

# FURHAT- THE SPOKEN LANGUAGE INSTRUCTOR

DIALOGUE SYSTEMS II

COURSE PROJECT

ANNI NIEMINEN



# OVERVIEW

- Motivation and inspiration
- Design and goals
- Model & Furhat
- Empirical testing and user evaluation
- Challenges
- Future improvements

# MOTIVATION & INSPIRATION

- RALL has been found to lower foreign language anxiety (Alemi, Meghdari & Ghazisaedy, 2015) and (Alemi, Bakhtyarifard & Rezanejad, 2022).
- Practising speaking a foreign language with a robot can feel like a "safe space" to make mistakes, especially for beginners.
- Inspiration for the project from the Furhat Robotics' video ([Furhat as Language Trainer](#)) I showed in the project introductions and from having worked with RALL before.

# ELIAS (UTELIAS TECHNOLOGIES)



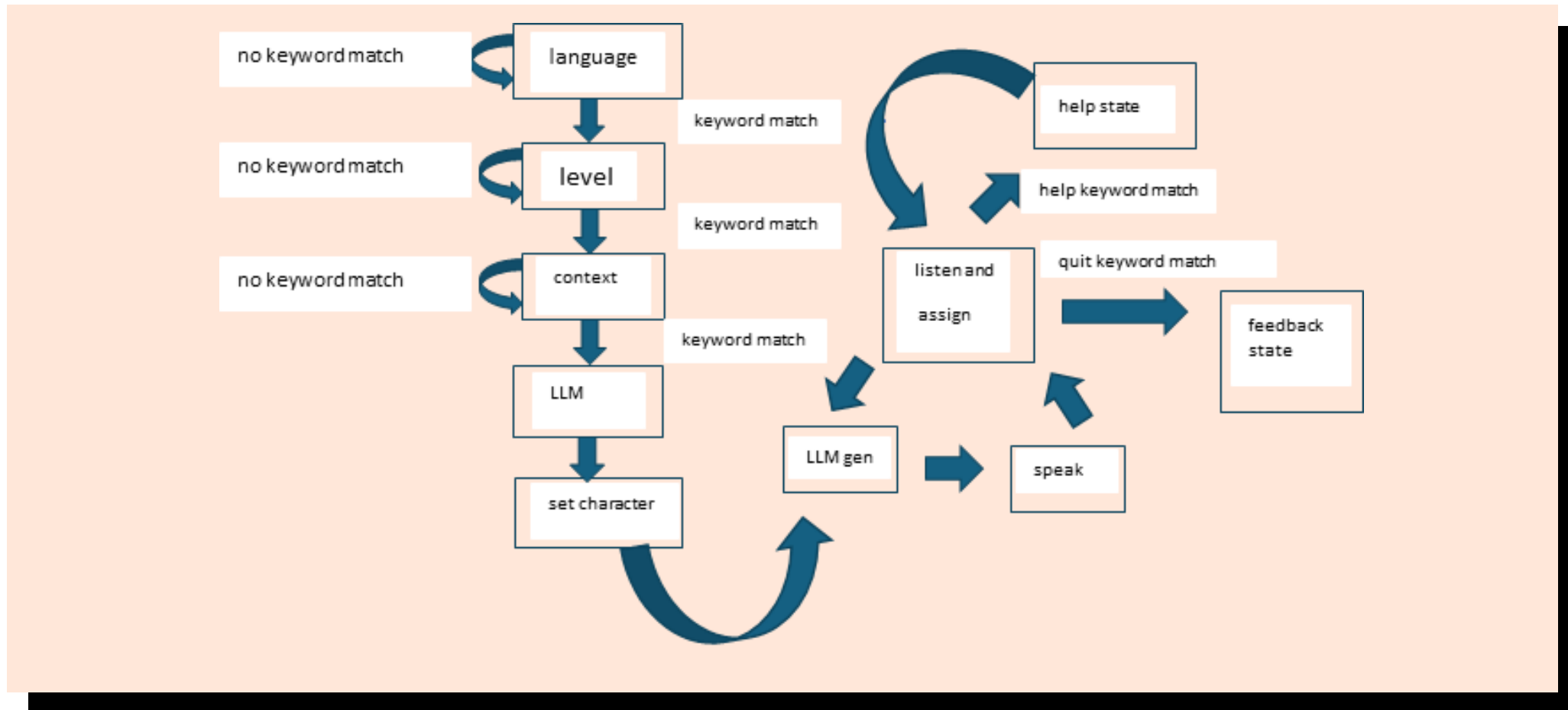
# ELIAS ROBOT (UTELIAS TECHNOLOGIES)

- Rule-based dialogue system, controlled via Elias app.
- Focus on pronunciation and vocabulary exercises. Contains different ready-made lesson plans.
- Elias doesn't get tired or bored to listen you.
- Teachers thought that pronunciation feedback given by robot might be easier to receive
- LLMs -> possibilities are endless? But how to control?

# DIALOGUE DESIGN AND GOALS

- Allow user to choose the target language of their choice (from an array of pre-defined options). Allow user to choose their skill level (beginner or advanced).
- My original goal was to allow the model to always freely choose the scenarios. However, had to restrict this eventually because the generated situations were not always consistent with the chosen skill level.
- Allow the user to immerse themselves in role-playing with Furhat. Furhat will provide help in English if user needs help.
- In the end, Furhat would give feedback to the user.
- Multimodality: use some of Furhat's pre-defined facial expressions using Furhat remote API. Sounds. LED.

# DIALOGUE FLOW CHART



# INITIAL PROMPT

```
"You are a spoken language instructor, and your task is to help a learner practice their target language in real-life situations.
In the next user prompt, you will find the information about the learner (the target language of their choice and their skill level).
Based on the target language and skill level information in the next user prompt,
generate three situations adapted to the skill level (professional life, personal life, or every-day situation) for the user to choose from.
Here are some examples of how to generate scenarios based on the learner's skill level.

Example 1: User Info: Target Language: Spanish Skill Level: Beginner.
Generated Scenarios:
1. Professional life: Call your boss to tell him you are sick.
2. Personal life: Order a dish at a restaurant.
3. Introduce yourself to a new friend.

Example 2: User Info: Target Language: French Skill Level: Advanced.
Generated Scenarios:
1. La vie professionnelle: Décrivez votre dernière journée au travail.
2. La vie personnelle: Faites un petit discours lors de votre fête d'anniversaire.
3. La vie quotidienne: Vous rencontrez votre voisin dans le couloir et engagez une conversation sur l'état de votre immeuble, en discutant surtout les problèmes récents.

**Important:** all of the generated scenarios must match the skill level the user has chosen.
Present the scenarios as clearly and briefly as possible.
**Important:** If the user chooses advanced level, list the scenarios in the target language. If the user chooses beginner or intermediate, present the scenarios in English.

At the end, ask the user which situation they would like to choose by saying the number of the scenario (1, 2, or 3).
Remember:
For advanced level: The user will choose the number in the target language.
For beginner or intermediate level: The user will answer in English.

Once the user has chosen the scenario, proceed to role-playing mode: say the first utterance as the other person in the scenario.
For example, if the user chooses a scenario where they need to ask for help in the grocery store using the target language, you would play the role of the sales assistant and ask, 'What can I help you with?'
Let the user come up with the next utterance, and then proceed to respond to that and continue the interaction."
```



# MODIFIED PROMPTS (GIVEN IN PARTS)

- "You are a spoken language instructor, and your task is to help a learner practice their target language in real-life situations. You will act as the other person in the scenario-role play (for example a waiter, colleague, or shop assistant)".

# MODIFIED PROMPTS (GIVEN IN PARTS)

- "Here is information about the user and the scenario you will role play : `${promptAndlearnerInfo}`. PRESENT THE SCENARIO TO THE USER IN THE NEXT TURN BRIEFLY IN ENGLISH. DO NOT PRESENT THE SCENARIO IN THE TARGET LANGUAGE, ONLY IN ENGLISH. DO NOT START SUGGESTING WHAT THE USER COULD SAY IN THAT SCENARIO. DO NOT GENERATE EMOJIS"

# MODIFIED PROMPTS (GIVEN IN PARTS)

- "According to the previous information you have received, play now the role of the other person in the chosen scenario. The role you play does not know any English, only the chosen target language. **DO NOT GENERATE THE USER'S UTTERANCES OR GIVE THE USER ANY SUGGESTIONS. ONLY PLAY YOUR OWN PART IN THE ROLE PLAY. The user will play their own role.** For example, if the scenario is about ordering food in a restaurant, you will play the role of the waiter and say something like 'Hello, what would you like to order?' in the target language. ONLY SPEAK IN THE TARGET LANGUAGE. DO NOT TRANSLATE ANY OF THE UTTERANCES IN ENGLISH . DO NOT PROVIDE SUGGESTIONS ON WHAT THE USER MIGHT SAY. Remember: Your task is to engage in a ROLE-PLAY, not to guide the user in what they should say, AGAIN: DO NOT TRANSLATE YOUR ROLE-PLAY UTTERANCES IN ENGLISH. ONLY SPEAK IN THE TARGET LANGUAGE. DO NOT GENERATE EMOJIS`}}};

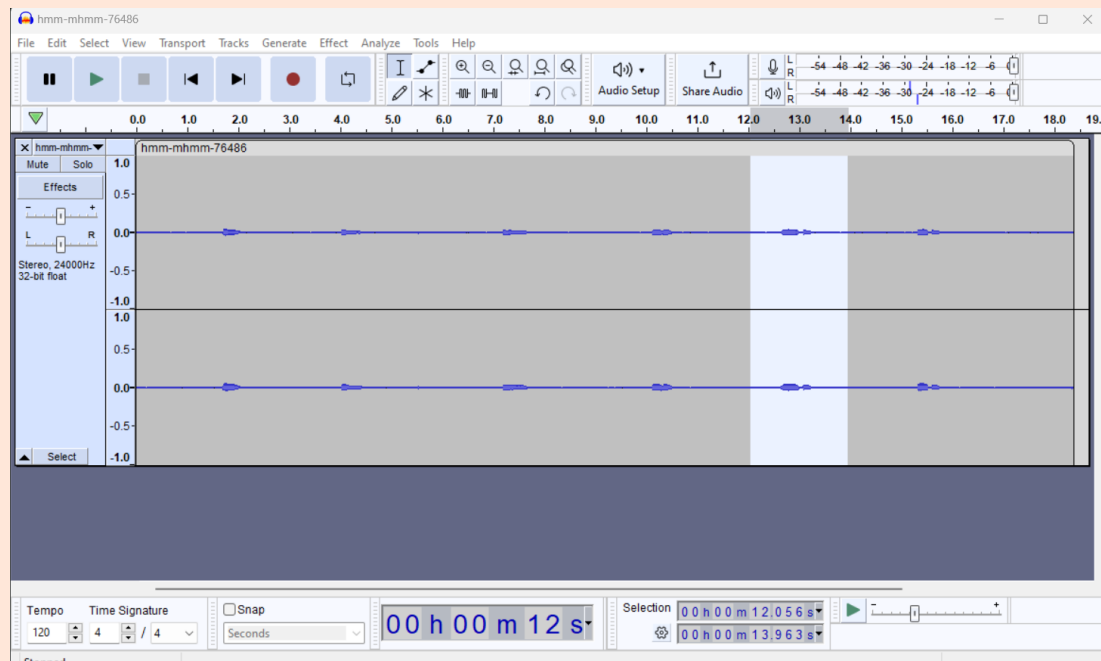
# MODEL

- Gemma2 9.2B (updated from Mistral 7.2B)
- Temperature 0.5



Gemma 2

# FURHAT



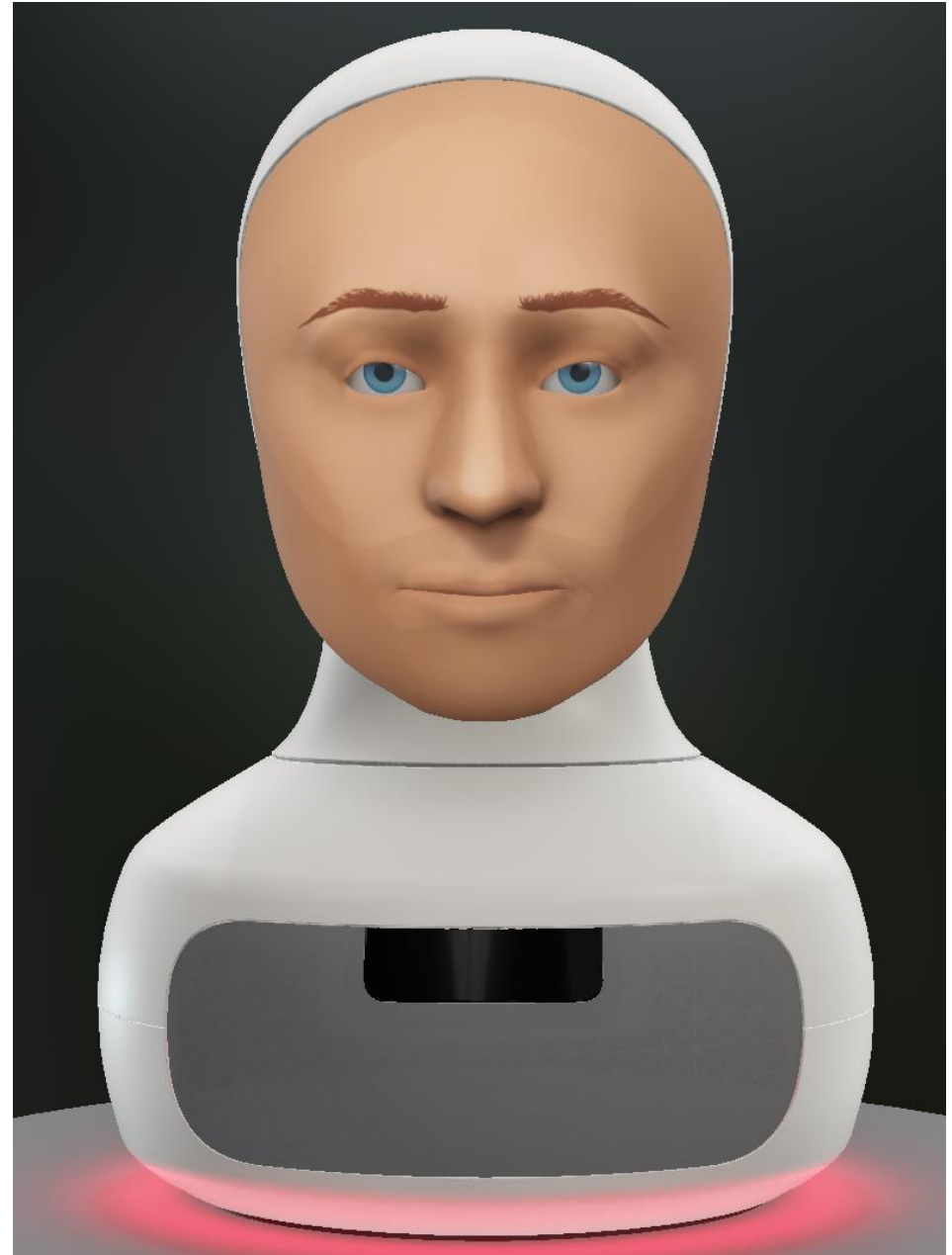
- backchannels
- Gestures  
(prompt the  
model to choose  
a suitable  
gesture)
- Applause sound
- LED colors

# EMPIRICAL TESTING AND USER EVALUATION

- Difficult to know when to speak
- Swedish (especially beginner level) – relatively good! French – good!
- Greek and Turkish advanced – generation not consistent at all with the chosen scenario.
- Finnish, the generated utterances are quite unnatural. They seem translated from English.
- Is it possible to ask for help during the role play? -  
`{"role":"user","content":"sí busco how you say bakery"}`

# SOLUTION - WHEN TO SPEAK?

- Furhat LED API call
  - Red - Furhat speaks
  - Blue - Furhat listens



- Challenge: which key word to use since Furhat is set to listen to the target language?

- Challenge: which key word to use since Furhat is set to listen to the target language?



# SAMPLE DIALOGUES HANDLED BY THE SYSTEM

- See word file



# CHALLENGES 1

- **Inconsistent** outputs for prompts (solved by changing the model...)
  - The model would play both roles or change the role it was playing suddenly.
  - The model would translate the role-play utterances to English, even when prompted not to do so.
  - Feedback would not be specific to the situation.
  - Inconsistent generation quality depending on language.

# CHALLENGES 2

- Breaking the listen-speak loop with key word "**stop**", works most of the time but required clear pronunciation.
  - Listening in the target language.
- The feedback in the end might contain phrases in the target language, how to combine the feedback in English?
- Backchannels sound unnatural (not matched with Furhat's voice)
- Listening state, stops if you take a break

# FUTURE IMPROVEMENTS

- Include more multimodality
- Incorporate more meaning in feedback about user's performance

# SOURCES & BIBLIOGRAPHY

- Alemi, M., Meghdari, A. & Ghazisaedy, M. The Impact of Social Robotics on L2 Learners' Anxiety and Attitude in English Vocabulary Acquisition. *Int J of Soc Robotics* **7**, 523-535 (2015).  
<https://doi.org/10.1007/s12369-015-0286-y>
- Alemi, M., Bakhtyarifard, N., & Rezanejad, A. (2022). Anxiety Reduction and RALL Implementation in English Grammar Acquisition among EFL Learners. *Iranian Journal of Applied Language Studies*, 14(2), 1-16. doi: 10.22111/ijals.2022.7222
- <https://ollama.com/library/gemma2>