

Universitetet i Oslo – Institutt for informatikk

Emne: IN5620

Studenter: daniegb, danieafa, tinenor, ratungla, emmarm

Dato: 14.09.2022

Description of the group

The group consists of Daniel Gloppestad Bajer, Daniel Faour, Tine Nørregaard, Roy André Øvrelid Tungland-Knudsen and Emma Rist Malkomsen. Most of the group got their bachelor's in computer science at UiO, while one got his at NTNU in Trondheim. All five are now in their first year of a master's in Informatics: Design, use and interaction.

Background

In this project we choose to set our focus on certain language processing areas such as Natural Language Processing (NPL) and Second Language Acquisition (SLA) in a digital communication context (Meurers, 2012). We want to take a closer look into AI technology within NPL while observing digital communication to understand how people can use this technology in a SLA context. Therefore the educational aspect of foreign language learning is more interesting to us than language linguistics. Our approach in this project will then be looking at the involvement of users in the design process (Maloney-Krichmar & Preece, 2004). Which is why we want to specifically look at how text processing and AI is used in chatbot technologies.

Chatbot is a form of computer-mediated communications and is essentially a computer program or AI that is capable of carrying intelligent conversations with users in particular domains (Haristiani, 2019). Chatbots are developed for general and educational purposes and have huge potential as a language learning medium. People learning foreign languages say they would like to use a digital language tutoring software more than a real human due to the availability and being a low threshold way to learn a language (Haristiani, 2019).

Position your work relative to existing knowledge and practice:

There has been some research on chatbots for language learning. Fryer, Coniam, Carpenter, & Lăpușneanu (2020) argue that chatbots will eventually become the perfect language-learning partner. Giving us the ability to learn multiple languages anytime, anywhere, and at our own pace.

Haristani's (2019) paper indicates that chatbots have a high potential to be used as a language learning medium. They conducted a descriptive method study, where the data collection they

used were reviews on previous chatbot research. In addition, they developed their own chatbot-based language learning medium: Gengobot.

The results from Haristani's (2019) study showed that language learners are interested in using chatbots. This is because of its flexibility - chatbots can be used anytime and anywhere. By analyzing the data collection and the observation results from their own Gengobot, they found six advantages chatbots can help language learners with. First, students feel more relaxed talking to a computer than humans. Chatbots can also repeat the same material endlessly without getting tired. Third, many chatbots have the functionality for both text and synthesized speech, providing students practice material for both listening and reading skills. The fourth advantage is that chatbots are new and interesting technology. Students also have an opportunity to use different language structures that they normally would not have used and, lastly, chatbots could (if complex enough) provide quick and effective feedback when it comes to spelling and grammar.

Another interesting study about chatbot and language learning is Tu's (2020) research. They observed how chatbots could be used to tutor English for learners in a non-English speaking setting, and aimed to provide contextual conversations to learners. The chatbot from the study actively engaged with the learner by implying the theme of the next statement together with its response. This resulted in chatbot generated responses across multiple contexts, and that it was possible to get chatbots to become more capable of natural conversations.

There was one noticeable limitation with the Tu's (2020) study. The chatbot was easily distracted in specific contexts. This issue happened when vocabulary overlapped, and overly similar speaking patterns happened, in intents.

In this project we want to look into creating our own framework for how we judge chatbots' ability to teach users new languages. Since we are no experts in this matter we choose to base our framework on some of Microsoft's eighteen AI guidelines. These guidelines are sorted into four different categories depending on the phase of user interaction to which they apply (Amershi et al, 2019). The different categories ensure a better user-experience.

The "initial phase" is the first category where the AI system should let the user know what to learn and expect from the system. The second category is "during interaction" which makes sure that the right and appropriate information is displayed during human-system-interaction.

In case of any problems with the system there is a “when the system is wrong” category that involves having plans for potential uncertainties, asking the user for clarifications for instance. The last category, “over time”, helps the user catch up to improvements and changes to the system as an AI system improves over time, in addition to learning from user interactions and providing feedback based on these interactions.

Getting feedback is important during the development of learning (Carless, Salter, Yang & Lam, 2011). One can look at feedback as tutors giving comments on assignments or tasks, but it is considered more helpful if it can be acted upon by the learners. How the learners interpret and use the feedback is an essential part of their learning development. In the study by Carless et al. (2011), they discussed the concept of sustainable feedback and made some principles related to this. They define it as “a process, or activity, in which the feedback can support and inform the learner of the task at hand, but also develop the ability to self-regulate performance on future tasks” (p. 397). We find the concept of sustainable feedback interesting when it comes to language learning and we will incorporate the principles by Carless et al. (2011) as part of our framework.

Questions

Many educational theories argue for how one must practice in order to learn something new. The term “learning by doing” or “[l]earn to do by doing” has been an important term in pedagogy ever since John Dewey first presented it in 1889 (Dewey & McLelland, s. 129-130). Language learning is no exception and we wonder if artificial intelligence can be used by people who are in the process of learning a new language. Can conversations with a chatbot contribute to using learned linguistic elements in a conversation? And what will the experience of this way of practicing be like? One of the main questions we ask is therefore:

How can AI in the form of a chatbot facilitate language learning?

Methods

Throughout the project we have had a focus on communications and group dynamics. Setting a time for weekly meetings, choosing a communications channel and file sharing service,

have been the backbone for the project. Because of the size of the group and differing schedules of the group members, establishing this structure was important for the design process.

In order to approach the research questions stated above, we started by planning and conducting an expert interview with a language pedagogy expert. Additionally we did a review of existing technologies to get an overview over the field and evaluated chatbots that were relevant for our area of research.

This was part of an exploratory approach (*Swedberg, R. (2020))* where we in addition to getting a better understanding of the research field also wanted to figure out the direction we wanted for the project. This led us to slightly changing some of our initial research questions. Initially we had the idea of focusing on *testing* the efficiency of chatbots in language learning, but after looking at existing chatbots and learning about the pedagogy of language learning, we chose to change focus towards developing a framework for *designing* and developing language learning chatbots. Through learning about sustainable feedback our scope turned towards chatbots that could facilitate and share resources, in addition to acting as conversational partners.

Realizing the limited time frame of our project we started a User Centered Design process (Rubin, J., & Chisnell, D, 2008), but were aware that there would not be many iterations, and that there would be a need for further research in order to refine and evaluate the suggested framework. However, based on our initial data obtained through the explorative approach, we conducted a focus group with people that had experience from language learning. This enabled us to get further insight into considerations to be taken into account in the design of chatbots for language learning.

At this point in the project we are able to make a suggestion about how a chatbot should be designed for language learning. The set of guidelines we propose are presented in the Discussion chapter. They are based on our initial data and literature review as well as on the insight gained through conducting the focus group. The task of developing and testing a chatbot based on our suggestions will be left for future work.

Findings

In our initial research we did a review of two existing chatbots for language learning. They were both conversational chatbots, meant for practicing language skills, but one of the bots were Norwegian and the other was English speaking. Some of our take-aways from this review was that the learning outcome of chatbots with too strict of a script was very limiting. Double-barreled questions and closed questions also limited the experience further. On the other hand we found that using a familiar interface (similar to a normal messaging app) made it easier to focus on the language and that chatbots offering a "read-aloud-function" could help with pronunciation practice.

After this initial review we interviewed a language pedagogy expert. The expert we talked to is in charge of a Norwegian language course at a university, meant as an offer for foreign students and staff. The course combines digital resources like videos and a chatbot with physical lectures. In this interview we focused on how the existing chatbot is used in this course, what unrealized potential chatbots may inhabit, and what the expert considered to be the best methods for learning language in general.

The expert told us that the chatbot was mainly used by the students to repeat what they had learned in the course. She also pointed out that the chatbot they use does not give feedback on how the students perform, it only provides examples of the Norwegian language in use. Regarding general language learning, she advocated the term *sustainable feedback*. The main focus in this pedagogical practice is, according to our informant, that it is a two-way process, that the students have to evaluate themselves and that the feedback provides a further path for learning more. This last element of sustainable feedback also points directly at one of the expert's suggestions for taking advantage of the learning potential in chatbots better; namely that chatbots should offer the student appropriate resources for further learning. Additionally the expert pointed out how the use of sound and speech could help students with practicing pronunciation. This coincides with Haristani's (2019) argument that a reason for using chatbots for language learning is that they may provide students with practice material for both listening and reading skills.

Through conducting a focus group with language learners we were able to deepen our understanding of our initial findings. We also became aware of new factors important for

students of a foreign language. One of the participants recommended the service *språklab* at the University of Oslo, because it enabled students to practice speaking and listening to their own voice. Another participant mentioned that she does not like autocorrect services because they don't allow her to practice and learn the spelling herself. She could however see the benefit of a chatbot providing language help and translation. Both of these experiences indicate the importance of feedback, and the active role of the student in its own feedback.

The participants all agreed that surrounding yourself with the language in different forms was important. This was something they all practiced, and they could see some potential benefit of a chatbot since it can provide consistency to language learning. One participant also said that she would like it if the chatbot took an active role and could initiate conversations. That way it might remind students to study when their motivation is low. The participants were, however, mostly skeptical towards chatbots in general. They particularly did not like when chatbots pretended to be humans, and perceived this as dishonesty.

The participants were discouraged from using chatbots because of their experiences with existing commercial chatbots. They emphasized the importance of talking to real human beings as a motivating factor for keeping a conversation. The main insight gained through this is the understanding of the importance of a chatbot clearly and honestly communicating its possibilities and limitations.

Discussion

As introduced in the background section of this assignment, there are several frameworks that are relevant to language learning with AI and chatbots in particular. The guidelines of Amershi et al. (2019) are made to make AI solutions more user friendly, but is not particularly made for chatbots and does not take in the language learning aspect. The framework presented by Carless et al (2011) is on the other hand concerned with the learning aspect, but on a general level, and not language learning in particular. And although technology-assisted learning is at the core of one of the three principles, it does not address chatbots as a learning tool. We therefore wish to extract the most relevant parts of these frameworks and combine them with our findings. The result is a new framework that describes how a language learning chatbot should be designed to assist language learners in the best way.

Even though all the guidelines of Amershi et al. (2019) are probably applicable for all chatbots, including language learning chatbots, we want to highlight the ones that are the most relevant for this purpose. The following guidelines are therefore chosen:

Guideline	**Source**	**Description/argumentation**
Make clear what the system can do	Amershi et al. (2019)	important for the user to know the chatbots´ capabilities when it comes to language learning. This helps with expectation management
Show contextually relevant information		display content relevant to human-system interactions to avoid information clutter during usage
Make clear why the system did what it did		this is to avoid confusion and heresy and give the user a clear learning experience
Remember recent interactions		to give the best learning results it is important that the chatbot remembers what the user already knows/has done. The user will then avoid repetitive learning
Learn from user behavior		same as above, but can also make recommendations or rank certain results higher than others based on past interactions
Scope services when in doubt		in case of any uncertainties, the system will provide alternatives or ask the user to clarify

Facilitate feedback processes	Carless et al. (2011)	stimulate development of capacities in monitoring and evaluating own learning
Enhance ongoing lifelong learning		support development of skills for goal-setting and planning learning
Offer dialogue with the aim of promoting autonomy and reflective interaction		motivate the user to pursue their language goals and become less reliant on direction from the chatbot
Provide further (appropriate) resources	expert interview	the user will have the appropriate tools to develop their language learning
Include speech for pronunciation practice		give the user the opportunity to learn how they speak and pronounce words, in addition to grammar
Encourage user self-evaluation		give the user an opportunity to reflect over their language learning process on their own
Comment on input		ref Haristani: chatbots could provide quick and effective feedback when it comes to spelling and grammar.

Before the last delivery we plan on merging all the guidelines into a new set of guidelines with the help of this table:

AI/technology	Learning	Language learning with chatbots		
Amershi et al. (2019)	Carless et al. (2011)	Expert interview	Focus group	Our guidelines

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 37(4), 445-456.

Carless, Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. Studies in Higher Education (Dorchester-on-Thames), 36(4), 395–407. https://doi.org/10.1080/03075071003642449

Dewey, J. & McLellan, J.A. (1889). *Applied Psychology: An Introduction to the Principles and Practice of Education*. Boston: Educational Publishing Company

Fryer, L., Coniam, D., Carpenter, R., & Lăpușneanu, D. (2020). Bots for language learning now: Current and future directions. *Language Learning & Technology*, *24*(2), 8-22.

Haristiani, N. (2019, November). Artificial Intelligence (AI) chatbot as language learning medium: An inquiry. In *Journal of Physics: Conference Series* (Vol. 1387, No. 1, p. 012020). IOP Publishing.

Meurers, D. (2012). Natural language processing and language learning. *Encyclopedia of applied linguistics*, 4193-4205.

Tu, J. (2020). Learn to speak like a native: AI-powered chatbot simulating natural conversation for language tutoring. In *Journal of Physics: Conference Series* (Vol. 1693, No. 1, p. 012216). IOP Publishing.

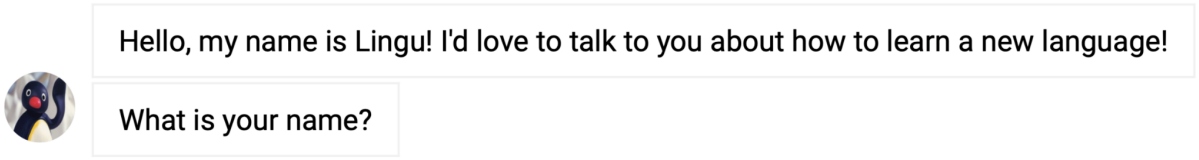
Swedberg, R. (2020). Exploratory Research. In C. Elman, J. Gerring, & J. Mahoney (Eds.), The Production of Knowledge: Enhancing Progress in Social Science (Strategies for Social Inquiry, pp. 17-41). Cambridge: Cambridge University Press.

Rubin, J., & Chisnell, D. (2008). Handbook of usability testing: how to plan, design and conduct effective tests. John Wiley & Sons.

Microsoft (2022). Introduction to guidelines for human-AI interaction. DOI: https://learn.microsoft.com/en-us/ai/opbuildpdf/guidelines-human-ai-interaction/toc.pdf?branch=live

Appendix 1 - Reflection document | Chatbot

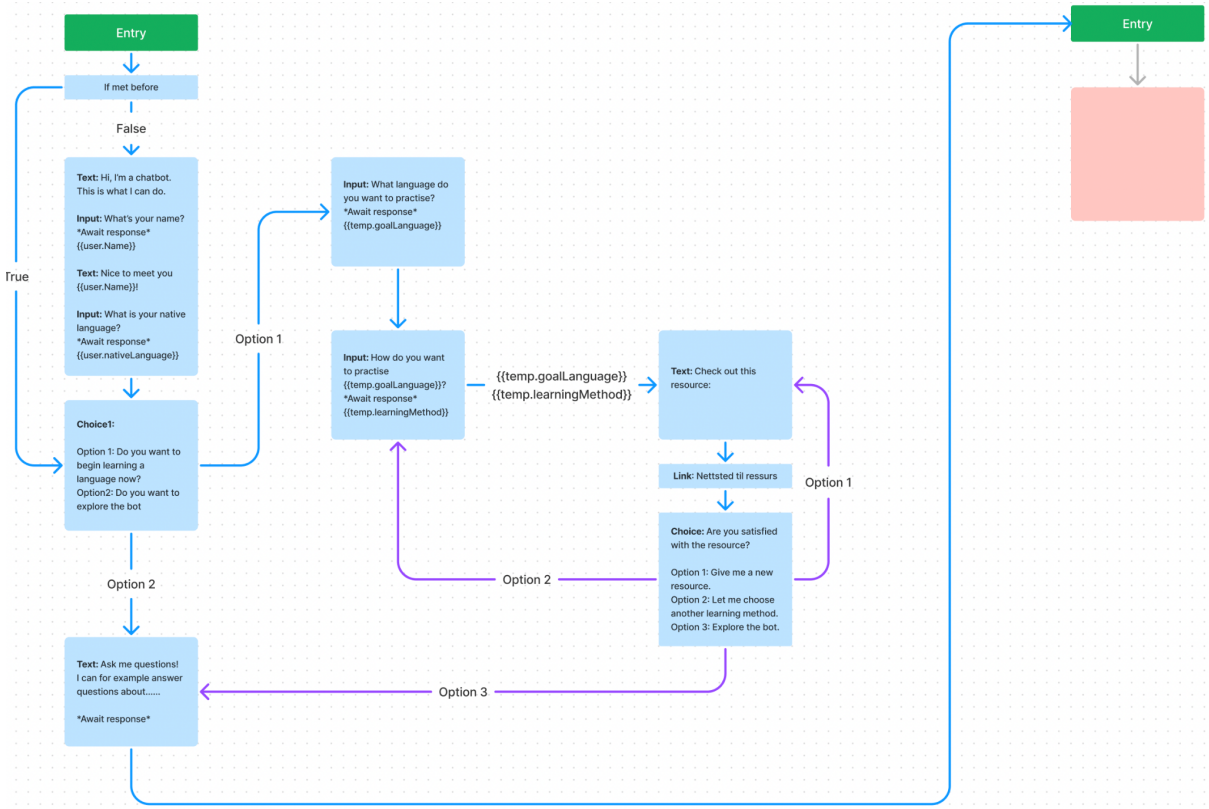
During the second module of the course IN5620, our group were tasked with creating a chatbot with self-chosen functionality. Our project aims to explore if and how AI-infused systems can function as tools for language learning. We therefore decided on creating Lingu, a chatbot that users could talk to about the topic of learning a new language and that can lead the user to appropriate language learning resources.



We sat together as a group and planned out two main functionalities for the chatbot:

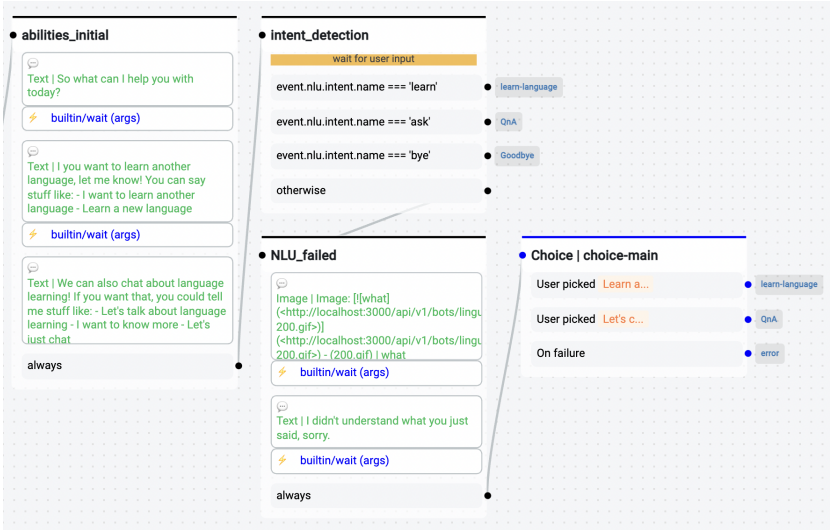
- A question and answer section (QnA) about language learning
- Directing users to appropriate language learning resources

We decided on using the Botpress-system in order to create the chatbot, as it offers some natural language processing (NLP), an easy to use QnA-functionality and flows with programmable logic in order to handle conversational interaction. We then made an initial flowchart using Figma, in order to plan out our intended conversational flows and functionality:



Creation and training of Lingu the chatbot was then started, based on the flowchart. Originally we planned a functionality that would be able to recognize if the user had talked to the chatbot in the past, causing it to skip the introduction. Due to the scope and intent of this prototype, this functionality was scrapped.

We also intended on using button-choices whenever the user had to make a choice, but instead utilized the NLP-functionality in Botpress and programmed the choices as intents, in order to let the user type out whether they wanted to learn a new language or just have a chat. If Lingu can not understand which choice the user is attempting to make, the user will be prompted with buttons instead, making sure they're able to follow the intended flow. A snapshot of the intent-detection implementation can be seen below:



Lingu is programmed to ask the user about their name, main language and which language they're interested in learning (if they want to learn another language). These are stored as variables and used by Lingu in conversation in order to make the conversation feel more natural. Lingu also uses GIFs in order to express itself and seem more natural and fun.



The stored variables for chosen languages are currently not used when deciding which external language resources are presented to the user, but could be used for that purpose in later iterations of the chatbot.

Furthermore, the user is able to choose if they want language resources available on the web or if they prefer to use an app. The user is then provided with a "carousel" of multiple appropriate web services or apps, with links and descriptions.

If the user just wants to have a chat about languages and language learning, we programmed 16 unique questions and answers, each with 8-10 language-variations in order to train the NLP. There's functionality implemented allowing the user to navigate between the language learning functionality and the QnA, until they wish to end the chat.

A showcase of Lingu, its capabilities and how it handles failure-events can be seen on the following link: https://youtu.be/5-PDNPn1CGw