would occur, and it was easier to fix any errors happening on our side as the sessions went on. In the cases where that was possible and safe, a member of the team was present in person. Before starting each session, we informed each of the participants what ingredients they would need and asked them to lay everything out in advance. In this way, every experimental session

was focused solely on the cooking process. All of the experiments were recorded on video for later analysis.

Evaluation

In order to evaluate cooking with the help of Furhat as opposed to cooking with a written recipe, we took down several metrics. For both groups, we measured the total time it took to make the recipe as well as the time needed for each separate step: in this way, we may operationalize cooking efficiency and compare it between the groups. In the Furhat group, we also noted down each time a participant asked the robot a question or wanted to hear a step again. This was done to determine whether people were taking advantage of the option to be “taught” rather than simply following a recipe as best they can. After the cooking sessions, each participant received a questionnaire with the following items asking them about their experience and how they would rate it. Each item was ranked on a five-point scale from “Strongly disagree” to “Strongly agree.” Some of the items are reversed (i.e. “would not use recipe again” vs. “would use recipe again”) in order to avoid automatic answering and reduce bias.

Q1. I enjoyed this experience.

Q2. I learned new things.

Q3. I would like to continue learning to cook new dishes.

Q4. I did not feel I needed additional help to complete the task.

Q5. I would not make this dish for myself again.

Q6. I found this experience to be difficult.

The short questionnaire is designed to evaluate the participants’ perception of how enjoyable and effective the learning process was in each group, and the answers figured into our overall assessment of whether Furhat increased the efficiency of the process in terms of time and effort. While H1 is solely evaluated based on the time it took participants to complete their task, questions 1, 2 and 5 interrogate H2 by asking the participants about how much they enjoyed the session, whether they felt it was useful in terms of learning, and whether they would like to repeat the recipe, establishing any preference differences between the two groups. Questions 3 and 5 pertain to motivation for future learning (H3), and questions 4 and 6 aim to evaluate the ease or difficulty with which each group went through the process (H4). In the next section, we present the evaluation results with regard to each of the hypotheses.

Evaluation Results

In terms of the time taken to complete the recipe, as well as the time per step, the group who was working with Furhat was faster. In Group 1 (no Furhat), total times ranged between 13 min 21 sec and 17 min 59 sec for the whole recipe (1 min 14 sec and 1 min 45 sec per step). In Group 2 (Furhat), the numbers were lower: 11 min 30 sec to 14 min per recipe (1 min 6 sec to 1 min 18 sec per step). The average time per recipe for Group 1 was 15 min 19 sec and for Group 2 - 12 min 36 sec. While keeping in mind that 8 is a small sample size, and our study is only a pilot, an average difference of 2 min 43

sec between the two groups leads us to believe that there could very well be a significant difference between the two experimental conditions in terms of time efficiency. Therefore, with regard to H1 we conclude that indeed, users take less time to complete the recipes when they have help from the Furhat Culinary Instructor.

When it comes to preference and overall enjoyment, the questionnaire responses showed a somewhat similar picture in both groups. All participants except one from Group 2 answered Q1 with “Agree” or “Strongly agree,” indicating that it is not possible to claim that the experience of learning to cook is more enjoyable or preferable with the Furhat Culinary Instructor than without it - in fact, in both cases it seems to be about equally delightful. When it came to whether participants wanted to repeat the experience (Q5), the responses were more of a mixed bag across groups: some commented that they did not like the result of the recipe, but would like to continue to cook, while others enthusiastically declared their intention to continue learning. In both groups, three of the participants wanted to repeat the recipe in the future while one did not, once again not showing a difference between the experimental conditions. Across groups, there was agreement among the participants that they had learned something new - with the exception of one person in Group 1. However, because of the small sample size, it is difficult to claim that the difference in their response means people felt like more learning occurred in Group 2. To

sum up, in regard to H2, we may not conclude from the data gathered in the pilot study that learning is preferable or more pleasant with Furhat than without. Nevertheless, it is worth noting the difference in the participants' behaviour in the two groups. In Group 1, people were largely neutral throughout the experience and simply went through the steps of the recipe, while in Group 2 they showed a variety of emotional responses ranging from amusement and laughter all the way to anger (in one specific case where Furhat was not able to hear instructions over loud frying sounds). Participants reacted to Furhat's facial expressions, smiling when he smiled and in one case, laughing when he winked. While this didn't seem to translate into more positive answers on the questionnaire, the participants engaged with Furhat, especially when it showed emotion.

As already discussed, it was not possible to draw a conclusion from our pilot study on whether Furhat increased the participants' motivation to continue learning (based on Q3 and Q5). All that can be said on H3 is that their motivation did not seem to decrease in comparison with those who did not use Furhat, and overall almost everyone declared their intention to keep learning to cook and/or to repeatedly cook the dish they learned in the session.

When it came to the perceived level of difficulty of the novel recipe task, the majority of participants deemed it to be quite easy, answering Q6 in the negative (“Disagree” or “Strongly disagree”). One participant who completed different

recipes in both groups found the experience difficult both times but slightly more difficult in the Furhat group - this was the same participant who encountered difficulties with Furhat understanding instructions due to the sounds of cooking on the stove. Once again due to the small sample size, we may not draw a definite conclusion, but it is possible to note the kind of situation which worsens the experience: not being understood, needing to repeat instructions, or becoming confused and/or stressed out mid-session. Furthermore, even though participants seemed to perceive a similar level of difficulty in both groups, those in Group 2 declared on Q4 that they felt they needed additional help, while in Group 1, half the participants said they did not need any help. Interestingly, those in Group 2 commented that they felt they got at least some of the help they needed from the robot, with one person saying they would like “even more help,” while another specified they were more comfortable knowing that Furhat was available to help them and to “ask when you are stuck,” especially as opposed to trying to resolve the confusion themselves by rereading a written recipe.

In summation, at least within the scope of our small study, the null hypothesis has been rejected due to the fact that there was a quantifiable difference in time efficiency both when it came to completing the whole recipe and on the level of singular steps. In addition, participants claimed that they felt Furhat helped them with the process, and using it made a large difference in

their emotional response throughout the sessions. Unfortunately, we cannot draw any certain conclusions regarding preference, motivation, or difficulty; these might be interesting areas for future research to focus on.

Discussion

The biggest weaknesses of the experiment came from the lack of opportunity to meet in person both within the team and when working with the participants. Due to the unusual circumstances caused by the COVID-19 pandemic, collaboration happened through chat apps and Zoom calls. This also led to less control over the experimental environment: participants were trying out our project at home, introducing some risks to the evaluation process, for example the possibility of a bug occurring in the interface; to solve this, Furhat was run on one of our own machines over Zoom whenever it was not possible to be in the same room as the participant.

There were also some minor issues with background noise causing undefined behaviour in Furhat. At the beginning of the project, we expected such issues to occur in relation to different accents and conversational language; however, we had no such problems during the sessions, and instead had a couple of instances of environmental noises confusing the robot, resulting in the participant having to repeat requests. This could potentially be solved by the use of a background-noise-canceling microphone being used instead of the

default microphone that comes with most laptops.

Aside from the technical issues, it is important to note that the small user pool (partially due once again to the difficulty of recruitment at this time) prevents us from drawing any large-scale conclusions from our evaluation. This is why we chose to focus on the users' experience and opinions instead of going for any sort of serious statistical analysis that would be expected in a full-scale experiment/user evaluation. Nevertheless, the results of this small study can pave the way to more extensive research by pointing to some potentially interesting questions, including but not limited to the hypothesis statements which ended up with inconclusive results. Going even further, future research could broaden the focus of the current study by expanding to many different kinds of robot-assisted learning rather than just the acquisition of new culinary skills.

References

1. Al Moubayed, Samer, et al. "Furhat: A Back-Projected Human-Like Robot Head for Multiparty Human-Machine Interaction." *Cognitive Behavioural Systems: COST 2102 International Training School, Dresden, Germany, February 21-26, 2011 Revised Selected Papers*, by Anna Esposito et al., Springer, 2012, pp. 114-130.
2. Al Moubayed, Samer, et al. "Interspeech 2013, 14th Annual Conference of the International Speech Communication Association, August 25-29, 2013." *Interspeech 2013 - Show and Tell*, pp. 747-749.
3. Farmer, Nicole, et al. "Psychosocial Benefits of Cooking Interventions: A Systematic Review." *Health Education & Behavior*, vol. 45, no. 2, 2017, pp. 167-180., doi:10.1177/1090198117736352.
4. Hastie, Helen, et al. "The Interaction Between Voice and Appearance in the Embodiment of a Robot Tutor." *Social Robotics: 9th International Conference, ICSR 2017, Tsukuba, Japan, November 22-24, 2017, Proceedings*, by Shuzhi Sam Ge et al., Springer, 2017, pp. 64-74.
5. Kontogiorgos, Dimosthenis, et al. "Embodiment Effects in Interactions with Failing Robots." *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1-14.
6. Zhang, Yanxia, et al. "Look but Don't Stare: Mutual Gaze Interaction in Social Robots." *Social Robotics: 9th International Conference, ICSR 2017, Tsukuba, Japan, November 22-24, 2017, Proceedings*, by Shuzhi Sam Ge et al., Springer, 2017, pp. 556-566.

Appendix 1. Preliminary User Interviews

User 1. MK

- 25 years old, from Bulgaria
- Works as React developer
- Cooking novice

User 2. DRH

- 52 years old, from Germany
- Currently between jobs
- Digital immigrant
- Very advanced cooking skills

User 3. JH

- 56 years old, from Germany
- Owner of a company for distribution of gastronomy and hygiene items
- Digital immigrant
- Cooking novice/ moderately skilled in some areas

Q1. Do you prefer to learn new recipes on your own or have them taught to you by someone else? Why?

MK: I prefer to have them taught to me because I am a terrible cook and don't particularly like cooking. Having a simple easy to follow explanation makes it less likely I'll set the kitchen on fire.

DRH: I don't really like following along recipes, which is why I prefer cooking over baking. I feel restricted when I need to follow exact instructions and the chance of doing something wrong is a lot higher. I usually cook based on intuition and experience, although I do get inspired by actual recipes.

JH: I am not the most creative person when it comes to cooking and I don't try anything new just for myself, so the

motivation of having a recipe taught by someone else might be nice.

Q2. What is the most important feature a cooking instructor robot should have?

MK: A clear voice and an easy way to go back to a previous statement.

DRH: I find robots quite scary and creepy, so facial expressions and voice would probably be important to me. The right balance between being too mechanical and too human-like. I also would find it important to be able to ask follow-up questions for separate steps.

JH: I need very clear and detailed instructions and the option to switch in-between steps if needed. I would also want to have an overview of the whole recipe, to get a feeling for where I am and what might come next.

Q3. What do you imagine the most annoying thing about a cooking instructor robot would be?

MK: If it misunderstands what I'm saying due to my accent.

DRH: The most annoying thing would be if I would have to repeat myself several times until he understands what I am saying or if he mistakes other conversations as me talking to him.

JH: Getting lost in the recipe would be most annoying to me and probably something I might be worried about with recipes in general. I would need enough time for each step to have the feeling like I am still in control of what is happening.

Appendix 2. Example recipe

Bulgarian Pancakes

Ingredients:

- 2 eggs
- 1.5 cup of flour
- 2 cups of milk
- ½ teaspoon of salt
- 1 teaspoon of sugar
- Oil to fry
- Butter to coat

Procedure:

1. Crack the eggs into a large bowl and whisk until homogenous.
2. Add the flour a little bit at a time, whisking until incorporated before you add more.
3. Add the salt and sugar and whisk them in.
4. Slowly add the milk, making sure no lumps form when you're whisking. The goal is to get a smooth mixture.
5. The resulting liquid should be thin enough to pour out but thick enough to cling to your spoon. If you do not have the desired consistency, adjust with a little more flour or milk.
6. Heat your pan on medium to high heat. Add about a teaspoon of oil.
7. Once the oil is hot, ladle in enough batter to thinly cover the bottom of the pan.
8. Wait until the top of the pancake looks dry.
9. Flip.
10. Wait another minute and take the pancake out onto a plate.
11. Coat the top side with butter.

12. Proceed in the same manner, continuously adding more oil to the pan when it seems to have run dry.

Appendix 3. Furhat Flow Diagram

